
TandemTrack: Promoting Consistent Exercise Leveraging Multimodal Training and Tracking

Daniel Smolyak

Dept. of Computer Science
University of Maryland
College Park, MD 20742, USA
dsmolyak@umd.edu

Bongshin Lee

Microsoft Research
Redmond, WA 98052, USA
bongshin@microsoft.com

Eun Kyoung Choe

College of Information Studies
University of Maryland
College Park, MD 20742, USA
choe@umd.edu

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

Copyright held by the owner/author(s).
CHI'18 Extended Abstracts, April 21–26, 2018, Montreal, QC, Canada
ACM 978-1-4503-5621-3/18/04.
<https://doi.org/10.1145/3170427.3188536>

Abstract

As smart speakers such as the Amazon Echo have grown in popularity, these devices have presented new opportunities for exploring voice interaction, particularly in the domain of exercise. We have created TandemTrack, a multimodal system comprised of a mobile app and an Alexa skill whose capabilities include exercise guidance, data capture, feedback, and reminder. We propose using TandemTrack in a deployment study to explore and analyze individuals' preferences and use of voice and visual interactions, and to test TandemTrack's effectiveness as a whole in promoting consistent exercise habits while additionally comparing the effectiveness of the app against the skill. Not only does TandemTrack make important steps in addressing gaps in currently available exercise assistants, but also, more importantly, we believe this is the first study to investigate the specific value that voice interactions may provide over other technologies in promoting consistent exercise.

Author Keywords

Voice interaction; multimodal interaction; smart speaker; self-tracking; exercise assistant.

ACM Classification Keywords

H.5.2 [Information interfaces and presentation (e.g., HCI)]: User Interfaces.

Introduction

Technologies ranging from wearable devices to mobile apps have provided people with the tools to, among other tasks, track exercises, organize workout plans, and receive feedback on their progress towards their health goals. While these technologies are likely sufficient for fulfilling people's needs related to exercise, a sizable percentage of the population that could benefit from these technologies still does not engage with them [10], whether because of concerns over cost or convenience, or a lack of motivation on the part of individuals. In this work, we examine how smart speakers, leveraging their voice interaction capabilities, may help lower the barrier of entry to exercise assisting technologies.

Voice interaction has recently grown dramatically in importance and popularity. Starting with the creation of personal assistants such as Siri, Alexa, and Cortana, smart speakers such as Google Home and Amazon Echo have come into widespread use in the home. While its primary use lies in its ability to play music and set timers & reminders, there have been a large number of *Alexa skills*, the equivalent of apps, developed for the Amazon Echo & Alexa devices, that range from playing trivia games to controlling other "smart" devices in the home [3]. Particularly of interest, many Alexa skills provide people with support for their workouts (e.g., 7-minute Workout [2], 30 Day Pushup Challenge [1]), by guiding them through a series of exercises. (As of January 2018, more than 750 Alexa skills are listed under the "Health & Fitness" category.) In addition, the smart speaker's lack of mobility can be seen as a feature, helping force consistency in exercise location.

However, the smart speaker's main drawback as an exercise support device is in its lack of a visual component, a valuable medium for feedback and self-reflection. Popular mobile apps for exercise, on the other hand, provide a vi-

sual interface while neglecting the voice interactions that the smart speaker may provide. Thus, we propose providing people with both an Alexa skill and a mobile app through *TandemTrack*, to explore if and how voice interaction complements and augments the mobile app in terms of supporting self-tracking and promoting consistent exercise. Additionally, we aim to discover how the voice and visual interactions compare against each other as well as people's preferences and use of the two for several of the capabilities both the Alexa skill and mobile app provide, focusing specifically on four aspects: exercise guidance, data capture, feedback, and reminder. We expect the skill to have greater popularity for exercise guidance and data capture, as voice interaction is likely more convenient during exercise, given its hands free nature, and we suspect that the mobile app will have greater popularity for feedback, as visuals can more easily convey information and detail.

TandemTrack Design Goals

1. Leverage an Integrated, Multimodal Interaction

The Amazon Echo and mobile devices each have their own set of features and capabilities, some overlapping and some unique to each. Because our overarching goal is to leverage the potential synergy that would come from effectively combining these two interaction modalities, TandemTrack was designed to support seamless switch between either modality to best suit people's needs and preferences.

2. Leverage Feedback to Encourage Physical Activity

Through TandemTrack, we would like for people to receive feedback on their progress from the two modalities by tracking data on their level of exercise and consistency of exercise habits. Providing the capability for people to view or hear their exercise history and how their current status compares to their past performance is essential for encouraging physical activity [7].

3. Prescribe a Simple Exercise Regimen

Our third design goal is to provide basic exercise guidance via voice interaction in order to decrease people's burden of researching and deciding on a proper indoor exercise regimen. We do not seek to replace a personal trainer who can provide tailor-made workouts and real-time suggestions; instead, we aim to promote healthy exercise habits in the simplest way possible. Common barriers to exercise include lack of time and inconvenience [10], so it is our goal to decrease these barriers through the choice of a routine that is quick and can be completed in any location.

4. Promote Consistent Exercise Habits

Performing a task every day, especially at the same time and place, leads to habit formation or strengthening of existing habits, where habits are behaviors that become automatic after they have been performed consistently and repeatedly [9]. Habits can help reduce or overcome barriers to any behavior, as their automatic nature decreases the need for conscious effort or deliberation to engage in the behavior. Thus, building a daily habit would be incredibly useful for encouraging people to engage in consistent exercise, leading to healthy behavior change.

TandemTrack

Here we first describe the exercise regimen details and justification. We then explain how we designed and implemented TandemTrack to support the exercise regimen, while achieving the four design goals mentioned above.

Providing Exercise Guidance

Many existing mobile apps provide a wide variety of guided workouts of differing length, intensity, and purpose alongside a number of useful features. For instance, Nike Training Club [6] provides over 100 different workouts, where each workout takes the user through a series of exercises

and is accompanied by videos to show the proper form for each exercise. Existing Alexa skills for workouts also provide various levels of guidance, with some doing as little as to name a random exercise and number of repetitions for the user to complete. However, on the more capable end, the 7-minute workout skill [2] guides the user through 14 sets of exercises, each of 30-second duration. This is done by informing the user of the upcoming exercise prior to the beginning of the next set, starting the timer when it hears the word "ready," and repeating this process for every set.

Our exercise regimen consists of three sets of a chosen exercise to be completed every day, where the chosen daily exercise alternates between sit-ups and push-ups. Each set of the exercise chosen for the day lasts for 30 seconds and one minute rest will take place between each set. While Centers for Disease Control and Prevention (CDC) guidelines recommend at least two days of strength exercise per week [4], we suggest a seven-day per week regimen in order to emphasize the importance of consistent exercise routines, which more effectively create healthy exercise habits.

We specifically chose push-ups, sit-ups, and the timed set format for several reasons. First, these two exercises are convenient to do in most locations, especially in the home, and the daily amount of time necessary to complete the assigned workout is approximately 5 minutes, thus lowering both the time and inconvenience barriers to exercise. Additionally, these two exercises are well-known and commonly found in most general fitness tests, and they rely on two different muscle groups, allowing each muscle group to rest on the day of the other exercise. The timed set format allows people to track their progress through the number of repetitions of the exercise they can complete in the 30 second intervals: the better the person becomes at the exercise, the more repetitions they can complete in the same

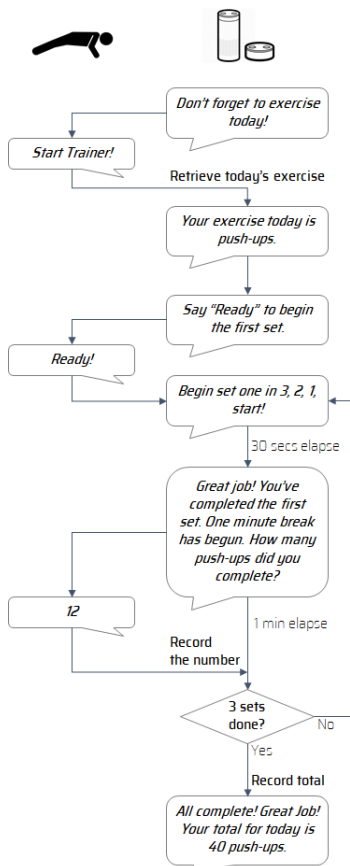


Figure 1: Example daily interaction with the TandemTrack skill.

interval. While myriad other options for similar and possibly better-designed exercise plans exist, this regimen is sure to provide health benefits through strength training to our intended participants, healthy adults, and is suitable for examining the feasibility of voice interaction in realizing the design goals described in the previous section.

TandemTrack's "Trainer" is designed specifically for our exercise regimen. It guides the user through the daily workout by using a timer and occasional audio cues to switch between rest and exercising, and allows the user to start, pause, or reset their workout at any time. Figure 1 shows an example interaction flow with the TandemTrack skill. When it is not convenient for people to access the smart speaker, they can follow the same exercise guidance using a graphical user interface on the mobile app.

Augmenting Data Capture with Voice Interaction

Exercise data can be captured on many levels of specificity, various relevant contexts, and on different ranges of automation. For example, JEFIT (www.jefit.com) provides options to record the number of repetitions completed for each set in a workout, jot down free-form notes, keep track of weight and height, and even take photos of oneself. While the TandemTrack app does not provide as many data capture features as existing apps, it focuses on making data capture as convenient as possible, given that there is exactly one routine that data needs to be captured for. The TandemTrack app collects data by prompting the user to record the number of repetitions they performed in each set immediately after their completion of that set, when the number is still fresh on their mind.

More importantly, the TandemTrack skill allows people to capture data independently of other devices. Some current skills, such as Fitbit [5], rely on the user to link their Alexa

skill with an account created and maintained on another device, and then only provide the capability to query data captured on those other devices. Other skills, such as 7-Minute Workout [2], completely ignore the need for data capture and feedback, focusing only on providing workout guidance. The TandemTrack skill uses voice interaction to capture data as shown in Figure 1: after each set, the skill simply asks the user how many repetitions they completed, thus minimizing the time between the completion of the set and recording of data, while also leveraging the convenience of voice interaction.

Enriching Exercise Feedback

Using the data collected from users, TandemTrack provides feedback in several different ways, each addressing varying time frames and specificity. From the TandemTrack app, the top section of the homescreen serves as a daily summary, either listing the repetitions for the day or directing the user to go to the Trainer, and swiping left across the top of the screen will allow the user to view summaries for previous days (Figure 2). In the middle of the screen are two workout metrics—the number of consecutive days and total days that the user has exercised—to emphasize exercising consistently every day. On the bottom half of the screen are a set of time-based visualizations, that the user can swipe left and right to switch between. The bottom visualization shown in Figure 2 also highlights daily consistency by connecting together contiguous circles where the user exercised on both days: longer periods of daily exercise would form longer chains of circles.

As a complement to the mobile app, the TandemTrack skill provides audio feedback, which can be more convenient than the app in certain situations. Most Alexa skills that do have access to the user's data provide feedback through basic dialogue interactions, such as the user asking "how

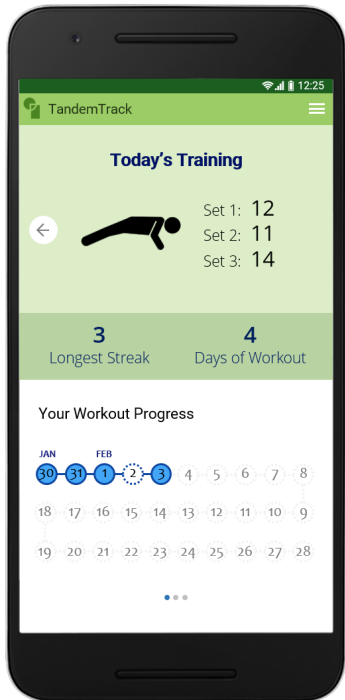


Figure 2: The TandemTrack mobile app's homescreen provides a daily summary, two workout metrics, and time-based visualizations.

many steps I've taken" (e.g., [5]). TandemTrack functions on the same principle, where users can ask a set of simple questions relevant to the data and progress. For example, users may ask, "what was my average number of push-ups for last week?" or "what is my longest streak?"

Facilitating Daily Routines with Reminders

Reminders help to prevent forgetfulness and are essential to building effective habits through daily routines [12]. TandemTrack allows people to set the time of day that they want to exercise: either using voice interaction via the TandemTrack skill or from the settings page in the mobile app. Given that the person may not be present in the right place at the right time to hear the reminder from the smart speaker, TandemTrack reminds people using both modalities.

Implementation Details

The TandemTrack skill is implemented in Node.js through the Alexa Skills Kit and uses Amazon Lambda to upload to the *Amazon Echo Dot*, a smart speaker we use in our study, while the TandemTrack app is written in Kotlin for Android. As a central aspect of our study with TandemTrack is data synchronization between the two devices, the app and the skill share access to a Firebase database. Thus, both devices must have an internet connection in order to send data to the database after collecting repetition numbers from the user during a workout, and likewise receive data after a workout has been completed on another device.

Study Plan

We plan to conduct a 30-day deployment study using TandemTrack and the Amazon Echo Dot. Our goal is to explore if and how voice interaction complements and augments the mobile app rather than how it competes against the mobile app in terms of supporting self-tracking and promoting consistent exercise. Thus, we will use a between-subjects

design with two conditions: the TandemTrack mobile app only versus all of TandemTrack (both app and skill). We will recruit approximately 24 adults who have Android phones and are specifically interested in forming a daily exercise habit (i.e., short indoor strength training).

Our quantitative measure of TandemTrack's success would be focused on consistency of exercise, where longer streaks and fewer lapse days indicate a favorable outcome. More importantly, we will explore user interactions both quantitatively, by analyzing how often participants decide to use either the TandemTrack skill or app (in the all of TandemTrack condition), and qualitatively, by conducting post-study interviews to receive feedback on TandemTrack and asking about their preferences between the two modalities.

Discussion and Future Work

To the best of our knowledge, this is the first study to investigate voice interaction in terms of the value that it might provide over or together with a mobile device. While over 750 Alexa skills relating to health & fitness exist, their advantages over other modalities for promoting behavior change have not been proven yet, with many currently available skills seeming to have been created simply because smart speakers allow it and are increasingly popular. Thus, through this study, we hope not only to compare and explore user preferences between the two modalities, but also to understand how smart speaker and voice interaction can best be leveraged to aid the behavior change process, in particular with promoting consistent exercise habits.

TandemTrack, combining the mobile app and the Alexa skill, fills important gaps in currently available exercise systems. While leveraging a mobile app equipped with the exercise assistant capabilities highlighted in the design section, the TandemTrack skill would be among the first exercise skills

to provide data capture features, let alone providing other exercise assistant capabilities in such a way that promotes our design goals. Existing exercise skills exclusively focus on providing thorough exercise guidance or serving as an alternative for receiving feedback on data collected by other devices, never combining the two.

As research has already suggested a growing trend in users giving a more social role to Alexa in their homes [11], it is also worth exploring how the personification of Alexa and the Echo device affects the motivation of the users to exercise. Social aspects are often key elements of mobile apps that encourage exercise [7], and Alexa, through its voice interaction, could act as an exercise partner in a way that mobile apps would have a hard time replicating. Conversational interaction, which has already been found to have several relevant use cases [8], could also improve the personification and overall quality of the Alexa skill.

Acknowledgment

This work was in part supported by the National Science Foundation under award number CRII:CHS-1753453.

REFERENCES

1. 30 Day Pushup Challenge. <https://www.amazon.com/dp/B074YCDWH9>
2. 7-Minute Workout. <https://www.amazon.com/dp/B018WUNBE6>
3. Amazon.com: Alexa Skills. <https://www.amazon.com/b?node=13727921011>
4. Centers for Disease Control and Prevention: Physical Activity Basics. <https://www.cdc.gov/physicalactivity/basics/adults/index.htm>
5. Fitbit Skill. <https://www.amazon.com/dp/B01CH4BP28>
6. Nike Training App. https://www.nike.com/us/en_us/c/nike-plus/training-app
7. Sunny Consolvo, Katherine Everitt, Ian Smith, and James A Landay. 2006. Design requirements for technologies that encourage physical activity. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 457–466.
8. Rafel Kocielnik and Gary Hsieh. 2017. New Opportunities for Dialogue-based Interaction in Behavior Change Domain. In *Proceedings of Workshop on Talking with Conversational Agents in Collaborative Action Workshop at CSCW '17*. ACM.
9. Phillippa Lally and Benjamin Gardner. 2013. Promoting habit formation. *Health Psychology Review* 7, sup1 (2013), S137–S158.
10. Leib Litman, Zohn Rosen, David Spierer, Sarah Weinberger-Litman, Akiva Goldschein, and Jonathan Robinson. 2015. Mobile exercise apps and increased leisure time exercise activity: a moderated mediation analysis of the role of self-efficacy and barriers. *Journal of medical Internet research* 17, 8 (2015).
11. Amanda Purington, Jessie G Taft, Shruti Sannon, Natalya N Bazarova, and Samuel Hardman Taylor. 2017. Alexa is my new BFF: social roles, user satisfaction, and personification of the amazon echo. In *Proceedings of the SIGCHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 2853–2859.
12. Katarzyna Stawarz, Anna L Cox, and Ann Blandford. 2014. Don't forget your pill!: designing effective medication reminder apps that support users' daily routines. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2269–2278.