# The Positive Impact of Push vs Pull Progress Feedback: A 6-week Activity Tracking Study in the Wild

JESSICA R. CAUCHARD, Ben Gurion University of the Negev, Israel
JEREMY FREY, Ullo, France
OCTAVIA ZAHRT, Stanford University, USA
KRISTER JOHNSON, Stanford University, USA
ALIA CRUM, Stanford University, USA
JAMES A. LANDAY, Stanford University, USA

Lack of physical activity has been shown to increase disease and reduce life expectancy. In response, mobile devices are increasingly being used to support people's health and fitness by tracking physical activity. Prior work shows that the type of feedback, either ambient or via notification, affects users' behavior towards their physical activity. Yet, these phone- and watch-based interactions and notifications have primarily been visual in nature. Inspired by prior research, we explored the impact of feedback modality (visual, tactile, and hybrid: visual/tactile) on 44 participants' behavior and exercise mindset in a 6-week field study. We present the differences between modalities and the notion of push vs. pull for interface feedback and notifications. Across 1,662 days of study data, we found statistically significant impacts of feedback modality and, in particular, the positive effects of push feedback on participants' mindset about the process of exercise. Our results also highlight design guidelines for wearables and multimodal notification systems.

CCS Concepts: • **Human-centered computing → Empirical studies in ubiquitous and mobile computing**; *Field studies*; *Interaction paradigms*; Haptic devices;

Additional Key Words and Phrases: Activity tracker, Step counter, Personal informatics, Vibrotactile display, Vibrations, Multimodal, Wearables, Behavior change, Health, Fitness, Quantified self-tracking devices.

**ACM Reference Format:**
Jessica R. Cauchard, Jeremy Frey, Octavia Zahrt, Krister Johnson, Alia Crum, and James A. Landay. 2019. The Positive Impact of Push vs Pull Progress Feedback: A 6-week Activity Tracking Study in the Wild. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 3, Article 76 (September 2019), 23 pages. https://doi.org/10.1145/3351234



## 1 INTRODUCTION

Lack of physical activity has been shown to increase disease and reduce life expectancy [39]. In a *Nature* paper from 2017, Althoff et al. [2] leverage usage of smartphones to provide a large-scale measurement of physical activity across 111 countries. They show inequalities in activity distribution across and within countries, and that it has become a global public health issue. Tremendous research and public health efforts over the past

Authors' addresses: Jessica R. Cauchard, jcauchard@acm.org, Ben Gurion University of the Negev, Be'er Sheva, Israel; Jeremy Frey, jfrey@ullo.fr, Ullo, Bordeaux, France; Octavia Zahrt, zahrt@stanford.edu, Stanford University, Stanford, California, USA; Krister Johnson, krister.h.johnson@gmail.com, Stanford University, Stanford, California, USA; Alia Crum, crum@stanford.edu, Stanford University, Stanford, California, USA; James A. Landay, landay@stanford.edu, Stanford University, Stanford, California, USA.


2474-9567/2019/9-ART76 $15.00
https://doi.org/10.1145/3351234

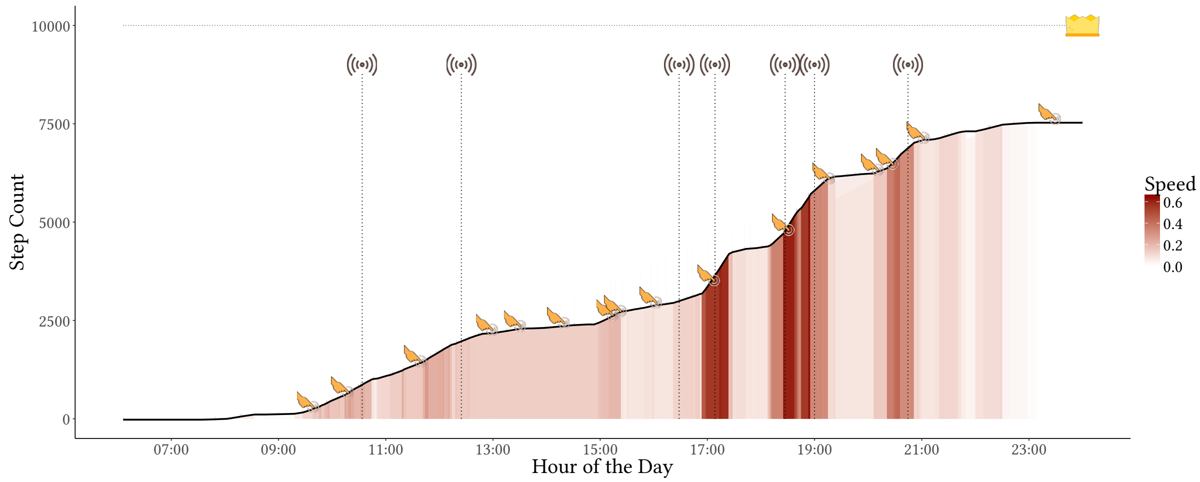



Fig. 1. Step Counts, Vibrations, and Button Presses: a day in the life of a participant. This figure highlights one day of activity from Participant P18 (Hybrid: Visual/Tactile condition). The curve shows the step count activity progress up to 7,549 steps and the color corresponds to the speed (in steps per second). The seven vibrations correspond to 70% of the goal being reached. The yellow hands correspond to button presses on the watch.

decades have aimed to educate populations about the importance of physical activity for health using public health campaigns and guidelines, such as [52]. These health promotion programs have targeted individuals' intentions to meet recommended levels of physical activity. In addition, research shows that there is an important intention-behavior gap. In fact, intentions alone explain only an average 28% of variance in future behavior [56]. Yet, more recent research suggests that one important driver of behavior is the degree to which the behavior is viewed as a positive, approach-worthy process and experience, regardless of how beneficial the outcome [6]. Additionally, research from the psychological and medical sciences is accumulating to suggest that health and longevity outcomes depend not only on actual health-related behavior (e.g., number of steps in a day), but also on individuals' perceptions (or mindsets) about their health and behavior [15, 16, 18, 33, 66]. In this work, we suggest that individuals' mindsets about the process of exercising are often overlooked in the design of wearable trackers to focus solely on encouraging increased engagement in exercise.

Researchers in the HCI and Ubiquitous Computing communities have been working in the fields of personal informatics and persuasive technology [26] to develop systems that can support people in tracking and monitoring their physical activity [13]. This has influenced the emergence of commercial mobile and wearable devices specifically designed to support health and fitness tracking. Across the world people own and use smartphones and, increasingly, wearable technology, which as a whole is projected to generate US$95.3 billion in revenue and 560 million in shipments by 2021 [35]. Their increased adoption is correlated to wearable technologies that are constantly improving in sensing accuracy, computing power, battery life, design, interaction, and functionality [55]. One main functionality of wearable technologies is health and fitness monitoring. Several studies have shown that wearing a step counter can significantly increase a person's physical activity and improve health (i.e., decrease body mass index and blood pressure) [7]. Prior work also shows that adequate monitoring and proper notification systems using mobile devices [11] may lead to behavior change and better habits.

While activity trackers are increasingly being used, research findings on their effect on individuals' health and behavior perceptions is often overlooked to focus on exercise behavior. In addition, activity trackers often experience high levels of abandonment [10]. Researchers have identified specific barriers as to why people do not continue using these devices over time, including: tracking accuracy, reward system, social comparisons, and aesthetics and form [32, 38]. Prior work also shows that without "motivational affordances, informational affordances do not sustain long-term use of the device" [34].

This research work focuses on the human-computer interaction aspects of the wearable, and in particular on how the "reward system" and the way the data is conveyed to the user affects them. Indeed, prior work shows that the type of feedback, either ambient or via notification, affects users' physical activity behavior [11]. Yet, these phone- and watch-based interactions and notifications have primarily been visual in nature. Inspired by Cauchard et al.'s ActiVibe experiment [9], we explore the impact of notification modality (visual, tactile, and hybrid: visual/tactile) in a 6-week longitudinal study with 42 participants. This research highlights how the activity tracking feedback modality can lead to differences in users' behavior and mindset regarding their activity. Across 1,662 days of study data, we found a statistically significant impact on the participants' exercise mindset when the progress feedback is *Pushed* (tactile conditions) compared to *Pulled* (visual condition).

This work makes the following contributions to the field:

(1) We found an effect on behavior and mindset when the information is *Pushed* to the user compared to them having to *Pull* it.
(2) The modality of feedback (visual or tactile) of an activity tracker has an effect on the user's behavior and mindset.
(3) Vibrations have a positive impact on users and can be considered a part of the reward system.
(4) To the best of our knowledge, this is the first research differentiating visual and tactile feedback for activity trackers.

The paper is structured as follows. We first present prior work, the prototype design, and the study design. We then discuss our results and findings on how the differences in feedback modality affect people's activities, habits, behavior, and mindset. The paper concludes with design guidelines for wearable fitness trackers.

## 2 BACKGROUND

This section reviews prior work on activity trackers and behavior change, user interfaces and feedback modalities for mobile and wearable trackers, and people's mindsets about exercise.

### 2.1 Activity Tracking Applications

Many mobile phone applications and trackers have been developed and researched over the years to increase a person's physical activity and support behavior change. Consolvo et. al [13] present a large literature survey of these applications. We find that most applications provide immediate feedback about one's progress using concrete visual representations of the logged data.

Prior research investigated how interfaces can be designed to best support users. Habito [29] proposed a 3-section design strategy that displays goals, contextualizes physical activity, and provides textual feedback. de Vries et al. [17] proposed different motivational messages fit for different stages of support to the behavior change and Duro [20] suggested that the motivational message can have an affective impact on the user. Kocielnik et al. [37] raised that users can struggle to notice and understand their activity data, and as such propose a mobile conversational agent to support reflection.

Alternative user interfaces have included graphical representations of activity via a garden [11, 14], a fish tank [44], and informative art [23]. The main characteristic and success of the garden representations [11, 14] is its ambient display on the mobile phone's background that provides glanceable reminders of one's activity level every time the phone is used. While these interfaces have been researched over the years, the compact shape of wearable devices introduces new challenges in terms of designing glanceable and meaningful interfaces adapted to these new form factors.

### 2.2 Wearable Activity Trackers

Most commercial wearable fitness trackers present real-time data on a small-sized display such as numbers, dots of colors, and graphics such as a flower[1]. Gouveia et al. [30] show that users spare 2.8 to 4.9 seconds on average glancing at their wearable device to check their activity level and summarize it as: "You have 5 seconds" [31]. Prior work proposes to visualize activity logs as a glowing light on the wrist [8] or on the foot [43], or by digitally changing light patterns on a wearable bracelet [3, 40]. Recently, interactive clothes have been suggested as a way to reflect one's exercise levels [28]. Jarrahi et al. [34] suggest that the material features and physical affordances of a wearable device play a role in its long-term use for behavior change. We build upon this assumption in observing how the modality of interaction (visual, tactile, or hybrid) affects activity tracker usage. While some commercial activity trackers have started using vibrations to indicate reaching one's goal [19, 46], to the best of our knowledge, none of the interfaces previously studied the impact of tactile feedback in activity tracking.

### 2.3 Push/Pull Mechanism

Push feedback, such as notifications, alerts, and nudges, is used to set reminders, draw people's attention to an event, or engage them with specific content. In a 10-year survey on activity tracking and behavior change, Okeke et al. [50] found that "digital nudges" were used in 21 out of 26 research works to prompt users to take an action. Yet, we argue that Push feedback can be also be used as a way to raise the user's awareness towards their progress and change their mindset about the process of exercising. When an activity log is visual, it is available for the user to look at when they wish (*Pull*). However, when the information is in the audio or tactile modalities, it is presented to the user whether they want to attend to it or not (*Push*). Meyer et al. [46] report that direct vibration feedback on achieving a goal was effective, which we argue is due to the *Push* mechanism that is a form of notification in itself. In this work, we compare the two mechanisms where information is either sent through *Push* feedback in the tactile modality or can be accessed by pressing a button (*Pull* in either visual or tactile modalities).

### 2.4 Mindsets about Exercise and Their Impact on Behavior and Health

Despite the presence of guidelines reminding people of the importance of physical activity, most (78%) of the US population falls short of physical activity guidelines [48]. Recent psychological research suggests that this is in part due to the mindsets people have about exercise. First, people hold mindsets about the adequacy and health benefits of their level of physical activity [66, 67]. Even though it is widely known that sufficient exercise is important for health, it is harder for individuals to evaluate their own physical activity level. Therefore, individuals look to external standards such as wearable activity trackers to form the mindset that their activity level is beneficial to their health or not. For example, if individuals' pedometer consistently displays a step count that falls short of their goal (e.g., 10,000 steps a day), they may adopt the mindset that their physical activity level is inadequate to their health. Initial research suggests that this negative mindset reduces people's self-efficacy, engagement in exercise, and perceived health [67], and is even associated with higher mortality risk [66].

Second, recent research demonstrates the importance of people's mindsets about the process of engaging in exercise [6]. Even though most people are aware of the long-term benefits, they frequently view the process of exercising as "cringe-worthy" (e.g., hard work, unpleasant, or boring). For these people, engaging in exercise is not inherently rewarding but instead requires considerable self-control and can lead to depletion [4]. Thus, people who view exercising as "cringe-worthy" are less likely to initiate and sustain an exercise routine. In contrast, when people view the process of exercising as "crave-worthy" (e.g., enjoyable, easy, and fun), exercising becomes an inherently rewarding behavior. People are intrinsically motivated to engage in rewarding behaviors [5], which in turn improves performance, persistence, and general well-being [54]. Several experiments show that step

[1]For example Misfit Shine, http://www.misfit.com and Fitbit One, http://www.fitbit.com

count tracking can lead users to experience walking more like hard work rather than a fun activity. This mindset that exercise is cringe-worthy in turn reduces enjoyment of walking and undermines general well-being [22]. Importantly, the pedometer used in this study provides users with a simple visual display of the number of steps, available for the user to look at when they wished (*Pull*).

Here we examine how feedback modalities (*Push* vs. *Pull*) can buffer individuals from the negative effect of tracking on mindset. We reason that *Push* feedback might be more adaptive because it frequently highlights positive progress towards the goal even while users are engaging in tasks unrelated to the goal of exercising. Users' increased awareness of their progress may induce a more adaptive exercise benefits mindset, which can in turn increase engagement in exercise and wellbeing. Additionally, *Push* feedback about goal progress signals that exercising can be easy, convenient, and enjoyable, as exercise can be accomplished without extra effort while pursuing everyday activities. The resulting positive exercise process mindset can in turn promote motivation and sustained engagement in exercise. This positive feedback may be especially useful for people who are just starting out pursuing a goal as it increases their commitment [25]. The following section defines the hardware choices and interface design.

## 3 HARDWARE AND INTERFACE DESIGN

Our interface was developed for a smartwatch with accelerometer sensor for tracking step counts and activities. The smartwatch used in this study was a Pebble Classic (52 x 36 x 11.5mm, 38g), which is robust and easy to wear. The research team lent each participant a smartwatch for the duration of the study. Two apps were installed on the participant's iPhone, the Pebble app and the study app. When the watch and the phone were connected by Bluetooth (BLE), the study app on the phone received the watch data and sent it to a server for monitoring. The participants interacted exclusively with the watch and only had to interact with the phone Pebble app in case of connection issues. All watch notifications were disabled for the duration of the study.

In this study, we focused on step count activity since walking is a cheap and easy form of exercise that is accessible to most. Step count tracking is accurate and pedometers have been shown to be a good motivation for physical activity and behavior change [49, 61]. We designed variations of the same interface (Figure 2) for each condition: Visual Only, Tactile Only, and Hybrid: Visual/Tactile. They were designed to display a person's current step count and percentage achieved towards their daily goal. In each condition, the interface displays the time by default, as prior work shows that people often glance at the time and their physical activity concurrently [30]. We wanted to separate occurrences when people looked at the time from when they checked their step count. Step count data is pushed to the user in the two tactile conditions and can be pulled in all three conditions by pressing on the middle button on the right side of the watch. The three conditions were designed as follow:

***Visual Only (VO)***. This interface is Pull-only, so there is no visual feedback unless the user presses the button. When the user presses the button, the watch displays the total step count and percentage of completion towards the goal between 0% and 100%, to the 1% (Figure 2-1).

***Tactile Only (TO)***. The user receives a meaningful vibration (based on ActiVibe [9] with pre-vibration) every time they reach a 10% increment towards completing their goal. The encoding is based off of the Roman numeral system. One to four short vibrations (150ms) indicate values 1 to 4, and 5 is represented by a long vibration (600ms), 6 is then defined as 5 + 1: a long followed by a short vibration. For example, when reaching 30%, the user feels a pre-vibration indicating the signal, followed by 3 short vibrations corresponding to 30%. When the user presses the button, the last vibration sent is re-played. If their current status is between 0% and 10%, they will feel the pre-signal only (Figure 2-2).

***Hybrid: Visual/Tactile (V+T)***. The user receives a single short vibration for each 10% increment reached towards completing the goal. When the button is pressed, the watch displays the percentage of completion


Fig. 2. Smartwatch displays when the button is pressed. (1) Visual Only: displays step count value and percentage towards the goal. (2) Tactile Only: the screen does not change and the last encoded vibrations are re-played. (3) Hybrid: Visual/Tactile: displays percentage achieved towards the goal.

between 0% and 100%, to the 10% (Figure 2-3). The V+T condition is designed with a single vibration instead of an encoded vibration to tease apart the causal mechanism, either Push (V+T) itself or the encoded Push feedback (TO).

Both the TO and V+T interface designs follow the concept of a micro-planning activity such as supporting people in "reaching 1,000 steps in the next hour" [30]. Since the vibrations are designed to raise awareness and do not require any user action, we do not time them to be the least disrupting [24, 51]. In all conditions, to give participants a sense of accomplishment, we cap the daily 10,000 steps goal to 100% and do not display additional information when participants have already completed the goal. The next section describes the longitudinal study designed to evaluate the differences between the three interfaces.

## 4 6-WEEK LONGITUDINAL STUDY

To assess whether the modality of feedback on wearable activity tracking devices would affect people, we ran a 6-week longitudinal study across the three conditions of the system: Visual Only, Tactile Only, and Hybrid: Visual/Tactile. This study helped us to assess how the feedback modality of an activity tracker affects people's behavior and mindset. Six weeks is a limited time to look at behavior change, but it is sufficient to gather data from both weekends and weekdays, days with a routine and days without, and see differences based on the human-computer interface design. We explored the following hypotheses in the study:

- *H1*: The modality in which feedback is sent to a user affects their behavior and perception.
- *H2*: There is a positive effect on behavior and perception when the information is sent using a Push mechanism.
- *H3*: Encoding vibrotactile information has an impact over conveying a single vibration using the Push mechanism.
- *H4*: Push feedback leads to higher step counts.

### 4.1 Study Design and Conditions

The study was a between-subject design with 3x15 participants who were randomly assigned to one of three experimental conditions. The design choices and related user interface are detailed in Section 3 *Hardware and Interface Design*.

## 4.2 Goal

We opted for an assigned goal of 10,000 steps per day, which corresponds to an acceptable step count goal for healthy adults [62, 63]. While prior works suggest that people are more receptive to self setting goals [12, 45], this single goal-setting model is well established in commercial devices and prior work shows that "the default 10,000 steps was dominant" across Fitbit users [59].

## 4.3 Baseline

Several studies recruit participants that are already wearable step counter users, and as such use their existing average step count as a baseline [37, 59]. Others use the first week of the study to generate a baseline. In this study, we consider all types of users, who may or may not already be tracking their step count. When discussing the option for a week to establish the baseline, we considered that wearing the watch and being informed about the study would raise participants' awareness and potentially affect their behavior. As such, we decided against and opted for a no-baseline model as in [60, 64].

## 4.4 Procedure

To verify the protocol was sound and that the data collection was properly working, we ran an 8-day pilot user study with 12 participants. The study procedure and interview steps are detailed below.

*Step 1: Background questionnaire and sign up.* The study was advertised online via public forums and Facebook groups. People interested in the study were asked to fill in a background questionnaire including questions to verify their eligibility. Selected participants were asked to complete a consent form and a questionnaire about their demographics, work and transportation habits, and physical activity attitudes and routines prior to the first in-person session.

*Step 2: Pre-study session.* This first in-person session was run in our laboratory and lasted 30 to 45 minutes per person. Participants completed questionnaires about their exercise process mindset and exercise benefit mindset. Participants had their height and weight measurements taken by the research team. The apps were installed on their iPhone and they received the study smartwatch as well as care instructions. Participants in the Tactile Only condition learned the vibration encoding and were tested to make sure they could recognize it as in [9].

*Step 3: Longitudinal user study.* Participants were encouraged to carry their phone and wear the watch every day until at least 8pm to capture data from their time at home, at work, commuting, as well as while performing various activities. The compensation at the end of the study was based on their wearing the watch and not on performing activities or meeting goals. This allowed us to see how well participants would do towards their goal without the compensation influencing their step count behavior. Yet, this meant participants would be encouraged to wear the technology so we could not observe lapses or abandonment of the technology in this study.

*Step 4: Post-study interview.* The second in-person session was conducted post-study, approximately six weeks after the first session, and lasted 30 minutes to an hour. Participants returned the smartwatch, they were interviewed about their experience in the study, had their weight measurement retaken, and repeated the two mindset questionnaires.

## 4.5 Participants

The study ran July to August 2016 in the San Francisco Bay Area, California. Forty-four people (22 female, 22 male) between the ages of 19 and 70 years old ($\mu$=41, SD=16) participated in the study[2]. All participants were iPhone users with a data plan. They had a wide range of lifestyles and professions as shown in Appendix A. They

[2]One if the original 45 didn't come in for the pre-study session and was removed

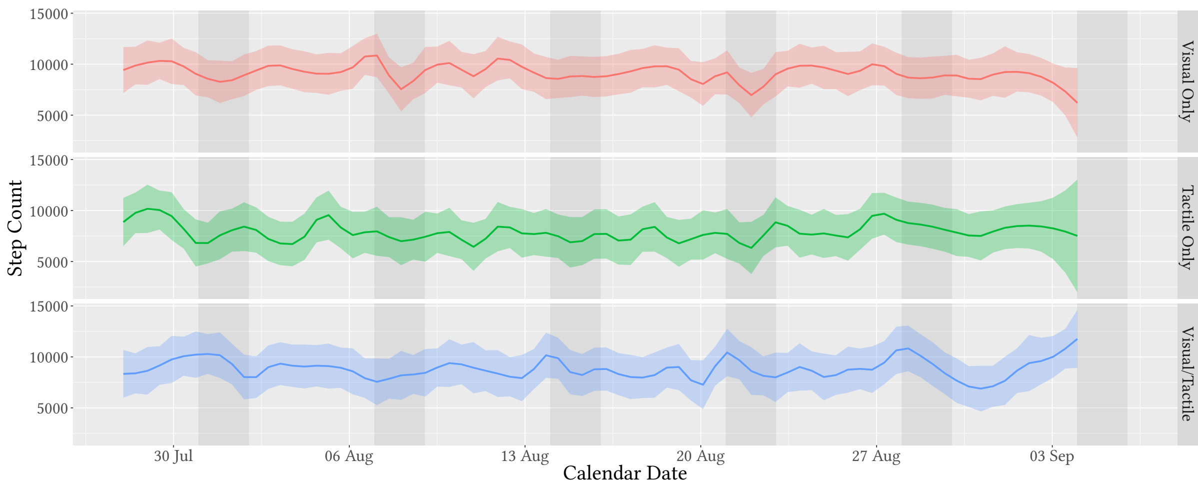



Fig. 3. Number of steps taken per day for all participants across the three conditions between July 26th and September 5th. Average value is represented as a bold line. Week-ends are shaded in grey.

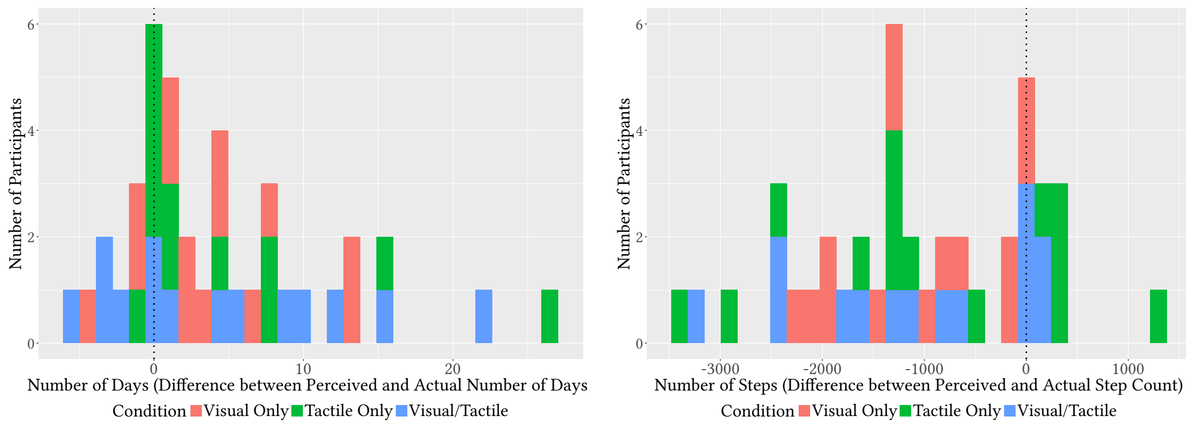
had different workday schedules and means of transportation. Sixteen (16) were classified as normal weight, sixteen (16) as overweight, and twelve (12) as obese according to Body Mass Index (BMI) calculations performed on their height and weight measurements taken during the pre-study session. This diversity was essential to verify our theory with different populations.

While their workout habits differed, most participants (38) wanted to increase their physical activity by the end of the summer and some (6) wanted it to stay at the same level, with 40 participants considering walking as a form of exercise. Seventeen participants already used wearables (e.g., Fitbit, Nike Surge, or Apple Watch) and/or step count apps (e.g., iPhone health app or Google fit) and agreed to not use any of them during the study. Participants were compensated US$50 for taking part in the study and up to another US$50 for complying with it (i.e., how many days they wore the watch regardless of their performance towards their goal).

## 4.6 Measures

We collected several measures to understand both how people were using the technology and how they were affected by it.

- *Step Count.* We measured participants' step count throughout the day and over the course of the study.
- *Vibration.* We collected the time and value of the vibrations felt by the users in the Push conditions, which corresponds to the time they reach 10% of the goal (i.e., a 1,000 step sub-goal).
- *Button Press.* We collected the time of each button press, when participants check their current activity status on the watch.
- *Height and Weight.* The participants' measurements were taken during the pre and post-study sessions and their BMI was calculated accordingly.
- *Exercise Process Mindset.* This measure (1-4 scale) refers to an individual's mindset about the process of engaging in exercise or physical activity [6] (see Appendix B).
- *Exercise Benefit Mindset.* This measure (1-7 scale) determines the role of participants' perceptions about the benefits associated with exercise [6].




(a) (b)

Fig. 4. (a) Difference between perceived and actual number of days where the goal of 10,000 steps was reached. (b) Difference between perceived and actual daily step count across all participants and conditions.

Figure 1 shows an example of a participant's step count and progress towards their goal throughout a day, including when vibrations and button presses occurred.

## 5 QUANTITATIVE DATA ANALYSIS

We gathered data from 44 participants over 6 weeks. We processed each of the corresponding data sets to test the assumption of normality. The Shapiro-Wilk's test was significant ($p < 0.05$), so that the assumption of normality could not be met. As such, we used Aligned Rank Transform (ART), a transformation performed on ANOVA for nonparametric factorial analyses [65] in all analyses below. Where there was a significant main effect, we performed post-hoc tests by comparing pairwise the least-square means, using Tukey correction for multiple comparisons [41]. The following subsections with a * in the title present results where statistical significance was found across conditions.

### 5.1 Compliance and Number of Days Used

We considered that a participant was compliant and wore the watch if at least 100 steps were recorded in a day. Participants who did not wear the watch for more than a third of the total study days were removed from further analysis. As such, P4 and P19 did not meet the criteria for compliance and their data was removed from further analyses. We further analyze data from 42 compliant participants (14 per condition) and data from days where the step count is at least 100 steps (i.e., the watch was worn).

On average, the study lasted 39 days per participant and we gathered 1,662 days of data. After removing data from non-compliant days and participants, we found a total of 1,608 days of data where the watch was worn. Participants wore the watch, on average, 98.9% of the days (SD = 3.3). We did not find a significant effect of the condition on percentage of days the watch was worn ($F(2,39)=1.39$, $p=0.26$).

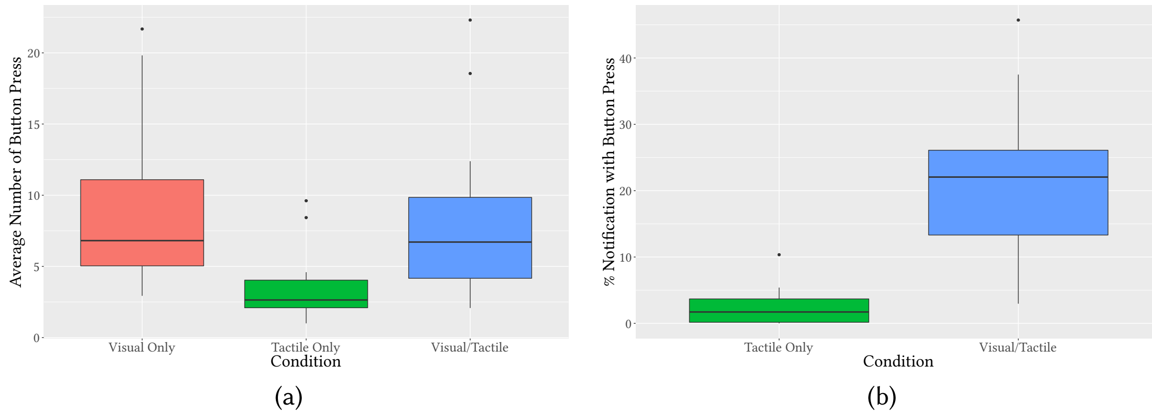



Fig. 5. (a) Average number of button press. The TO condition shows significantly less presses than the other 2 conditions. (b) Percentage of notifications where a button press occurred within 60 seconds of a *Push* notification.

## 5.2 Step Count

On average participants walked 8,517 steps per day (SD = 2,978). We did not find a significant effect of condition (F(2,36)=0.81, p=0.45) or gender (F(1,36)=1.52, p=0.23) on the average daily step count or a significant interaction between condition and gender (F(2,36)=0.07, p=0.94). We performed linear regressions to assess the evolution of the step counts over the course of the study (Figure 3) and did not find significant trends across subjects or significant differences between conditions (p>0.05).

## 5.3 Users' Perception of Step Count and Goal

Post study, we wanted to verify participants' awareness of their step count. We asked them what percentage of completion of their goal they thought they had reached on average each day. We compared these numbers with the actual numbers of steps and percentage towards the 10,000 steps goal. With an average difference of -1,045 steps (SD = 1,102), participants overall under-estimated their daily step count (Figure 4b). We did not find a significant effect of condition (F(2,36)=0.19, p=0.83) or gender (F(1,36)=0.50, p=0.48) on these estimates, nor a significant interaction between condition and gender (F(2,36)=1.34, p=0.27).

On average participants reached the 10,000 steps goal 32.8% of the study days (SD = 24.1). Post study, participants were asked how many times they thought they had reached the goal[3]. We compared these numbers with the number of days when they actually reached the goal. With an average difference of 4.5 days (SD = 7.0), participants overall over-estimated the number of days they reached their goal (Figure 4a). We did not find a significant effect of condition (F(2,34)=0.01, p=0.99) or gender (F(1,34)=0.03, p=0.87) on these estimates, nor a significant interaction between condition and gender (F(2,34)=0.13, p=0.88).

[3]Note: P8 and P13 did not provide an answer to this specific question.

## 5.4 Button Press on the Smartwatch*

We recorded 6,905 button presses over the course of the study[4]. In the VO and V+T conditions, a press would display the current number of steps and percentage toward the goal on the watch screen, while in the TO condition, the press would replay the last sent encoded vibration message.

*Overall Number of Button Presses*.* We found a significant effect of the condition on the number of button presses, with less presses in the Tactile Only (TO) condition as compared to the Visual Only (VO) or Hybrid (V+T) conditions (Figure 5a). On average participants in the VO condition pressed to see their step count 8.9 times per day (SD = 5.9), in TO: 3.6 times per day (SD = 2.5) and in V+T: 8.3 times per day (SD = 5.9) ($F(2,39)=9.31$, $p < 0.001$).

*Button Press Timing*.* We measured the percentage of vibrations that were followed within 60 seconds by a button press. There was a significant effect of the condition ($F(1,26)=51.3$, $p < 0.001$) (Figure 5b). On average participants in the V+T condition pressed the button following a vibration 21.8% (SD = 11.7) of the time compared to only 2.5% (SD = 2.9) of the time in the TO condition.

## 5.5 Exercise Process Mindset*

The Exercise Process Mindset (EPM) measure had good reliability (time-1 measurement: Cronbach's alpha = 0.82; time-2 measurement: Cronbach's alpha = 0.81). We first analyzed EPM for each condition and then within conditions at time-1 (pre-study) and comparing the change from time 1 to time-2 (post-study). Note that EPM higher values represent more positive mindsets.

At time-1, the baseline measurement of EPM appeared to be higher in the Hybrid (V+T) condition compared to the other conditions, but pre-treatment differences were not statistically significant as shown by a one-way ANOVA ($F(2,39)=0.803$, $p=0.455$). At time-2, EPM measured after completion of the study was lower for participants in the Visual Only (VO) condition than in the two tactile conditions (Table 1). To examine whether EPM changed significantly from time-1 to time-2 *within* each feedback condition, we conducted a series of paired t-tests. Figure 6 shows EPM at time-1 and time-2 depending on feedback condition (also shown in Table 1), along with error bars representing standard errors. In the VO condition, EPM decreased slightly time-1 to time-2, but this difference was not significantly different from zero ($t(13)=-1.995$, $p=0.067$). In the TO condition, EPM increased significantly from time-1 to time-2 ($t(13)=2.222$, $p=0.045$). In the Hybrid (V+T) condition, EPM remained virtually unchanged

[4]It is possible that not all presses were recorded because of technical limitations, such as if a participant did not have network connectivity for a long period. The users would still have seen/felt the step count nonetheless.

Table 1. Means (and Standard Deviations) for Time-1 EPM, Time-2 EPM, and EPM change from pre to post-study.

| Condition | Time-1 EPM | Time-2 EPM | EPM Change |
|---|---|---|---|
| VO | 2.85 (0.36) | 2.68 (0.49) | -0.16 (0.31) |
| TO | 2.83 (0.61) | 3.00 (0.52) | 0.17 (0.29) |
| V+T | 3.04 (0.48) | 2.99 (0.33) | -0.05 (0.38) |

Table 2. Results of Regression Predicting Time-2 EPM (Controlling for Time-1 EPM).

| | Estimate | Std. Error | t-value | p-value |
|---|---|---|---|---|
| VO vs. TO & V+T | 0.083 | 0.033 | 2.550 | 0.015* |
| TO vs. V+T | 0.081 | 0.057 | 1.413 | 0.166 |
| Time 1 EPM | 0.708 | 0.097 | 7.324 | < .001* |

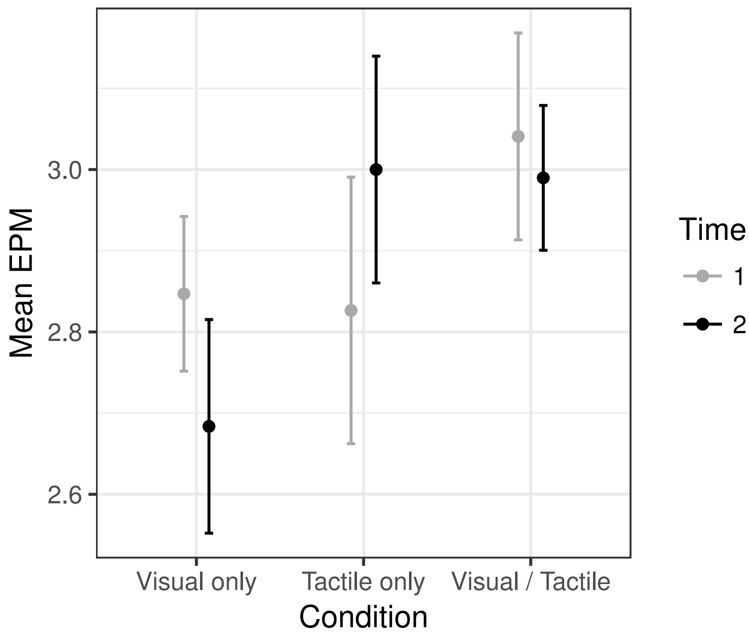



Fig. 6. Exercise Process Mindset (EPM) at time-1 and time-2 by condition. The EPM values range from 1 to 4 with higher values representing more positive mindsets. Error bars represent standard errors.

(t(13)=0.504, p=0.622). Note that when using Bonferroni adjusted alpha levels of .017 (.05/3) to account for multiple testing, none of these changes reached statistical significance.

To examine whether participants' time-2 EPM differed significantly depending on whether they received feedback using *Pull* or *Push* mechanisms (*H2*), we conducted regression analysis with feedback condition as predictor and controlling for time-1 EPM as a covariate. Planned contrasts were used to reveal whether the VO condition (*Pull*) had significantly different effects on mindset than either of the two tactile conditions (TO & V+T, *Push*), and whether these two tactile conditions differed from one another. We found that the tactile conditions were associated with significantly higher scores on time-2 Exercise Process Mindset than the VO condition (b=0.08, t(38)=2.550, p=0.015). However, we did not find a significant difference between the two tactile conditions (p=0.166; see Table 2 for full results).

These results show that participants exposed to any of the two tactile conditions were more likely to believe that exercising is easy, pleasurable, and fun, whereas participants exposed to visual feedback were more likely to believe that exercising is hard, painful, and boring after the study – controlling for how participants felt about exercise before the study.

## 5.6 Exercise Benefits Mindset

The measure included 7 items on 7-point scales, such as "My current level of physical activity is healthy" (Strongly agree - Strongly disagree); "How much does your current level of physical (in-)activity increase or decrease your risk of disease?" (Increases my risk very much - Decreases my risk very much). The measure had good reliability (Cronbach's $\alpha$ = 0.89), yet we did not find significant effects.

### 5.7 Additional Measures

We found no significant differences between male and female participants across conditions. While several participants mentioned having different walking patterns and reporting using the technology differently between work days and weekends, we found no significant differences in our quantitative data analysis. We also did not observe any significant changes in weight loss or BMI over the course of the study.

## 6 QUALITATIVE DATA ANALYSIS

This section presents the main findings gathered in the post-study interviews.

### 6.1 Activity Level and Behavior Change

When asked if they thought they were more active as a result of wearing the step counter, most participants (24) answered "yes", 13 answered "no", and 5 thought they were about the same. Participants who considered having actively changed their behavior reported making the time for it (P4), taking the stairs more often, taking *"the long way around from place to place"* (P14), trying *"to find ways to get more walking in"* (P34), walking *"to a further away bus stop to get to the goal"* (P44), and sometimes trading their bike for walking (P3, P15). Participants reported how wearing the step counter and trying to achieve the goal raised their awareness of activity or sedentarity. "*At first I didn't realize how much I was sitting. It gave me an increased awareness of what I was NOT doing"* (P36).

A few participants mentioned the novelty effect and how "after the first two weeks, the novelty wore off and I went back to my usual patterns" (P30). This effect was however not visible in the quantitative data analysis of the step counts over time. People who considered that they did not modify their activity behavior mentioned, *"Some days I was just busy with other things"* (P2) and *"if I had worked harder at it, I would have gone on walks in the evening"* (P18). We observed several barriers to why people did not reach the goal.

### 6.2 Barriers to Reaching the Goal

We received heterogeneous feedback to how challenging it was to reach the goal, with participants finding it easy (15), hard (15), and some (12) who did not try to hit the goal. We identified several barriers such as scheduling, injuries, and disappointment when the user wanted to track an activity not supported by the step counter. Despite these barriers, participants reported being satisfied when they would reach the goal: "I was amazed how quickly you get to 10,000. I felt good. I felt proud of myself" (P25). Specifically, participants commented on the vibration notification and how they were *"getting positive feedback on the process of becoming healthy"* (P27).

### 6.3 Notifications and Vibrations

Most participants in the tactile conditions (26/28) felt positively about the vibration feedback. Most (20) reported that vibrations were "not annoying at all", while three (3) were sometimes disturbed by it or bothered when it would go off by mistake, such as picking up arm movement instead of walking. In the Tactile Only condition, all participants reported understanding the value conveyed by the vibrations without effort. Both in TO and V+T, the Push feedback was found to be helpful and all but one found the vibrations comfortable. Some participants reported getting used to them and looking forward to receiving a vibration.

Participants enjoyed receiving the regular feedback, which "made me feel I achieved something" (P8), and P9 qualified as *"almost comforting"*. While most enjoyed the 10% sub-goal, some participants wanted the ability to adjust when the vibration would occur to either be more *"granular"* (P5) or less frequent. In TO, only one person (P8) mentioned they would have liked an additional visual display. Other considerations that arose from the qualitative data analysis are listed below.

### 6.4 Additional Considerations

Most people were comfortable wearing the smartwatch, yet some reported discomfort because they either weren't used to wearing a watch or would get perspiration on their wrist[5]. Some mentioned issues around trust and accuracy. There were also battery considerations despite the watch being connected using Bluetooth Low Energy (BLE), which led to P19 abandoning the study. The following section discusses the limitations of this work.

## 7 DISCUSSION

The results show that hypotheses H1 and H2 were partially supported, H3 was fully supported, and H4 was refuted. In this section we discuss the findings of the 6-week longitudinal study.

### 7.1 Visual vs. Tactile Feedback

- *H1*: The modality in which feedback is sent to a user affects their behavior and perception.

We found a significant effect on the Exercise Process Mindset between the visual condition and the two tactile conditions, however we did not find any significant differences of activity behavior. This hypothesis is therefore partially supported by this research.

People maintained their step count over the course of the study. Prior work had found that participants using an awareness display (i.e., glanceable activity tracker on their phone) would maintain their activity level throughout a 3-month study [14]. Our results are in par with theirs for maintaining activity level with all 3 conditions being comparable to results with glanceable and ambient displays.

Step count did not show significant differences across conditions and participants had a high awareness of their step count, which they overall underestimated by 1,000 steps on average across conditions. This is particularly interesting as the two tactile conditions present information at a lower resolution, to the 10% rather than to the 1%; yet, it led to same level of understanding of the activity log.

In both tactile conditions participants view the exercise process as easier, more pleasurable, and more relaxing than in the visual condition. We conclude that in our work: ***Tactile feedback is better suited to wearable activity tracking than purely visual feedback.***

### 7.2 Push vs Pull Mechanisms

- *H2*: There is a positive effect on behavior and perception when the information is sent using a Push mechanism.

This hypothesis is partially proven. We found a significant effect on the Exercise Process Mindset between the visual condition (*Pull*) and the two tactile conditions (*Push*) as presented in the previous section but we did not find differences in terms of activity behavior.

In the tactile conditions (TO and V+T), which intrinsically provide a (*Push*) mechanism, participants pulled the data 3.6 and 8.3 times per day on average for TO and V+T, respectively. Since participants only pressed the button within 60 seconds of the encoded vibration 2.5% of the time in TO against 21.8% in V+T, we assume that the vibrations were well understood as in [9]. In TO, participants pressed the button *"when it didn't vibrate for a while"* (P1,P3,P14) or *"when I am actually trying to figure out where I am"* (P12).

This shows that while the *Push* mechanism has a positive effect on mindset, ***Push and Pull mechanism should both be available for wearable activity trackers.***

[5]Temperatures went up as high as 31°C over the summer we ran the study.

### 7.3 Vibration Reward

- *H3*: Encoding vibrotactile information has an impact over conveying a single vibration using the Push mechanism.

This hypothesis was supported as we found a significant difference in the number of button presses and the time when they occur between the two tactile conditions (TO and V+T). In the tactile only (TO) condition, participants pressed the button 2.5 times less often than in the hybrid condition (V+T); yet, they had an equivalent amount of daily step count and equivalent awareness of how well participants were doing towards their goal.

Both tactile conditions created awareness: *"I was starting to get a sense for what was 1,000 steps, an unconscious sense that I should get a buzz and [...] I would get a buzz"* (P10). Vibrations were seen as a positive motivator, *"it is a good feeling, I have achieved something"* (P8). Yet, P16 (V+T) mentions *"I would have liked a celebratory 100%. Since it was only just one buzz always the same, then it didn't mean anything [...] no feeling of doing better."* about the single vibration, showing the interest in encoded tactile notifications.

Prior work explored graphical representations of rewards for activity monitoring with varying degrees of success. For example, trophies and ribbons [47] did not succeed while a flourishing fish tank [44] and garden [11, 14] did support users in maintaining their activity level. We show that ***vibrations can provide meaningful feedback and be used as a reward mechanism.*** The vibrations are considered positively as they are reflective of a positive behavior.

### 7.4 High Step Count

- *H4*: Push feedback leads to higher step counts.

We expected that there would be an effect of modality on the participants step count, however this hypothesis was refuted.

We found an average daily step count of 8,580 steps across participants. While our study design did not allow for a comparative baseline (see Section 9 Limitations), data gathered from 2013 to 2014 shows that on average in the United States people walk 4,774 steps per day [2]. The high step count in our study may be linked to an initial self-selection bias of the people interested in taking part in the study, or possible to the Hawthorne effect as participants were conscious of being monitored, such as P27 who mentioned that having the team *"Big Brother"* watching their progress made them want to do more.

The section below discusses the design implications of this work.

## 8 DESIGN IMPLICATIONS

Our findings led to implications for the design of future wearable activity trackers and we propose recommendations for future longitudinal studies in behavior change.

### 8.1 Interface Design & Reward System

We received many comments around goal setting and the display being capped at 100%. Regarding the goal itself, some participants mentioned *"If I would set a higher goal, maybe I would be more motivated"* (P31), which is consistent with the prior literature. Locke and Latham [45] define two factors that facilitates a person's commitment to the goal: the importance of the goal and the self-efficacy, a person's belief that they can reach the goal. In our study, self-efficacy seems to have played a stronger negative role, probably due to the fact that the goal was assigned and not self-set. Although P42 reported *"I am competitive so I would want to complete the goal."* In terms of display, participants wanted to go beyond the 100% mark: *"I wished it had something past 100%, to have more incentive"* (P12). In the tactile conditions, the Push feedback became associated to "little milestones" or sub-goals. These results are consistent with prior work showing that goals can be divided into daily and hourly goals [59]. The vibrations transformed into a physical reward system to the participants and a positive motivator.

### 8.2 Reflective vs. Impulsive Thinking

Li et al. [42] define personal informatics systems, such as activity trackers, as a five-stage process. Epstein et al. [21] build upon it and propose a modified version to reflect a "lived informatics view." We focus on the "tracking and acting" stages [21], from collection to integration to reflection, where *integration* corresponds to the efforts needed to visualize log information. Our interfaces minimize *integration*, especially in the tactile conditions where activity data is *Pushed*. Prior work suggests the importance of the "impulsive" over the "reflective" mode [57]. Gouveia et al. [30] show that over 70% of the time, users glanced at the wearable for under 5 seconds with no further exploration. They conclude that these glances may serve towards learning and discuss the need to "develop mechanisms to support learning through these frequent glances". Our interface could be an answer since it provides immediate *reflection* and possibly triggers the "impulsive" mode when participants feel their current activity level through the tactile modality. This notion of "immediate impact" was previously reported in [27] for long term activity tracker users. In all three conditions users were aware of their overall step count despite only seeing a current value without a holistic view. Many participants mentioned tailoring their behaviors to reach the goal, successfully reaching the *action* stage [42].

### 8.3 Longitudinal Studies in The Age of Big Data

Finally, we address some of the challenges faced in running this longitudinal study, as considerations for future researchers. Despite our instructions to only use the smartwatch for step count, some participants admitted to having either: voluntary double-checked their step count on a phone app (i.e., Apple Health, or FitBit app), or inadvertently seen it, such as when playing Pokemon on their Gameboy (P7). This is a common behavior in personal tracking, referred to as "Interweaving Personal Trackers" [53]. It is an issue to consider for future research where participants have access to many devices and apps that are not trivial to monitor and require addressing major privacy concerns. We here investigated single person usage, and did not dive into social aspects of activity tracking and behavior change. Yet, we found out post-study that P24 & P29 participated as a couple, by chance, allocated to the same condition. Several participants mentioned discussing their step count on a regular basis with family members, friends, and colleagues interested in the study *"I would tell my friend to walk and talk so I would make my step count"* (P26). Once again, we enter an age where people become increasingly comfortable sharing personal data and these social interactions are difficult to measure with traditional research methods.

The following section discusses the limitations of this work.

## 9 LIMITATIONS

This study was six weeks long and was too short to witness actual behavior change. Another limitation is the lack of a clear baseline. Several studies recruit participants that are already wearable step counter users, and as such use their existing average step count as a baseline [37, 59]. Instead, we recruited a larger range of participants where most (25) did not use a step counter prior to the study, which prevented us from using existing step count as baseline. We considered that wearing the step counter itself would raise participants' awareness and potentially affect their behavior, so decided against creating a baseline with the study hardware and opted for a no-baseline model as in [60, 64]. While we could not see differences in terms of step count or behavior, from a user interface standpoint, we identified differences across modalities and showed that over the course of six weeks participants kept their step count up.

## 10 FUTURE WORK

Future work will investigate alternative *Push* mechanisms such other haptics sensations or even audio feedback, which might present different challenges around privacy and location of the device on the body. As wearable devices are increasingly being used, we see new form factors without screens and also devices which are not necessarily positioned to be easily glanced at. Because of this major difference compared to smartphones, we anticipate that tactile displays are a viable alternative to current visual only displays, and in particular in the context of activity tracking and behavior change. This may also increase the acceptability of wearable devices that do not display information visible to all.

Prior work suggests that abandonment and lapsing, when a person stops actively using their self-tracking tool, is a major issue with activity tracking [1]. However, since we investigated the differences in perception across modalities, our study design encouraged participants to keep using the watch for the duration of the study. Future work will investigate whether the lapses would happen in the same fashion with tactile feedback as they do with visual feedback. Future work will investigate whether there are differences across age groups as suggested in [36].

## 11 CONCLUSIONS

In this work, we explored the effect of Push vs Pull feedback modality for wearable fitness trackers. As insurance companies are starting to only propose "interactive" life insurance policies via mobile and wearable fitness trackers [58], it has never been more crucial to understand how these devices can best support users. In a 6-week longitudinal study (N=44) across 3 conditions: visual, tactile, and hybrid, we found that tactile feedback is a viable alternative to visual feedback for progress monitoring on wearables. We found specific differences on participants' mindsets about the process of exercise depending on the feedback mechanism. In particular, users maintained a positive exercise process mindset with *Push* feedback, viewing the process of exercising as easier, more pleasant, and fun, compared to users exposed to a traditional visual Pull interface. In contrast, users of the visual Pull interface saw a negative impact on their exercise process mindset, such that they were more likely to believe that exercising is hard, painful, or boring after the study. This is important as mindsets about exercise have powerful effects on engagement in exercise behavior and can shape psychological and physiological health outcomes. We also found that vibrations were considered positively and as a reward by the participants. This work was the first to uncover the positive impact of tactile displays for activity tracking.

## ACKNOWLEDGMENTS

We would like to thank our colleagues for their help and support with this project. In particular, we are grateful to Rina Imari Horii and Elijah John Sampson Zenger for their help with the intake and debriefing for the user study. Many thanks to Dr. Martin Pielot for his feedback on an earlier version of this paper, and to Jenna Goldberg for proof reading the manuscript for submission.

## REFERENCES

[1] Elena Agapie, Daniel Avrahami, and Jennifer Marlow. 2016. Staying the Course: System-Driven Lapse Management for Supporting Behavior Change. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*. ACM Press, New York, New York, USA, 1072–1083. https://doi.org/10.1145/2858036.2858142

[2] Tim Althoff, Rok Sosič, Jennifer Hicks, Abby C King, Scott Delp, and Jure Leskovec. 2017. Large-scale physical activity data reveal worldwide activity inequality. *Nature* 547 (07 2017), 336. http://dx.doi.org/10.1038/nature23018

[3] Swamy Ananthanarayan and Katie A. Siek. 2010. Health sense: a gedanken experiment on persuasive wearable technology for health awareness. In *Proceedings of the ACM international conference on Health informatics - IHI '10*. ACM Press, New York, New York, USA, 400–404. https://doi.org/10.1145/1882992.1883051

[4] Roy F. Baumeister, Ellen Bratslavsky, Mark Muraven, and Dianne M. Tice. 1998. Ego depletion: is the active self a limited resource? *Journal of personality and social psychology* 74, 5 (5 1998), 1252–1265.

[5] Kent C. Berridge and Terry E. Robinson. 2003. Parsing reward. *Trends in neurosciences* 26, 9 (9 2003), 507–513. https://doi.org/10.1016/S0166-2236(03)00233-9

[6] Danielle Z. Boles, Maysa DeSousa, Bradley P. Turnwald, Rina Horii, Taylor Duarte, Hazel Rose Markus, and Alia J. Crum. 2019. If being healthy is hard, painful and depriving, it is no wonder we are all fat and sick: Mindsets about the process of health provide a new avenue for motivating health behavior. *Work in Submission* (2019).

[7] Dena M. Bravata, Crystal Smith-Spangler, Vandana Sundaram, Allison L. Gienger, Nancy Lin, Robyn Lewis, Christopher D. Stave, Ingram Olkin, and John R. Sirard. 2007. Using pedometers to increase physical activity and improve health: A systematic review. *Journal of the American Medical Association* 298, 19 (11 2007), 2296–2304. https://doi.org/10.1001/jama.298.19.2296

[8] Patrick Burns, Christopher Lueg, and Shlomo Berkovsky. 2013. Colours That Move You: Persuasive Ambient Activity Displays. In *Proceedings of the 8th international conference on Persuasive Technology*. Springer-Verlag, 27–32. https://doi.org/10.1007/978-3-642-37157-8{_}5

[9] Jessica R. Cauchard, Janette L. Cheng, Thomas Pietrzak, and James A. Landay. 2016. ActiVibe: Design and Evaluation of Vibrations for Progress Monitoring. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*. ACM Press, New York, New York, USA, 3261–3271. https://doi.org/10.1145/2858036.2858046

[10] James Clawson, Jessica A. Pater, Andrew D. Miller, Elizabeth D. Mynatt, and Lena Mamykina. 2015. No longer wearing: investigating the abandonment of personal health-tracking technologies on craigslist. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15*. ACM Press, New York, New York, USA, 647–658. https://doi.org/10.1145/2750858.2807554

[11] Sunny Consolvo, Predrag Klasnja, David W. McDonald, Daniel Avrahami, Jon Froehlich, Louis LeGrand, Ryan Libby, Keith Mosher, and James A. Landay. 2008. Flowers or a robot army?: encouraging awareness & activity with personal, mobile displays. In *Proceedings of the 10th international conference on Ubiquitous computing - UbiComp '08*. ACM Press, New York, New York, USA, 54–63. https://doi.org/10.1145/1409635.1409644

[12] Sunny Consolvo, Predrag Klasnja, David W. McDonald, and James A. Landay. 2009. Goal-setting considerations for persuasive technologies that encourage physical activity. In *Proceedings of the 4th International Conference on Persuasive Technology - Persuasive '09*. ACM Press, New York, New York, USA, 1. https://doi.org/10.1145/1541948.1541960

[13] Sunny Consolvo, Predrag Klasnja, David W. McDonald, and James A. Landay. 2014. Designing for Healthy Lifestyles: Design Considerations for Mobile Technologies to Encourage Consumer Health and Wellness. *Foundations and Trends®in Human-Computer Interaction* 6, 3-4 (2014), 167–315. https://doi.org/10.1561/1100000040

[14] Sunny Consolvo, Ryan Libby, Ian Smith, James A. Landay, David W. McDonald, Tammy Toscos, Mike Y. Chen, Jon Froehlich, Beverly Harrison, Predrag Klasnja, Anthony LaMarca, and Louis LeGrand. 2008. Activity sensing in the wild: a field trial of ubifit garden. In *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08*. ACM Press, New York, New York, USA, 1797–1806. https://doi.org/10.1145/1357054.1357335

[15] Alia J. Crum, William R. Corbin, Kelly D. Brownell, and Peter Salovey. 2011. Mind Over Milkshakes: Mindsets, Not Just Nutrients, Determine Ghrelin Response. *Health Psychology* 30, 4 (7 2011), 424–429. https://doi.org/10.1037/a0023467

[16] Alia J. Crum and Ellen J. Langer. 2007. Mind-set matters: Exercise and the placebo effect. *Psychological Science* 18, 2 (2 2007), 165–171. https://doi.org/10.1111/j.1467-9280.2007.01867.x

[17] Roelof A.J. de Vries, Khiet P. Truong, Sigrid Kwint, Constance H.C. Drossaert, and Vanessa Evers. 2016. Crowd-Designed Motivation: Motivational Messages for Exercise Adherence Based on Behavior Change Theory. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*. ACM Press, New York, New York, USA, 297–308. https://doi.org/10.1145/2858036.2858229

[18] Raymond Desharnais, Jean Jobin, Charles Côté, Lucie Lévesque, and Gaston Godin. 1993. Aerobic exercise and the placebo effect: A controlled study. *Psychosomatic Medicine* 55, 2 (1993), 149–154. https://doi.org/10.1097/00006842-199303000-00003

[19] Mitch Duncan, Beatrice Murawski, Camille E. Short, Amanda L. Rebar, Stephanie Schoeppe, Stephanie Alley, Corneel Vandelanotte, and Morwenna Kirwan. 2017. Activity Trackers Implement Different Behavior Change Techniques for Activity, Sleep, and Sedentary Behaviors. *Interactive Journal of Medical Research* 6, 2 (8 2017), e13. https://doi.org/10.2196/ijmr.6685

[20] Lígia Duro. 2018. Why Should Your Activity Tracker Care About How Motivational Messages Look Like to You. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (CHI EA '18)*. ACM, New York, NY, USA, Article SRC06, 6 pages. https://doi.org/10.1145/3170427.3180285

[21] Daniel A. Epstein, An Ping, James Fogarty, and Sean A. Munson. 2015. A lived informatics model of personal informatics. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15*. ACM Press, New York, New York, USA, 731–742. https://doi.org/10.1145/2750858.2804250

[22] Jordan Etkin. 2016. The Hidden Cost of Personal Quantification. *Journal of Consumer Research* 42, 6 (2016), 967–984. https://doi.org/10.1093/jcr/ucv095

[23] Chloe Fan, Jodi Forlizzi, and Anind Dey. 2012. Considerations for technology that support physical activity by older adults. In *Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility - ASSETS '12*. ACM Press, New York, New York, USA, 33. https://doi.org/10.1145/2384916.2384923

[24] Joel E. Fischer, Nick Yee, Victoria Bellotti, Nathan Good, Steve Benford, and Chris Greenhalgh. 2010. Effects of content and time of delivery on receptivity to mobile interruptions. In *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services - MobileHCI '10*. ACM Press, New York, New York, USA, 103. https://doi.org/10.1145/1851600.1851620

[25] Ayelet Fishbach, Tal Eyal, and Stacey R Finkelstein. 2010. How positive and negative feedback motivate goal pursuit. , 517–530 pages. https://doi.org/10.1111/j.1751-9004.2010.00285.x

[26] BJ Fogg. 2009. Creating persuasive technologies. In *Proceedings of the 4th International Conference on Persuasive Technology - Persuasive '09*. ACM Press, New York, New York, USA, 1. https://doi.org/10.1145/1541948.1542005

[27] Thomas Fritz, Elaine M. Huang, Gail C. Murphy, and Thomas Zimmermann. 2014. Persuasive technology in the real world: a study of long-term use of activity sensing devices for fitness. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*. ACM Press, New York, New York, USA, 487–496. https://doi.org/10.1145/2556288.2557383

[28] Çağlar Genç, Merve Erkaya, Fuat Balcı, and Oğuzhan Özcan. 2018. Exploring Dynamic Expressions on Soft Wearables for Physical Exercises. In *Proceedings of the 2018 ACM Conference Companion Publication on Designing Interactive Systems (DIS '18 Companion)*. ACM, New York, NY, USA, 147–152. https://doi.org/10.1145/3197391.3205427

[29] Rúben Gouveia, Evangelos Karapanos, and Marc Hassenzahl. 2015. How Do We Engage with Activity Trackers?: A Longitudinal Study of Habito. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. ACM, New York, NY, USA, 1305–1316. https://doi.org/10.1145/2750858.2804290

[30] Rúben Gouveia, Evangelos Karapanos, and Marc Hassenzahl. 2018. Activity Tracking in Vivo. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 362, 13 pages. https://doi.org/10.1145/3173574.3173936

[31] Rúben Gouveia, Fábio Pereira, Ana Caraban, Sean A. Munson, and Evangelos Karapanos. 2015. You Have 5 Seconds: Designing Glanceable Feedback for Physical Activity Trackers. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers (UbiComp/ISWC'15 Adjunct)*. ACM, New York, NY, USA, 643–647. https://doi.org/10.1145/2800835.2809437

[32] Daniel Harrison, Paul Marshall, Nadia Bianchi-Berthouze, and Jon Bird. 2015. Activity tracking: barriers, workarounds and customisation. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15*. ACM Press, New York, New York, USA, 617–621. https://doi.org/10.1145/2750858.2805832

[33] Ellen L. Idler and Yael Benyamini. 1997. Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies. *Journal of Health and Social Behavior* 38, 1 (3 1997), 21. https://doi.org/10.2307/2955359

[34] Mohammad Hossein Jarrahi, Nicci Gafinowitz, and Grace Shin. 2018. Activity trackers, prior motivation, and perceived informational and motivational affordances. *Personal and Ubiquitous Computing* 22, 2 (4 2018), 433–448. https://doi.org/10.1007/s00779-017-1099-9

[35] Aditya Kaul and Clint Wheelock. 2016. *Wearable Device Market Forecasts*. Technical Report. Tractica LLC. https://www.tractica.com/wp-content/uploads/2016/03/MD-WDMF-1Q16-Executive-Summary.pdf

[36] Jayden Khakurel, Antti Knutas, Helinä Melkas, Birgit Penzenstadler, Bo Fu, and Jari Porras. 2018. Categorization Framework for Usability Issues of Smartwatches and Pedometers for the Older Adults. In *Universal Access in Human-Computer Interaction. Methods, Technologies, and Users*, Margherita Antona and Constantine Stephanidis (Eds.). Springer International Publishing, Cham, 91–106.

[37] Rafal Kocielnik, Lillian Xiao, Daniel Avrahami, and Gary Hsieh. 2018. Reflection Companion: A Conversational System for Engaging Users in Reflection on Physical Activity. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 2 (7 2018), 1–26. https://doi.org/10.1145/3214273

[38] Amanda Lazar, Christian Koehler, Joshua Tanenbaum, and David H. Nguyen. 2015. Why we use and abandon smart devices. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15*. ACM Press, New York, New York, USA, 635–646. https://doi.org/10.1145/2750858.2804288

[39] I. Min Lee, Eric J Shiroma, Felipe Lobelo, Pekka Puska, Steven N Blair, Peter T Katzmarzyk, Jasem R. Alkandari, Lars Bo Andersen, Adrian E. Bauman, Ross C. Brownson, Fiona C. Bull, Cora L. Craig, Ulf Ekelund, Shifalika Goenka, Regina Guthold, Pedro C. Hallal, William L. Haskell, Gregory W. Heath, Shigeru Inoue, Sonja Kahlmeier, Harold W. Kohl, Estelle Victoria Lambert, Grit Leetongin, Ruth J.F. Loos, Bess Marcus, Brian W. Martin, Neville Owen, Diana C. Parra, Michael Pratt, David Ogilvie, Rodrigo S. Reis, James F. Sallis, Olga Lucia Sarmiento, and Jonathan C. Wells. 2012. Effect of physical inactivity on major non-communicable diseases worldwide: An analysis of burden of disease and life expectancy. *The Lancet* 380, 9838 (7 2012), 219–229. https://doi.org/10.1016/S0140-6736(12)61031-9

[40] Moon-Hwan Lee, Seijin Cha, and Tek-Jin Nam. 2015. Patina Engraver: Visualizing Activity Logs as Patina in Fashionable Trackers. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*. ACM Press, New York, New York, USA, 1173–1182. https://doi.org/10.1145/2702123.2702213

[41] Russell Lenth. 2016. Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software, Articles* 69, 1 (2016), 1–33. https://doi.org/10.18637/jss.v069.i01

[42] Ian Li, Anind Dey, and Jodi Forlizzi. 2010. A stage-based model of personal informatics systems. In *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*. ACM Press, New York, New York, USA, 557–566. https://doi.org/10.1145/1753326.1753409

[43] Brian Y. Lim, Aubrey Shick, Chris Harrison, and Scott E. Hudson. 2011. Pediluma: motivating physical activity through contextual information and social influence. In *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction - TEI '11*. ACM Press, New York, New York, USA, 173–180. https://doi.org/10.1145/1935701.1935736

[44] James J. Lin, Lena Mamykina, Silvia Lindtner, Gregory Delajoux, and Henry B. Strub. 2006. Fish'n'Steps: Encouraging Physical Activity with an Interactive Computer Game. Springer, Berlin, Heidelberg, 261–278. https://doi.org/10.1007/11853565{_}16

[45] Edwin A Locke and Gary P Latham. 2002. Building a practically useful theory of goal setting and task motivation. A 35-year odyssey. *The American psychologist* 57, 9 (9 2002), 705–17. https://doi.org/10.1037//0003-066x.57.9.705

[46] Jochen Meyer, Jutta Fortmann, Merlin Wasmann, and Wilko Heuten. 2015. Making Lifelogging Usable: Design Guidelines for Activity Trackers. In *MultiMedia Modeling*, Xiangjian He, Suhuai Luo, Dacheng Tao, Changsheng Xu, Jie Yang, and Muhammad Abul Hasan (Eds.). Springer International Publishing, Cham, 323–334.

[47] Sean Munson and Sunny Consolvo. 2012. Exploring Goal-setting, Rewards, Self-monitoring, and Sharing to Motivate Physical Activity. In *Proceedings of the 6th International Conference on Pervasive Computing Technologies for Healthcare*. IEEE, 25–32. https://doi.org/10.4108/icst.pervasivehealth.2012.248691

[48] National Center for Health Statistics (US). 2017. *Health, United States, 2016: With Chartbook on Long-term Trends in Health.* Technical Report. Hyattsville, MD. https://www.ncbi.nlm.nih.gov/pubmed/28910066

[49] David Ogilvie, Charles E Foster, Helen Rothnie, Nick Cavill, Val Hamilton, Claire F Fitzsimons, and Nanette Mutrie. 2007. Interventions to promote walking: Systematic review. *British Medical Journal* 334, 7605 (6 2007), 1204–1207. https://doi.org/10.1136/bmj.39198.722720.BE

[50] Fabian Okeke, Michael Sobolev, and Deborah Estrin. 2018. Towards A Framework for Mobile Behavior Change Research. In *Proceedings of the Technology, Mind, and Society - TechMindSociety '18*. ACM Press, New York, New York, USA, 1–6. https://doi.org/10.1145/3183654.3183706

[51] Veljko Pejovic and Mirco Musolesi. 2014. InterruptMe: designing intelligent prompting mechanisms for pervasive applications. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '14 Adjunct*. ACM Press, New York, New York, USA, 897–908. https://doi.org/10.1145/2632048.2632062

[52] Katrina L Piercy, Richard P Troiano, Rachel M Ballard, Susan A Carlson, Janet E Fulton, Deborah A Galuska, Stephanie M George, and Richard D Olson. 2018. The Physical Activity Guidelines for Americans. *JAMA - Journal of the American Medical Association* 320, 19 (nov 2018), 2020–2028. https://doi.org/10.1001/jama.2018.14854

[53] John Rooksby, Mattias Rost, Alistair Morrison, and Matthew Chalmers Chalmers. 2014. Personal tracking as lived informatics. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*. ACM Press, New York, New York, USA, 1163–1172. https://doi.org/10.1145/2556288.2557039

[54] R M Ryan and E L Deci. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *The American psychologist* 55, 1 (1 2000), 68–78.

[55] Suranga Seneviratne, Yining Hu, Tham Nguyen, Guohao Lan, Sara Khalifa, Kanchana Thilakarathna, Mahbub Hassan, and Aruna Seneviratne. 2017. A Survey of Wearable Devices and Challenges. *IEEE Communications Surveys & Tutorials* 19, 4 (24 2017), 2573–2620. https://doi.org/10.1109/COMST.2017.2731979

[56] Paschal Sheeran. 2002. Intention - Behavior Relations: A Conceptual and Empirical Review. *European Review of Social Psychology* 12, 1 (2002), 1–36. https://doi.org/10.1080/14792772143000003 arXiv:https://doi.org/10.1080/14792772143000003

[57] Fritz Strack and Roland Deutsch. 2004. Reflective and Impulsive Determinants of Social Behavior. *Personality and Social Psychology Review* 8, 3 (8 2004), 220–247. https://doi.org/10.1207/s15327957pspr0803{_}1

[58] Suzanne Barlyn. 2018. Strap on the Fitbit: John Hancock to sell only interactive life insurance. https://finance.yahoo.com/news/strap-fitbit-john-hancock-sell-141135104.html?.tsrc=applewf

[59] Lie Ming Tang and Judy Kay. 2017. Harnessing Long Term Physical Activity Data – How Long-term Trackers Use Data and How an Adherence-based Interface Supports New Insights. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 2 (6 2017), 1–28. https://doi.org/10.1145/3090091

[60] Lyndall Thomas and Mark Williams. 2006. Promoting physical activity in the workplace: using pedometers to increase daily activity levels. *Health Promotion Journal of Australia* 17, 2 (2006), 97–102. https://doi.org/10.1071/HE06097

[61] Catrine Tudor-Locke. 2002. *Taking Steps toward Increased Physical Activity: Using Pedometers To Measure and Motivate.* Technical Report. https://eric.ed.gov/?id=ED470689

[62] Catrine Tudor-Locke and David R Bassett. 2004. How Many Steps/Day Are Enough? Preliminary Pedometer Indices for Public Health. , 8 pages. https://doi.org/10.2165/00007256-200434010-00001

[63] Catrine Tudor-Locke, Cora L Craig, Wendy J Brown, Stacy A Clemes, Katrien De Cocker, Billie Giles-Corti, Yoshiro Hatano, Shigeru Inoue, Sandra M Matsudo, Nanette Mutrie, Jean-Michel Oppert, David A Rowe, Michael D Schmidt, Grant M Schofield, John C Spence, Pedro J Teixeira, Mark A Tully, and Steven N Blair. 2011. How many steps/day are enough? For adults. *The international journal of behavioral nutrition and physical activity* 8 (7 2011), 79. https://doi.org/10.1186/1479-5868-8-79

[64] Diane B Wilson, Jerlym S Porter, Gwen Parker, and James Kilpatrick. 2005. Anthropometric changes using a walking intervention in African American breast cancer survivors: a pilot study. *Preventing chronic disease* 2, 2 (4 2005), A16. http://www.ncbi.nlm.nih.gov/

pubmed/15888227http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1327710

[65] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*. ACM Press, New York, New York, USA, 143–146. https://doi.org/10.1145/1978942.1978963

[66] Octavia H. Zahrt and Alia J. Crum. 2017. Perceived physical activity and mortality: Evidence from three nationally representative U.S. samples. *Health Psychology* 36, 11 (2017), 1017–1025. https://doi.org/10.1037/hea0000531

[67] Octavia H. Zahrt and Alia J. Crum. 2019. Effects of physical activity recommendations on mindset, behavior and health. *Under Review at Preventive Medicine Reports* (2019).

## A APPENDIX: PARTICIPANTS TABLE

Table of volunteers who participated in the study.

| Participant N# | Condition | Gender | Age | Profession |
|---|---|---|---|---|
| 1 | TO | Male | 28 | Research Associate |
| 2 | TO | Male | 24 | Student |
| 3 | TO | Female | 45 | Wellness Counselor |
| 4 | TO | Female | 58 | RN in Clinic |
| 5 | TO | Male | 30 | UX Designer |
| 6 | TO | Female | 20 | Student working part-time |
| 7 | TO | Female | 24 | Research administrator |
| 8 | TO | Male | 42 | Operations Manager |
| 9 | TO | Male | 19 | Student |
| 10 | TO | Male | 31 | Entrepreneur |
| 11 | TO | Female | 62 | Project Manager |
| 12 | TO | Male | 29 | Postdoctoral Fellow |
| 13 | TO | Female | 61 | Administrative |
| 14 | TO | Female | 52 | Attorney |
| 15 | TO | Male | 21 | Student |
| 16 | TO | Male | 40 | Facilities Manager |
| 17 | V+T | Male | 59 | Project Manager |
| 18 | V+T | Female | 43 | Childcare |
| 19 | V+T | Male | 24 | Marketing & Business |
| 20 | V+T | Female | 55 | Business Manager |
| 21 | V+T | Female | 19 | Gymnastics coach/Student |
| 22 | V+T | Female | 67 | Retired Executive Assistant |
| 23 | V+T | Female | 63 | Conference and Event Manager |
| 24 | V+T | Female | 47 | High School Teacher |
| 25 | V+T | Female | 68 | Executive Assistant |
| 26 | V+T | Male | 29 | Entrepreneur/Researcher |
| 27 | V+T | Female | 57 | Blogger & Small Business Owner |
| 28 | V+T | Male | 24 | Research Assistant |
| 29 | V+T | Male | 42 | Project Manager |
| 30 | V+T | Female | 26 | Campus Planner |
| 31 | VO | Female | 45 | Research Assistant |
| 32 | VO | Female | 28 | Postdoctoral Fellow |
| 33 | VO | Female | 35 | Executive Director |
| 34 | VO | Male | 50 | Computing Support Analyst |
| 35 | VO | Male | 26 | Program Analyst |
| 36 | VO | Female | 48 | Lab Manager |
| 37 | VO | Male | 66 | Business Manager |
| 38 | VO | Female | 53 | Administrative |
| 39 | VO | Male | 38 | Self-employed |
| 40 | VO | Male | 57 | Engineer |
| 41 | VO | Male | 70 | Clinical Psychologist |
| 42 | VO | Female | 33 | Assistant Communications Director |
| 43 | VO | Male | 24 | Nursing Assistant |
| 44 | VO | Female | 34 | Researcher |

## B APPENDIX: EXERCISE PROCESS MINDSET MEASURE

*Ranges from 1 - 4 (with higher values representing more positive mindsets)*

The following statements are different opinions about exercising. Please complete the sentence by marking the option that best indicates how you feel about exercising.

Exercising is _______.

| | | | |
|---|---|---|---|
| very easy | somewhat easy | somewhat difficult | very difficult |

Exercising is _______.

| | | | |
|---|---|---|---|
| very pleasurable | somewhat pleasurable | somewhat unpleasant | very unpleasant |

Exercising is _______.

| | | | |
|---|---|---|---|
| very relaxing | somewhat calming | somewhat stressful | very stressful |

Exercising is _______.

| | | | |
|---|---|---|---|
| very inconvenient | somewhat inconvenient | somewhat convenient | very convenient |

Exercising is _______.

| | | | |
|---|---|---|---|
| very fun | somewhat fun | somewhat boring | very boring |

Exercising is _______.

| | | | |
|---|---|---|---|
| very social | somewhat social | somewhat lonely | very lonely |

Exercising is _______.

| | | | |
|---|---|---|---|
| very indulgent | somewhat indulgent | somewhat depriving | very depriving |