The Impact of an In-Home Co-Located Robotic Coach in Helping People Make Fewer Exercise Mistakes

Nicole Salomons*, Tom Wallenstein*, Debasmita Ghose*, and Brian Scassellati
Yale University

Abstract—Regular exercise provides many mental and physical health benefits. However, when exercises are done incorrectly, it can lead to injuries. Because the COVID-19 pandemic made it challenging to exercise in communal spaces, the growth of virtual fitness programs was accelerated, putting people at risk of sustaining exercise-related injuries as they received little to no feedback on their exercising techniques. Co-located robots could be one potential enhancement to virtual training programs as they can cause higher learning gains, more compliance, and more enjoyment than non-co-located robots.

In this study, we compare the effects of a physically present robot by having a person exercise either with a robot (robot condition) or a video of a robot displayed on a tablet (tablet condition). Participants (N=25) had an exercise system in their homes for two weeks. Participants who exercised with the co-located robot made fewer mistakes than those who exercised with the video-displayed robot. Furthermore, participants in the robot condition reported a higher fitness increase and more motivation to exercise than participants in the tablet condition.

I. INTRODUCTION

An large number of studies have outlined the benefits of exercise [23, 24] including improved cardio-respiratory fitness [21], improved mental health [17], and diabetes prevention [4]. Despite the many benefits, there are risks associated with performing exercises with the wrong body posture or movement patterns [27, 30]. Studies have shown that weight-training-related exercises, when performed without supervision and corrective feedback from a trained professional, can put one at risk of sustaining musculoskeletal injuries [22]. Therefore, many choose to exercise with a personal trainer, who helps them learn correct techniques and tailors their regime according to their body type and strength levels, thus mitigating the risks associated with performing weight-training exercises without supervision [6].

In the COVID-19 pandemic, when access to gyms and personal trainers became limited, many people resorted to exercising at home with limited access to equipment and exercise partners [5, 20]. As the pandemic accelerated the growth of virtual fitness programs, videos created by trainers worldwide gained popularity among a broader audience. When people started learning to exercise with trainers virtually, they risked sustaining exercise-related injuries as they received little to no feedback on their technique while learning new exercises.

One potential enhancement to virtual fitness training systems could be robotic coaches. Social robots have been shown to be effective in providing corrective feedback and motivation while completing tasks [18]. Therefore, we designed a social robot system (Figure 1) that helps people adhere to an exercise routine while providing corrective feedback on their form. We deployed this system in homes over a 14-day study while providing corrective feedback using a machine learning algorithm. We compare our co-located robotic coach (robot condition) to a video of the same robot displayed on a tablet screen (tablet condition). To the best of our knowledge, this is the first in-home study to analyze the effects of a robot’s physical presence in helping people maintain correct exercise techniques.

Our results show that participants in the robot condition made significantly fewer mistakes while exercising than participants in the tablet condition. Additionally, our questionnaire results show that participants in the robot condition reported finding the workouts less difficult and reported a higher fitness increase than tablet condition participants. Finally, participants in the robot condition felt more motivated to exercise and found the system more entertaining.

II. BACKGROUND

A. Use of Robots in Physical Exercise Training

There have been a growing number of studies that show robots as personal coaches [1, 14], including assisting during repetitive, self-directed exercises in rehabilitative therapies [8, 11]. Other studies have deployed robot systems to engage the elderly in physical exercise [7, 9].

Several studies show that a robot can effectively help people learn correct movement patterns and exercise postures by providing real-time corrective feedback to the participant [1, 10]. However, most studies in this domain asked the participants to perform simple, rehabilitative, or injury-preventive exercises without external weights. Furthermore, previous

Fig. 1: Participants completed dumbbell exercises with a robotic coach during a two week in-home study

*Authors contributed equally
robotic coaches were mainly conducted in a controlled lab setting during one session. For robots to be effective coaches, they need to operate in unstructured environments for more extended periods and provide guidance during mainstream fitness exercises. Thus, we explore the impact of an in-home co-located robot in helping people practice weight-training exercises done typically at the gym.

B. Benefits of Physically Present Robots

A robot being physically located with the participant has many advantages. For example, research has shown that a physically present robot led to greater compliance [2] and learning gains [16] than the same robot displayed in a video. Studies have also shown that corrective feedback provided by co-located robots is more effective in helping people learn a given task correctly than video-displayed robots [28]. Therefore, this work studies whether a physically present robot will also benefit the user when acting as an exercise coach. Specifically, we aim to investigate if the feedback provided by a co-located robotic coach would impact the number of exercise mistakes people would make compared to people exercising with videos of a robot on a tablet.

III. METHODOLOGY

In our study, participants interacted with a system in their homes over 14 days, and the participants were expected to do one coaching session every day. Each coaching session took between 15-25 minutes, where they completed five upper or lower body exercises. We compared a co-located robot (Robot Condition) to the video of same robot displayed on a tablet (Tablet Condition). The video-displayed robot was recorded using a high-resolution camera. The physically present robot and the video-displayed robot appeared to be of similar size and performed the same exercise movements and utterances across the two conditions. Figure 2 shows the robots in the two conditions.

B. Exercises

A professional coach helped design the exercise routine, including a good sequence of exercises, the number of repetitions, break times, and common mistakes (Table I). The number of mistakes varied between the different exercises. We classified two mistakes for bicep curls and front raises. We identified one mistake for shoulder presses, single-arm triceps extensions, squats, and lunges. Calf raises and single-leg raises had no possible mistakes because of the simplicity of these exercises. Over a 14-day study period, the system guided participants through upper-body exercises on odd days and lower-body exercises on even days. The participants could choose to use an appropriate dumbbell set, weighing between 2-10 lbs, or perform the exercises without weights.

C. System Design

We used the MyKeepon robot, a 4-DOF, 15cm tall yellow-colored desk robot derived from a commercialized robot called the KeeponPro [15] (Figure 2). The robot can move up and down, side to side, or front and back during the exercise to give the participant the appropriate speed for each repetition. The MyKeepon robot was chosen due to its low cost and small form factor, making it an appropriate robot coach in people’s homes.

To deploy the robotic system in participants’ homes in the Robot Condition, we built a compact 16in x 12in junction box to house, a mini-computer, a router, and support

---

**Hypothesis 1:** Participants will perform fewer exercise mistakes with the co-located robot than with the video-displayed robot.

**Hypothesis 2:** Participants will perceive the co-located robot as smarter and more helpful than the video-displayed robot.

**Hypothesis 3:** Participants will be more motivated to exercise with the co-located robot than with the video-displayed robot.
equipment as shown in Figure 3a. Outside the box, we had the MyKeepon robot, a 12-inch tablet, an Intel RealSense camera [12] to track the participant’s pose while exercising, and an external speaker. The setup for the participants in the Tablet Condition was identical, except that the video of the robot was shown on the tablet instead of the robot being physically present (Figure 3b). Additionally, in the Tablet Condition, the same tablet was used to display the video of the human trainer demonstrating the exercises and the video of the robot exercising. Notably, when the video of the human trainer demonstrating the exercise was displayed, the video-displayed robot was not shown to the participants.

D. Procedure

After delivering the system to the participants’ homes, they filled out a consent form and a pre-study demographics questionnaire. The participants were asked to interact with the system every day for 14 days, with each interaction taking between 15-25 minutes. Each day a given participant experienced the following interaction sequence:

1) The participant would turn on the system and start an application on the tablet.
2) The robot would introduce itself on the first day and explain how the interaction was expected to proceed briefly. Each subsequent day, the robot would begin with a 1-2 minute long motivational greeting.
3) The robot would then guide the person to position themselves at an appropriate distance from the camera with the help of prompts on the tablet.
4) The robot would guide the participant through two sets of exercises, each set containing five exercises. For each of the exercises:
   a) A video of a human trainer performing the exercise was shown on the tablet for 15 seconds for both conditions.
   b) After the participant was prompted to begin exercising, the robot would either move up and down, side to side, or front and back indicating the primary body movement in a given exercise. The robot would instruct the participant to perform the same exercise following the pace of its movements. During this exercise interval, the tablet showed a video of the robot performing the movements in tablet condition, while it displayed a blank screen in the robot condition.
5) The robot would bid goodbye for the day and shut off the system automatically.

E. Mistake Correction

During the interaction, the images captured by the camera were used to track the pose of the participant. We used MoveNet based on TensorFlow.js’s pose detection API. Given a two-dimensional image, we run inference on a pre-trained MoveNet model to predict 17 keypoints on the human body with high accuracy in real-time. These keypoints were used in two ways:

1) Evaluating the participant’s position with respect to the camera: To evaluate if the person’s positioning was valid (i.e., facing the system frontal and about 5m away), we observed whether all 17 keypoints were present in the camera’s field of view with high confidence. If the participant’s position was not valid, they were asked to adjust their position appropriately until it was valid.

2) Evaluating the participant’s form while exercising: Participants were given corrective feedback on their form during most exercises. We designed an algorithm that separated the participants’ movements into repetitions and classified those repetitions according to the pre-defined mistakes in real-time. Given the predicted keypoints, the algorithm used the following steps to identify the appropriate feedback:

   Preprocessing: We normalized the predicted keypoints to reduce the dependence of the analysis on the position and person’s height with respect to the camera. First, we translated each of the keypoints with respect to the center of the body, which we defined to be the middle point of the quadrangle formed by the shoulder and hip keypoints. Then, we divided the translated keypoint positions by the body’s torso length (distance between the shoulder and hip keypoints) to account for different people’s heights. Afterward, we used a Kalman filter [29] to smooth out jitters between predicted keypoints for different frames of the video.

   Repetition Detection: To identify when a person had completed a valid repetition of an exercise, we calculated the minima and maxima for each movement by analyzing in real-time the trajectory of a chosen keypoint whose value changed prominently along the y-axis during a single repetition. We focused on the value of the wrist keypoint for all upper body exercises and the nose keypoint for all lower body exercises. For example, in a correct bicep curl movement, the y-values of the wrist in the trajectory must first strictly increase, then strictly decrease in the y-dimension. Therefore, a bicep-curl repetition is considered valid when a minimum is followed by a maximum and then another minimum.

   Mistake Classification: After a valid repetition has been detected, we use machine learning classifiers to detect if it was performed correctly. We detect mistakes for seven of the ten exercises. For each, we trained a different machine learning classifier. We collected data from twelve people under the supervision of a trainer. We asked each person to perform 10-15 repetitions of each exercise correctly and an additional 10-15 repetitions purposefully performing each of the pre-defined exercise mistakes. We experimented with three different classifiers for each exercise: support vector classifier (SVC) [19], k-nearest neighbor time series classifier (KNN) with k = 5 [26], and a feedforward neural network (FNN) [25]. The FNN consisted of two hidden layers, with 64 and 32 neurons respectively. It used a sliding window approach with a window size of n = 15, and was trained using cross-entropy loss. Since the focus of this paper was not a novel algorithm, we tested different classifiers and chose the one that worked best for each exercise. The chosen classifier for each exercise and its performance for leave-one-subject-out cross-validation on the training data is reported.
in Table I.

**Providing Feedback:** If the participant performed a given exercise correctly for at least three repetitions, the robot would provide a motivating utterance like “Keep going! You are doing well!”. However, if a specific exercise mistake is detected twice or more per set, the robot would provide a corrective utterance to the participant. An example utterance for the arm swinging mistake during biceps curls was “Don’t swing your arms so much! Keep your upper arm attached to the sides of your body”.

**F. Measures**

We collected both behavioral and questionnaire measures.

**Behavioral Measures:** Our behavioral measures included the percentage of days the participant exercised with the robot and the percentage of exercises performed correctly. Three people coded the first two days and the last two days that the system was in each person’s home to measure correct exercise execution. One person coded front raises and shoulder presses; one coded right triceps and left triceps; one coded squats and lunges. Bicep curls were not coded due to the difficulty of detecting swinging due to the low video frame rate. The coders were blind to condition. All three coders coded all exercises done by two participants for the first and the last two days of system deployment to measure the coders’ agreement with each other. The three coders had moderate agreement (Fleiss’ Kappa = 0.44, p < 0.001). The percentage of correct exercises only relates to the six coded exercises.

**Questionnaire Measures:** Our questionnaire measures included a demographics questionnaire asking about age, gender, and exercising habits. We also gave a RoSAS questionnaire [3], assessing the robot’s perceived warmth, competence, and discomfort. A post-experiment survey asked participants to answer the following questions regarding their perceptions of the interaction using a 1–7 Likert responding format: How difficult did you find the workouts? Do you feel an increase in strength and fitness after the last two weeks? How helpful did you view the instructions the robot gave you for doing the exercises? How helpful did you view the exercise corrections the robot gave you while doing the exercises? How important were the workouts in your daily routine? Did the robot motivate you to work out? Do you feel more motivated to continue exercising on a regular basis after the last two weeks? Did you exercise because the robot made you feel guilty if you did not?

**G. Participants**

There were 25 participants in our study. Fourteen participants were in the robot condition: five male, eight female, and one non-binary. Their average age was 21.91 years (SD=2.84). Eleven participants were in the tablet condition: four male and seven female. Their average age was 23.56 years (SD=5.85). There were no significant differences regarding robot familiarity (Robot Condition: M=3.64, SD=1.96; Tablet Condition: M=3.33, SD=1.00; p=.340).

Participants in the robot condition reported exercising on average 4.55 (SD = 1.44) hours a week before the system was in their home, while participants in the tablet condition reported an average of 2.67 (SD = 3.08). These differences were statistically significant (p = .044). 81.82% of participants in the robot condition reported having exercised with weights before, compared to 55.56% of participants in the tablet condition. These differences were not significant (p = 0.202). 27.27% of participants in the robot condition and 22.22% in the tablet condition had exercised with a personal trainer before the start of the study. These differences were not significant (p = 0.80).

<table>
<thead>
<tr>
<th>Day Mistakes Exercise Name</th>
<th>Reps</th>
<th>Mistake Description</th>
<th>Classifier</th>
<th>Validation Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper Body</td>
<td>Bicep Curls</td>
<td>Arm Swinging</td>
<td>12</td>
<td>Arm swings instead of moving only lower arm and keeping upper arm attached to body</td>
</tr>
<tr>
<td></td>
<td>Arm Half-Down</td>
<td></td>
<td></td>
<td>Movement doesn’t cover the full range of the arm’s motion</td>
</tr>
<tr>
<td></td>
<td>Front Raises</td>
<td>Arm Swinging</td>
<td>12</td>
<td>Arm swings to use momentum instead of controlled movement</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Arm Above Shoulder</td>
<td></td>
<td>Weights raised above the shoulder</td>
</tr>
<tr>
<td></td>
<td>Shoulder Presses</td>
<td>Elbows Out</td>
<td>12</td>
<td>Elbows point outward instead of keeping them at shoulder-width</td>
</tr>
<tr>
<td></td>
<td>Single-Arm Tricep Extensions (Right/Left)</td>
<td>Elbows Out</td>
<td>10</td>
<td>Elbows point outward not upward, upper arm far away from head</td>
</tr>
<tr>
<td>Lower Body</td>
<td>Squats</td>
<td>Knees Unstable</td>
<td>15</td>
<td>Knees move around or point inward while going down</td>
</tr>
<tr>
<td></td>
<td>Lunges</td>
<td>Knees Unstable</td>
<td>10</td>
<td>Struggles to keep balance, knee moves around while stepping back</td>
</tr>
<tr>
<td></td>
<td>Calf Raises</td>
<td></td>
<td>12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Single-Leg Raises (Right/Left)</td>
<td></td>
<td>20s</td>
<td></td>
</tr>
</tbody>
</table>

TABLE I: Participants completed upper body exercises on odd days and lower body exercises on even days. Upper body days had several different mistakes that were classified using our computer vision system. Whereas some lower body days had mistakes, and other ones did not. Each exercise had an appropriate number of repetitions that were completed according to the advice of professional coaches.
Fig. 4: Participants in the robot condition on average performed the exercises significantly more correctly than participants in the tablet condition.

IV. RESULTS

A. Behavioral Results

Participants in the robot condition completed on average 68.65% (SD = 11.41%) of the coded exercises correctly, while participants in the tablet condition completed 57.46% (SD = 16.78) of the coded exercises correctly. These differences were statistically significant \( t(25) = 1.98, p = .03 \). On the first two days participants in the robot condition completed more exercises correctly than the tablet condition (Robot: M=71.25%, SD:0.12%; Tablet: M=56.90%, SD:20.54%; \( t(25) = 2.14, p = .022 \)). Participants in the robot condition also completed more exercises correctly during the last two days than the tablet condition (Robot: M=68.11%, SD:15.25%; Tablet: M=52.90%, SD:21.15%; \( t(25) = 2.05, p = .026 \)). These results are presented in Figure 4.

On average, participants in the robot condition exercised 71.74% (SD = 19.79) days out of the days the system was in their home. Participants in the tablet conditions exercised on average 66.46% (SD = 23.48) days. These differences were not significant using a t-test \( t(25) = 0.61, p = 0.273 \).

B. Post-Experiment Questionnaire

Regarding the RoSANS questionnaire, there were no significant differences in warmth (Robot: M = 4.56, SD = 0.98; Tablet: M = 3.82, SD = 1.38; \( t(25) = 1.58, p = 0.064 \)), competence (Robot: M = 4.73, SD = 1.26; Tablet: M = 3.98, SD = 1.36; \( t(25) = 1.41, p = 0.086 \)), or discomfort (Robot: M = 1.71, SD = 0.51; Tablet: M = 2.17, SD = 1.30; \( t(25) = -1.19, p = 0.123 \)).

On the post-experiment questionnaire, participants in the robot condition found the exercises less difficult (Robot: M = 2.43, SD = 1.02; Tablet: M = 3.27, SD = 1.42; \( p=0.048 \)) and felt a larger fitness increase (Robot: M = 4.86, SD = 0.86; Tablet: M = 3.91, SD = 1.51; \( p=0.030 \)) than participants in the tablet condition. There were no significant differences between conditions in how helpful they found the instructions (Robot: M = 5.57, SD = 1.40; Tablet: M = 4.73, SD = 1.79; \( p=0.099 \)) or the corrections (Robot: M = 3.93, SD = 1.90; Tablet: M = 3.64, SD = 1.86; \( p=0.352 \)) given by the robot.

Participants felt more motivated to workout in the robot condition than the tablet conditions (Robot: M = 4.64, SD = 1.65; Tablet: M = 3.27, SD = 2.05; \( p=0.038 \), but there were no significant differences in motivation to continue exercising post-experiment (Robot: M = 5.57, SD = 1.40; Tablet: M = 4.73, SD = 1.79; \( p=0.099 \)). Participants in the robot condition reported seeing the workouts more important in their daily routine (Robot: M = 3.93, SD = 1.21; Tablet: M = 2.55, SD = 1.29; \( p=0.006 \)), and that they felt guiltier when they did not (Robot: M = 5.21, SD = 2.02; Tablet: M = 2.73, SD = 1.49; \( p=0.039 \)).

V. DISCUSSION

A. Hypotheses

Participants in the robot condition made fewer total mistakes than participants in the tablet condition. Furthermore, participants in the robot condition made fewer mistakes in the first two days and the last two days than in the tablet condition. The in-home systems were not deployed for a long-enough duration to see any reductions in the number of mistakes people made in either condition over the study period. Both conditions were consistent in the number of mistakes shown across the two weeks. On the questionnaire, Robot Condition participants reported finding the workouts less difficult and felt a higher strength and fitness increase. These results support Hypothesis 1: Participants performed fewer exercise mistakes with the co-located robot than with the video-displayed robot.

There were no significant differences regarding competence, helpfulness of the instructions, or helpfulness of the exercise corrections. Therefore we do not believe Hypothesis 2 to be true: Participants did not perceive the co-located robot as smarter and more helpful than the video-displayed robot.

There were no significant differences between conditions regarding the percentage of days the participants exercised with the robot. However, participants in the robot condition did report feeling more motivated to work out, placed more importance on exercising with the robot, and felt guiltier when they did not. Therefore we have partial support for Hypothesis 3: Participants felt more motivated to exercise with the co-located robot than the video-displayed robot.

B. Impact of Robot Co-Location on Exercising Mistakes

Having a co-located robot significantly reduced the number of mistakes people made while exercising. On average, participants in the tablet condition performed 43% of their exercises incorrectly. A person exercising with a co-located robot for two days (one upper-body and one lower-body) would have completed 61 fewer incorrect repetitions than a person in the tablet condition. Performing many incorrect repetitions could lead to injuries and sub-optimal strength improvements. This was confirmed by questionnaire results where participants in the tablet condition found the exercises more difficult and felt lower fitness and strength increase than participants in the robot condition.

There are multiple reasons why a person would have performed fewer mistakes with the co-located robot. One possibility is that they felt more engaged and entertained by
the physically present robot and therefore were paying more attention to the exercise demonstrations and corrections. Literature also shows that physically present robots increase learning gains [16], and therefore people might have learned more from the robot’s corrections. Lastly, research shows that co-located robots cause higher amounts of compliance [2]. Thus, participants in the robot condition could have been more willing to receive corrections from the robot.

One potential confound of the study is that participants in the robot condition had more time with the robot. The robot was temporarily not visible during the exercises in the tablet condition. However, we do not believe this significantly impacted the study, as the demonstrations were short, and the robot was mostly static during them. A second potential confound is that the movements of the robot representing repetition speed might have been more visible in 3D than in 2D on the tablet screen. However, this would have a minimal effect as most exercise speed movements were from left to right or up and down, which were equally visible in both conditions. The last potential confound is that participants in the robot condition had reported exercising more hours per week than participants in the tablet condition. They might have had more experience doing exercises and therefore made fewer mistakes. However, participants in the robot condition did not report significantly more experience with weight training nor more experience with personal trainers.

This study shows the great benefits of having a low-cost co-located robot present when exercising. Even if the robot cannot demonstrate the exercises themselves, the presence alone has people making fewer mistakes while exercising. Reducing errors increases exercise gains and reduces the potential for injuries.

ACKNOWLEDGMENT

This work was funded by the National Science Foundation (NSF) under grants No. 1955653, 1928448, 2106690, and 1813651.

REFERENCES


