ActiVibe: Design and Evaluation of Vibrations for Progress Monitoring

Jessica R. Cauchard* **Janette L. Cheng*** **Thomas Pietrzak‡** **James A. Landay***

*Stanford University
Department of Computer Science
{cauchard, janette.cheng, landay}@stanford.edu

‡Univ. Lille, Inria
thomas.pietrzak@univ-lille1.fr

ABSTRACT

Smartwatches and activity trackers are becoming prevalent, providing information about health and fitness, and offering personalized progress monitoring. These wearable devices often offer multimodal feedback with embedded visual, audio, and vibrotactile displays. Vibrations are particularly useful when providing discreet feedback, without users having to look at a display or anyone else noticing, thus preserving the flow of the primary activity. Yet, current use of vibrations is limited to basic patterns, since representing more complex information with a single actuator is challenging. Moreover, it is unclear how much the user's current physical activity may interfere with their understanding of the vibrations. We address both issues through the design and evaluation of ActiVibe, a set of vibrotactile icons designed to represent progress through the values 1 to 10. We demonstrate a recognition rate of over 96% in a laboratory setting using a commercial smartwatch. ActiVibe was also evaluated *in situ* with 22 participants for a 28-day period. We show that the recognition rate is 88.7% in the wild and give a list of factors that affect the recognition, as well as provide design guidelines for communicating progress via vibrations.

Author Keywords

Tactile display; vibrotactile icons; tactons.

ACM Classification Keywords

H.5.2. Information Interfaces and Presentation: User Interfaces: Haptic I/O, Interaction styles.

INTRODUCTION

We are seeing the emergence of wearable technologies with various form factors, which often include multiple sensors, and allow for regular environmental sensing and activity monitoring. Users can receive relevant information and feedback based on their preferences, such as incoming messages and calls, or progress towards a goal, such as the

Figure 1. Person exercising with the ActiVibe system

number of steps walked in a day. Despite their portability, these devices can be disruptive to users and others around them because of the high number of audio and visual notifications they produce.

In contrast, embedded vibrotactile displays alleviate some of those concerns by providing subtler, less intrusive notifications that only the user can perceive [1]. Moreover, prior work shows that vibration patterns can "be consumed at the periphery of a user's attention" [2]. Vibrotactile displays also provide an always-available display that comes in handy in situations where the user cannot take their eyes off their primary activity. For example, if the person is driving, biking, or giving a presentation [3], they may not be able to look at the information received. In this situation, the user receives information without having to explicitly attend to it or query for it, as they must do with existing devices (e.g., by pushing a button on a activity tracker such as the Fitbit).

However, vibrotactile displays often lack expressivity compared to visual or audio cues. Prior works suggest using complex vibrotactile displays to provide richer information, such as different intensities, or using multiple actuators to use the localization of the vibration on the skin. ActiVibe uses a different approach by encoding a vibration pattern in the low-resolution DC motor embedded in today's wearable devices. This is a more realistic approach to designing tactile displays given that the manufacturing costs of and lack of space in wearable devices, such as vibration rings [4], preclude including multiple actuators.

CHI'16, May 07-12, 2016, San Jose, CA, USA
ACM 978-1-4503-3362-7/16/05.
http://dx.doi.org/10.1145/2858036.2858046

In this paper we present the design of ActiVibe, a set of ten tactile icons for communicating progress using a basic vibration actuator embedded in a commercial smartwatch. This paper first describes our design rationale and the evaluation of six icon sets to determine which design choices impact the recognition of the vibrotactile message by the user. The most successful icon set, with a recognition rate of 96% in the laboratory, was chosen as the vibration pattern for ActiVibe. The system was evaluated in the wild for a 28-day period with 22 participants. The longitudinal study validates the success and potential of ActiVibe in naturalistic settings and helped us in developing design guidelines and conclusions on what factors impact the recognition rate and comfort of users.

RELATED WORK

This section discusses prior work in the field of tactile displays and longitudinal studies of feedback systems.

Parameters for Vibro-tactile Messages

Vibrotactile actuators use up to three signal parameters: *frequency* (pitch), *amplitude* (volume) and *duration*. High precision vibrators can control each of these parameters individually and the mechanism executes the command sent by the software with high fidelity. In contrast, it is impossible to simultaneously control both the frequency and amplitude of a low-resolution DC motor [5], like those embedded on most off-the-shelf wearable devices. Additionally, duration cannot be controlled precisely due to the inertia of the motor shaft rotation.

From a perception point of view, amplitude is sometimes referred to as *intensity*, as the perception of amplitude varies depending on the frequency of the signal. The *body location* of the stimulation is also a factor in the identification of the vibration, with a resolution between 1mm and 1cm depending on the body part. Compound parameters are built from these parameters: rhythm (succession of durations), direction (succession of body locations) or patterns (simultaneous body locations).

Our choice of a single low-resolution DC motor constrains the perception possibilities to intensity, duration and rhythm. Given that the control of intensity varies from one user to another or the way the smartwatch is fastened, we use the duration and rhythm parameters only.

Technology for Vibrotactile Displays

Several technologies in the literature allow communicating richer information, such as voice coil actuators [6, 7], pin arrays [8] or piezo ceramics [9]. The design of tactile icons can then take advantage of a richness of tactile parameters to represent structured messages [6-8]. Each parameter of the vibration pattern can encode a piece of information to convey richer messages. Another strategy consists of creating a tactile illusion, such as saltation [10] or a phantom tactile sensation [11]. Previous studies also propose richer interaction using several vibrators [12, 13]. In this case, information is either represented by the location or the pattern formed by multiple tactile stimulations.

Prior work on the design of tactile cues often tackled the lack of expressivity by either using high precision actuators or multiple actuators. Although more complex tactile feedback is becoming more widely available, e.g., in the Apple Watch, our objective is to maintain a high level of expressivity with a simple, inexpensive vibrator more commonly found in less expensive devices. Li et al. [14] propose changing the duty cycle of a mobile phone's vibrotactile motor to create different vibration signals. Some prior work used the semantics of the vibrotactile signal, such as the duration between two vibrations, as representation of a navigation angle [15].

Longitudinal Studies

Studying behavior change adds important requirements to already rigorous HCI-style evaluations: the interventions in question must be evaluated both *in situ* and over a longer period of time. Well-known examples include a three-month study of fitness behavior change supported by the UbiFit system [16], Klasnja's et. al's study of experience sampling method delivery frequencies and their related perceived intrusiveness [17], and a multi-week study of mental health behavior maintenance in the BeWell system [18]. All of these systems were tested in the field with users over weeks or months. We were motivated by this prior research to test how well a large group of people (22 participants) reacted to our ActiVibe patterns over a similarly long period (28 days) so as to uncover and solve the issues that may become apparent in later applications that use this vibrotactile technique to support awareness and behavior change.

ACTIVIBE DESIGN

ActiVibe was designed as a set of vibrotactile icons to represent progress. As performance is generally evaluated as a percentage or as a value on a scale, we created vibrations corresponding to the values 1 to 10, with the objective of representing 10% to 100% progress. Because of our choice of using a single basic vibration actuator, we encode the vibrations using the duration and rhythm parameters only. Since there was no prior encoding of discrete numbers found in the literature using duration and rhythm only, we first had to determine the best encoding pattern for ActiVibe. We designed a total of six patterns (Figure 2) and in our subsequent evaluations used the pattern with the highest accuracy rate for ActiVibe.

Figure 2 shows a visual representation of the sets that were evaluated. Each individual squiggly line represents a single short pulse, while a continuous line represents a longer vibration. We first designed the series of vibration sets (A-E) that were evaluated in a laboratory setting. The results of the first study helped us to design pattern F, which was then compared in a second laboratory study to the best sets from the first study (A, C, and E).

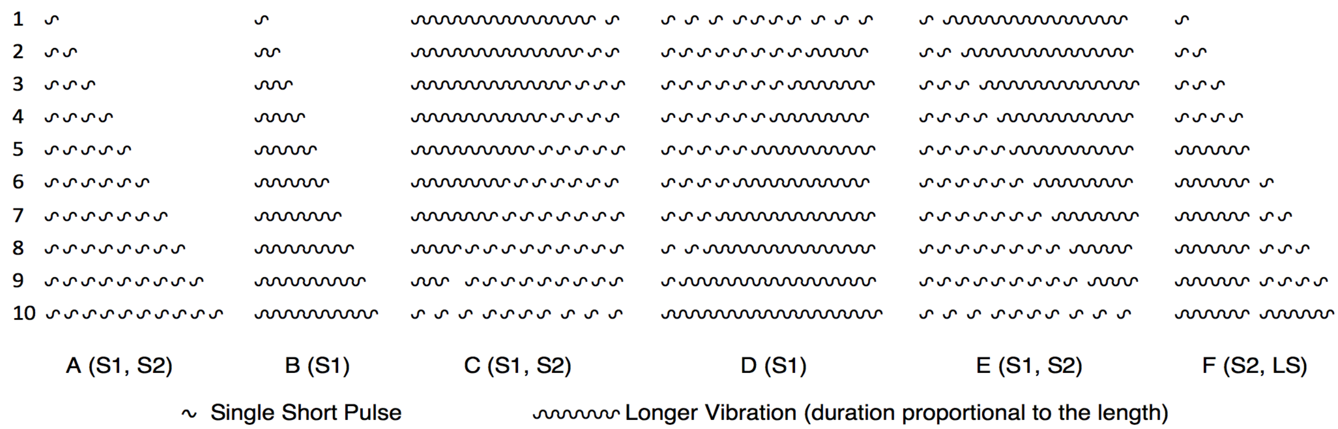


Figure 2. Sets of vibrations designed as candidates for ActiVibe. Sets labelled S1 were used in the first laboratory user study and S2 in the second study. The set labelled LS was used in the longitudinal study.

Design Rationale

Our design is driven by the semantics of the values we want to convey. Our intent is to represent discrete values of a progression. We explored two possibilities: 1) represent the actual value only; 2) represent the value as well as the scale.

Duration Only

Here, the duration of the vibrotactile pattern depends on the value it represents (Figure 1 A & B). The pattern has a short average duration, and thus it may be hard to understand the distance to the end of the event represented. We distinguish two variations. In set A, each value is represented by a series of short pulses, separated by short pauses. In set B, a continuous vibration represents each value with the duration corresponding to the value.

Duration and Rhythm

The disadvantage of representing only the value is that even if the user has an idea of the current value, there is no clue about the distance between this value and the maximum value. Introducing a scale enables the positioning of a value relative to the beginning and the end of a progression.

We represent both the value and the scale of a progression in several ways. Either the current value is represented by a series of short vibrations and the scale by filling the sequence with a long vibration, either before or after the value (Figure 2 C & E), or the current value is represented by a long vibration and the scale is represented by filling the sequence with a series of short vibrations (Figure 2 D).

Set F was defined using the results from the first study as a combination of short pulses and long vibrations.

Set	Information	Representation	Padding
A	Value only	Shorts	
B	Value only	Long	
C	Value + scale	Shorts/long	Before
D	Value + scale	Long/shorts	Before
E	Value + scale	Shorts/long	After
F	Value + scale	Shorts/long	

Table 1: Design space of the vibration pattern sets.

Table 1 summarizes the design space of the vibration pattern sets we developed for our laboratory studies. We were interested in knowing which sets of icons were most suitable for representing progression values, and what is the best precision we can obtain using these representations.

Short vs. Long Pulses

We ran short pilot studies to estimate the shortest vibration pulse that users are able to perceive with the apparatus (Pebble watch), as well as the shortest pause between two vibrations so that users can distinguish multiple short pulses. We obtained 100ms for the pulse and 150ms for the break between two pulses. The longest patterns last less than 3 seconds.

Laboratory User Study 1

We evaluated the ability of the icon sets to represent progress and, more specifically, the discrete values 1 to 10. We believed that participants would be able to match the visual representation to the vibrotactile pattern. We evaluated which patterns had a low error rate and hypothesized that erroneous answers would only occur near the target values.

Methodology

The experiment was conducted using a Pebble smartwatch tightly fit to the participants' left wrist, connected to an iPhone 5S via Bluetooth (controlled by the experimenter).

The experiment used a 5×10×30 within-subjects design, with factors *set* (A-E), *icons* (1-10) and 30 repetitions. The order of *set* was counter-balanced amongst participants. We recruited 10 participants (7 male), 24 to 37 y.o. (μ=28 y.o.). The participants were compensated for their participation with snacks.

Procedure and Task

The experiment was divided into a training phase and an experiment phase for each set of vibrations. Before the training started, the participant was shown a visual representation of the five sets (A-E). They were then sent the vibrations representing each discrete value (1-10) in the current set, in order, twice in a row. At the end of the training session, they chose which image they thought corresponded to the set. In the experiment phase, each value was sent 3 times in a random order. The participant was instructed to say out loud the value they felt out of the current set. The experimenter recorded the answer using the smartphone to ensure the study was based on perception only without distraction from input. Participants had a total of 8s between vibrations to give an answer. If no answer was given on time, the trial was counted as a "miss", and the next trial started. Participants were allowed to take small breaks between each set. Participants filled out a post-task questionnaire after each set and a final qualitative questionnaire at the end of the study.

Measures

We measured: missed rate (MR), error rate (ER), and absolute Difference between Input and Answered value (DIA). MR measures the percentage of missed trials, ER measures the percentage of trials for which the Answered value differs from the Intended value. DIA measures the precision of the answer.

Results

This section presents the results of the laboratory study.

	Set A	Set B	Set C	Set D	Set E
MR (%)	0.67	0.00	2.67	4.00	0.67
ER (%)	18.66	65.66	17.00	28.33	19.67

Table 2. Percentage of MR and ER for each set (S1)

Table 2 shows the results of *MR* and *ER* for S1 sets.

MR: We did not find any statistical significance; these low values show that participants were almost always able to recognize a value.

ER: The error rate is 29.8% overall, ranging from 17% (Set C) to 65.66% (Set B). Kruskal-Wallis analysis shows a significant effect ($\chi^2 = 20.3$, $p<.001$) with set B leading to significantly more errors than the other sets ($p<.005$ with A, C and E, $p<.05$ with D).

DIA: DIA shows how big the error was that people actually made. Our hypothesis was that errors would be close to the target values. Non-parametric Kruskal-Wallis rank sum test shows statistical significance ($\chi^2 = 257.4$, df = 4, $p<0.001$). Set B and D are significantly at higher distances with the following pairwise comparison using Bonferroni p adjustment: (B-A $p<.01$), (B-C $p<.01$), (B-D $p<.01$), (B-E $p<.01$), (D-A $p<.01$), (D-C $p<.01$), (D-E $p<.05$).

Subjective preferences: We found statistical significance in the subjective preferences on how easy it was to identify the values for a given set ($\chi^2 = 21.7$, df = 4, $p<.001$) and on perceived performance ($\chi^2 = 18.3$, df = 4, $p<.01$) with Sets A, C, and E performing significantly better than B and D (Table 3).

	Set A	Set B	Set C	Set D	Set E
Easiness	3.8	1.4	3.3	2.8	3.5
Performance	3.8	1.4	3.6	3	3.6

Table 3. Subjective preferences for ease in identifying the values and perceived performance.

Design Considerations

In this section we discuss the design considerations that resulted from our analysis of laboratory user study 1 and how they informed our next round of vibrotactile icon design and evaluation.

Understanding Vibrations:

All participants were able to identify the visual representation of the sets 100% of the time. Participants mentioned that this visual representation helped them understand and recognize the set. We found that the 2-minute training period was sufficient to understand and identify the sets of vibrations.

Design of Sets for Discrete Values

Sets A, C and E showed significantly less errors and smaller DIA than sets B and D (Figure 2). Participants also preferred sets A, C, and E to B and D. The main design issues we found were as follows:

- Set B: Participants struggled to identify how long each signal was. It appeared that most participants developed a methodology for counting the patterns with the other sets and could not use a similar methodology for set B. This appeared difficult to most users during the training phase.
- Set C: A few participants mentioned that they sometimes counted the long signal in addition to the short pulses and then felt that the signal was one pulse too long.
- Set D: Participants complained about having to count "backwards" as they were counting the number of short vibrations and so counting back from 10 to find the correct value. An equivalent measure to DIA but without the "absolute value" correction shows that participants inverted the results, by giving an icon a value of 4 instead of 6 or 3 instead of 7, for example.

It is worth noting that all of the qualitative data is backed up with the quantitative data.

Methodology for Counting Patterns

Participants found strategies to give meaning to each icon and pattern. Most of those strategies relied on counting the number of short vibrations.

Discrete Values vs. Progress

Although participants used a counting methodology to find the right discrete value, some noticed that the longer continuous vibration was helpful in showing progress. A participant mentioned that they would feel they had achieved a lot when feeling a longer vibration.

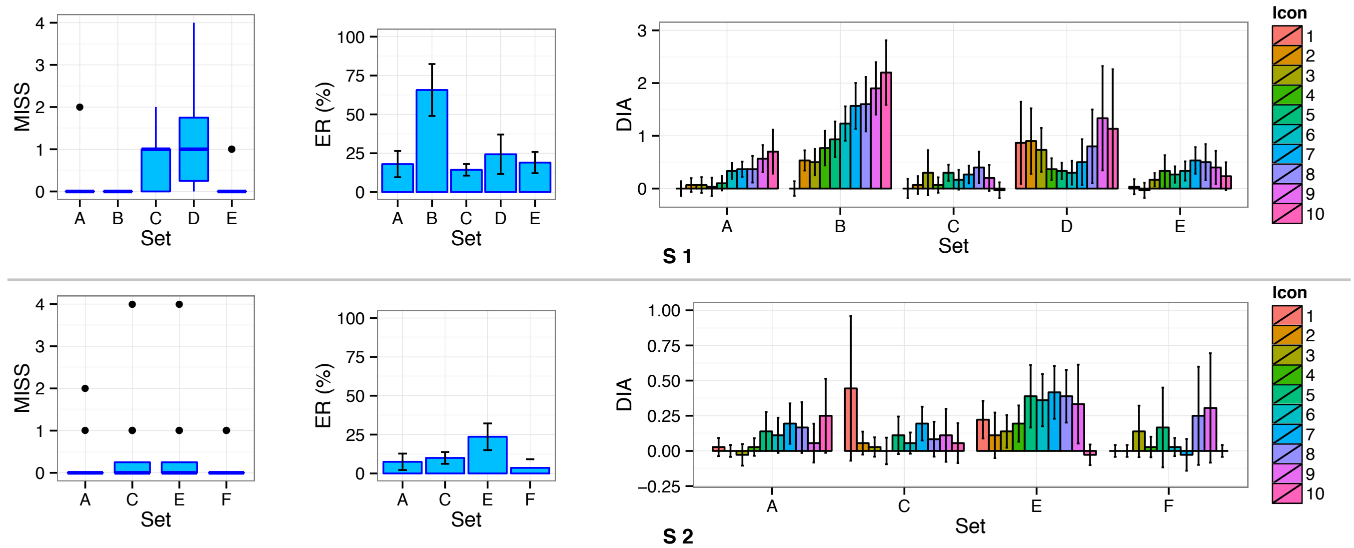


Figure 3. Experimental results. Top: Lab user study 1 (Sets S1) and Bottom: Lab user study 2 (Sets S2). Left: Missed (MR) trials for each set (where the participant did not provide an answer on time). Center: Error rate (ER) per set as a percentage of items that were not answered accurately. Right: Difference between Intended and Answered value (DIA) for each set and icon value with error bars (absolute value).

Design of Pattern Set F

Based on the results of laboratory user study 1, we decided to combine long and short vibrations to represent a value, such as in the Roman numeral system. One to four short vibrations indicate values 1 to 4, and 5 is represented by a long vibration, 6 is then defined as 5 + 1: a long vibration followed by a short vibration. This allows participants to not have to count too many digits, which is more prone to error, while having to discriminate between two lengths of vibration only: a short (100ms) and a long vibration (500ms) with a pause in between (150ms).

LAB USER STUDY 2

Set F was designed based on the results of laboratory user study 1 and was evaluated against the best pattern sets (S1) so we could determine which set to use for our longitudinal evaluation of ActiVibe.

Methodology

This second user study compares the results of Sets A, C, E & F (S2) using the same methodology as laboratory user study 1. We recruited 12 new participants (4 male), 23 to 32 y.o. (μ=26 y.o.).

Results

As per the first study, all participants could recognize the visual representation of all patterns after the study. A single participant could not recognize pattern F after the training phase, but then realized their mistake after the experimental phase.

	Set A	Set C	Set E	Set F
MR (%)	0.83	1.67	1.67	0.56
ER (%)	7.56	10.17	24.01	3.63

Table 4. Percentage of MR and ER for each set (S2)

MR: We did not find any statistical significance in the number of missed trials (Table 4).

ER: Kruskal-Wallis analysis shows a significant effect (χ^2=12.2, p<.01) with set E leading to significantly more errors than the other sets with post-hoc Wilcoxon pairwise comparison (A-C p<.05) and (A-F p<.05) (Table 4).

DIA: Non-parametric Kruskal-Wallis rank sum test shows statistical significance (χ^2 = 13.90, df = 3, p<0.01). Posthoc tests did not show significance in pairwise comparison.

User Preferences: 83% of the participants preferred Set F and 17% preferred Set A.

Figure 3 summarizes the results of the laboratory studies.

Design Considerations

Many participants mentioned that in Pattern C, they liked “the long vibrations [which] helped with concentrating on counting the upcoming bits” [P2]. [P5] mentions “using the long vibration period just to cue that I should prepare to count the number of pulses in the next phase” and [P6] said “It does give me some advance warning for when the numbers are coming up, which I feel could be useful if I am in an activity”. While Pattern C may not be the best choice of pattern, adding a pre-vibration for users to start paying attention to the vibration could be useful.

Laboratory Study Conclusions

Set E is statistically more prone to errors than the other three sets. Despite the lack of statistical significance between sets A, C, and F, set F had lower missed and error rates. With 96.4% accuracy, it was also widely preferred by participants—10 out of 12 chose it as their overall preferred set. We thus decided to use set F as the ActiVibe pattern in our subsequent longitudinal study.

ACTIVIBE LONGITUDINAL STUDY

After designing a pattern to represent progress that was recognized at an extremely high accuracy in a laboratory setting, we tested ActiVibe in a more naturalistic setting over a longer period of time. This study helps us to understand whether receiving regular vibrations is agreeable to users and gives us data to understand what factors are affecting the perception of the signal on a day-to-day basis when participants are not focused on interpreting it or even expecting it to arrive.

ActiVibe Pattern

The ActiVibe final pattern was designed based on Pattern F with a small increase in the durations given that participants will not be paying as much attention to the vibration as they had in the laboratory studies.

Each short vibration lasts 150ms, each long vibration 600ms, and the pause is 200ms long. Finally, the pre-vibration is a long vibration of 700ms.

Hardware

ActiVibe was implemented on a Pebble smartwatch connected via Bluetooth to an iPhone 5 or 6 running iOS 8.0. The smartwatches were lent to the participants by the research team and the iPhone app was installed on the participant's own phone. The Pebble app generated the vibrations and recorded the data, and the iPhone app was used to communicate the data to a server.

Pilot Study

To verify the feasibility of the study we ran an 8-day longitudinal user study with 12 participants (6 female, 6 male), from 19 to 33 years old (μ=21.8). The results of the 8-day study being consistent with the results of the 28-day study, we only report the results from the main study.

Longitudinal Study

This section describes the longitudinal study. The study hypotheses were:

- H1: ActiVibe pattern will be understood with high recognition accuracy in the wild.
- H2: ActiVibe pattern's recognition rate is higher in the laboratory than in the wild.
- H3: The pre-signal will help increase the recognition rate.
- H4: Activities will affect the recognition rate.

Participants

Twenty-two volunteer participants (Table 5) were recruited from a wide range of ages and lifestyles as we wanted to verify how well ActiVibe was working with different populations. Participants had different workday schedules, means of transportation, workout habits, and six had children living at home with them. The participants (11 female, 11 male) between the ages of 20 and 69 years old (μ=39.9, SD=16.0) were recruited through mailing lists and word of mouth. They were compensated $100 for taking part in the study and up to another $100 for complying with the study. The compliance was based on how many daily end-of-day questionnaires they answered and not on the in the moment vibration recognition tasks as we did not want to increase their motivation and hence their attention when performing these tasks.

Num	Age	Gender	Hand	Job Title
P1*	69	F	R	Bakery owner
P2	67	F	R	Physician educator
P3	60	F	R	Administrative assistant
P4	57	M	R	e-business Consultant
P5	53	M	R	IT helpdesk manager
P6	48	M	R	HR director
P7	54	M	R	Publications manager
P8*	50	F	R	Clinical research nurse
P9	52	M	R	Administrator
P10	37	F	R	Digital strategist
P11	39	M	R	Associate director
P12	39	F	L	Program manager
P13	29	F	R	Postdoc (Chemistry)
P14	28	M	L	Developer
P15	28	M	Ambid.	Postdoc and entrepreneur
P16	34	M	R	Unemployed
P17	29	F	R	Events coordinator
P18	20	M	R	Student (Political science)
P19*	20	M	R	Baseball coach
P20	21	F	R	Intern (self sufficiency program)
P21	23	F	R	Neuroimaging research assistant
P22	20	F	L	Student (Business)

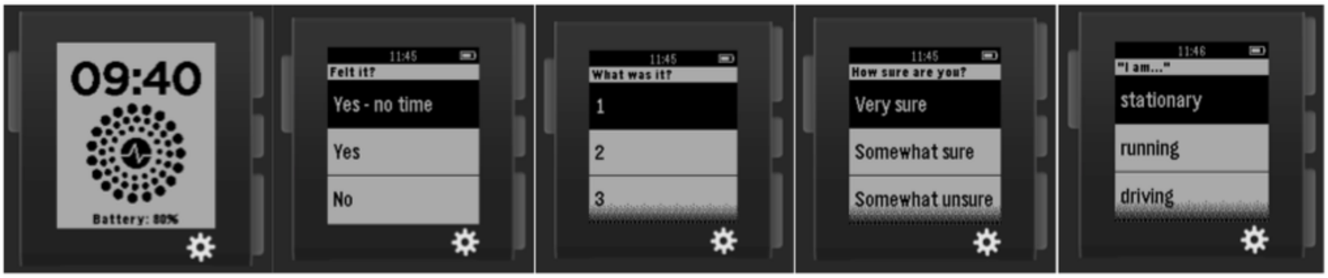
Table 5. Table of Participants including their age, gender, handedness and job title. The data from the 3 participants marked with a * had to be excluded from the data analysis.

Methodology

The experiment took place over a month (28 days). It included several work weeks and weekends. Participants were sent twelve vibrations per day at semi-random times within one-hour window between 7am and 8pm (as per [17]). This schedule helped to cover different activities from the users such as their commute, when they bring their kids to school, their day at work, and some evening activities.

The vibration sent was a random number between 1 and 10 from the ActiVibe pattern. While in the future we will help users monitor progress through an increasing value being sent each time they reach a step towards their goal, for example, we first needed to evaluate whether the pattern could be recognized when the user's attention was not focused on the vibration. We also wanted to overcome the novelty effects by sending the vibration over a long period of time (28 days).

Half of the participants (even participant numbers) received the ActiVibe pattern on its own, and the other half (odd participant numbers) received the pre-vibration pattern first.



(a) (b) (c) (d) (e)

Figure 4. Screenshots of the ActiVibe survey interface. a) Background display showing that the interface is working. b) Appears when the vibration occurs. c) Number of the vibration felt. d) Degree of certainty. e) Activity.

A training session was run in our laboratory to ensure all participants could learn the pattern and that the application could be installed on their phone. The training session followed the exact same methodology as both laboratory studies. The pre-study lasted about 30 minutes and participants were given the watch and a booklet with additional information about the study and how to take care of the watch. At the end of the study, we ran a post-study interview with each participant, which took approximately 30 minutes.

Procedure

Participants were asked to start wearing the watch and not remove it from their wrist from the moment they woke up until 8pm. They were also emailed a link to a questionnaire after 8pm every evening. Participants were asked to charge the watch and their phone every night.

Eighteen participants were right-handed, three left-handed, and one ambidextrous; sixteen wore the watch on the left side and four on the right. Participants were not given any instructions to how tight the band should be as it was not realistic to expect them to wear it tightly for the full duration of the study.

Task

Each time a vibration would be sent, a survey would appear on the watch (Figure 4). Participants did not need to interact with their phone as the full interface was on the smartwatch. Participants were asked to answer the survey as soon as possible, as long as it was safe to do so (this was added after the pilot study when we realized some participants answered the survey while driving). Participants had a total of 5 minutes to fill out the survey before it would time out. They also had the option to dismiss the survey by the press of a single button on the watch.

Interface

The watch had a background display (Figure 4a) showing the user that everything was running well. The display included the ActiVibe logo, the time, and the battery level. When a vibration would be felt, the survey (Figure 4b) would appear. The participants would answer the survey using the buttons on the side of the watch.

The first question asked the user if they felt the vibration. There were three possible answers:

- *Yes – no time*: to inform that the vibration was felt but that the user had no time to answer;
- *Yes*: to input the perceived vibration data;
- *No*: in case the user would see the survey on the watch but hadn't felt the vibration.

If the user pressed: *Yes – no time*: the interface would switch back to the background display; *Yes*: the interface would prompt for the vibration number (Figure 4c); *No*: the interface would directly prompt for the activity the user was doing (Figure 4e).

When the user was prompted for the vibration number they felt, they would use chose a number between 1 and 10 (Figure 4c) and then were asked to input how certain they felt about their choice (with five possible answers between *very sure* and *very unsure* – Figure 4d). Finally, they were asked about their activity when they felt the vibration. The list of activities we used is the same as the ones supported by the iOS activity detection functionality so as to compare the resulting data.

Phantom Vibrations

To account for phantom vibrations, participants had the option to press the back button (left side of the watch) to manually get to the survey screen. They were not informed about the existence of phantom vibrations so as to not make them feel bad about feeling something that they should not have felt. Instead, they were told that there might be some technical problems and if they ever felt a vibration and the survey did not appear, they had the option of making a manual entry.

Daily Questionnaire

The daily questionnaire that was sent to the participants after 8pm was generated based on the answers from the watch surveys on that day. For instance, if the participant answered that they were too busy to answer the survey (*Yes – no time*), they would be asked that night about what they thought they were doing at the time they were busy as well as a degree of certainty.

Measures

We recorded data for the following dependent variables:

- Answer Rate (AR): Percentage of surveys answered.
- Accuracy (ACC): Percentage of trials for which the Answered value and the Intended value are the same.
- Difference between Input and Answered value (DIA): This measures the precision of the answer.
- Activity: We record what activity participants were doing when they received the vibration and whether this affected their performance.
- Answer Time (AT): Time from the vibration to when the user pressed the first button on the interface – within the 5-minute window.
- *Qualitative data*: We conducted a post-study interview with each participant.

Results

This section describes the results of the longitudinal study.

Data Points

All participants but one complied with the user study. We had screened for participants who worked in the day and not at night, given that the vibrations were sent in the daytime only. P19 ended up working night shifts and waking up after the data collection ended on most days, we therefore had to discard the data from this participant.

P1 and P8 had some technical problems where the watch and the phone's Bluetooth connection was broken and had to be re-paired. The post-study interview revealed that notifications were turned on accidentally on the watch so that they ended up receiving vibrations for every single phone call and text message they received, which changed the study conditions from all other participants. Unfortunately, their data had to be discarded too since they did not have the same conditions as other users.

Answer Rate (AR)

Out of all the vibrations surveys sent, 79% were answered (4,944 survey responses) and 21% were not answered (1,295). The distribution of the responses in the answered surveys is shown in Table 6.

	Total	Yes	Yes-no time	No
Answered surveys	4,944	4,615 (93%)	291 (6%)	38 (1%)

Table 6. Distribution of responses for the answered surveys.

Accuracy (ACC)

Amongst the 4,615 surveys answered ("Yes"), 88.7% of the vibrations were recognized properly. The ACC rate is lower than in the laboratory condition (96%), but is still a very successful rate given the naturalistic conditions. The pre-signal did not have much effect on recognition (Table 7).

	Pre-signal	No pre-signal
Accuracy	90%	87%

Table 7. Comparison in Accuracy between the pre-signal and the no pre-signal conditions.

DIA

Table 8 shows the confusion matrix between the vibrotactile icon that is sent to the participant and their answer. The green circles represent the valid answers. The red circles highlight errors for which 10 or more were made.

	User Response									
Icon	1	2	3	4	5	6	7	8	9	10
1	404	9	4		11	1				
2	3	437	13	2	12	3	3			1
3	3	21	405	12	2	6	3	3		
4	4	1	30	383	3	1	10	1	8	1
5	15	11	5	4	442	3	1			2
6	3	9	11		15	415	7			1
7		11	7	2	2	9	440	6	1	2
8	1	2	10	10	7	4	24	381	9	2
9	2	1	4	12	2	3	7	56	383	3
10		13	7	3	5	26	8	1		393

Table 8. Confusion matrix between the icon sent and the user's response across all participants.

We observe two types of errors:

- Duration

We find that icons 1 (single short vibe) and 5 (single long vibe) are sometimes mistaken. We also find that 10 (2 long vibes) is most often mistaken for 2 (2 short vibes) and 6 (1 short and 1 long vibration). This is due to confusion in recognizing the short vs. the long vibration. In the post-study, some participants mentioned finding it sometimes hard to discriminate a long from a short pulse when only one vibration was felt. Extending the long pulse might solve this problem, although it would increase the information encoding time.

- Count

The DIA shows that more mistakes are made for icons with higher numbers of vibrations. Icons 4, 8 and 9 (encoded with 4 to 5 vibrations) present lower accuracy rate than icons encoded with 1 to 3 vibrations. This result is due to users losing count of the number of vibrations and is consistent with the results of the laboratory studies. Yet, ActiVibe still shows a high average recognition rate of 88.7% when the vibrations are sent in a random order.

Phantom Vibrations
We only recorded 11 phantom vibrations, corresponding to 0.2% of the total number of responses, which shows that phantom vibrations did not show any significance in our study. We also found that participants used the "back button" less than 0.5% of all responses to input the value after having originally pressed the "yes, no time" option.

Activities
Table 9 represents the percentage of time participants were performing a certain activity when the vibrotactile survey appeared. The data is collected both from the surveys and from the daily questionnaires. We found that participants were stationary 67% of the time, which is representative of the sedentary lifestyle in today's Western societies. They were active (walking, running, and biking) 13% of the time, and driving 6.5% of the time. Participants most often recorded "driving" when stuck in traffic jams or stopped at a traffic light. Some participants mentioned recording the data while driving. Although we discouraged this behavior, these participants also usually text while driving, feeling safe to do so. Participants chose "Other" for different reasons, such as other forms of exercising, cooking, or even when carrying their kids.

Activity	Distribution across surveys	Accuracy Rate
Stationary	67.01 %	90.5 %
Walking	11.71 %	87.4 %
Other	8.54 %	80.2 %
Driving	6.51 %	84.5 %
Sleeping	4.81 %	---
Running	0.80 %	54.1 %
Cycling	0.63 %	78.9 %

Table 9. Distribution of activities and Accuracy Rate per Activity across all surveys, generated using data from both the surveys and the daily questionnaires.

We observe that activity has an influence on the accuracy rate, with running presenting a much lower accuracy (54.1%) compared to the other activities. Other activities such as driving present a very high accuracy above 84%.

While running, participants could always feel the vibration but with a much lower accuracy rate. Some mentioned that they struggled to keep count while running and some reported that the vibration would be confused with the vibration felt when running. We expect that a stronger intensity of vibration could solve this problem.

Answer Time
On average, the participants answered the survey within 18 seconds of getting the vibration. Figure 5 shows the average answer time for each participant and overall.

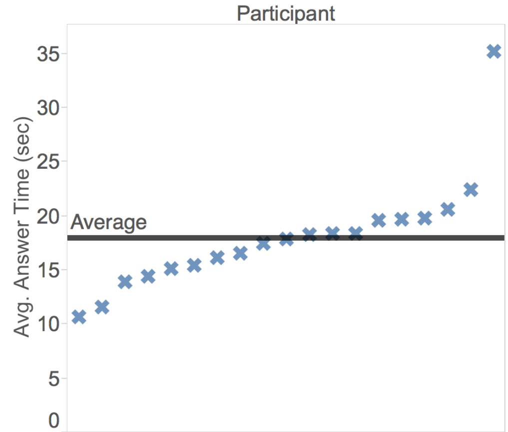


Figure 5. Average answer time for each participant and overall.

Qualitative results
Here we present the qualitative results of the longitudinal study.

- Comfort and Annoyance

We asked participants whether they found the vibrations comfortable. The large majority of them found them very comfortable (mean of 4.8 out of 5). Annoyance is rated at 1.84 (low) and participants mentioned that they found the vibration annoying mostly when they were talking to someone. Participants however mentioned that while the vibrations did not disturb them, the watchband itself was bothersome as the plastic would get hot during the summer and was not very breathable. Figure 6 shows the results of the qualitative data collected on a 5-point Likert scale.

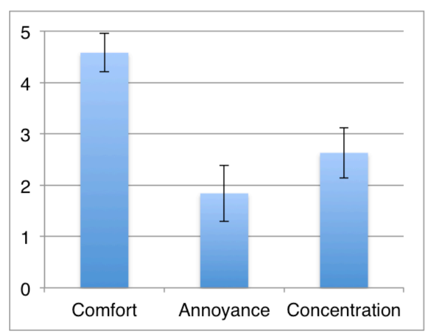


Figure 6. Qualitative data for vibration Comfort, Annoyance the vibrations caused, and how much Concentration recognizing the vibrations required on a Likert scale of 5.

- Pre-vibration

When discussing the pre-vibration, the group who didn't have the vibration answered at 78% that a pre-vibration would have been useful and the group who had it answered at 91% that the pre-vibration was useful to them.

- Activities

Participants mentioned that they could easily recognize the vibrations when they were stationary, walking, driving, or exercising at the gym. Motorcyclists could not recognize the vibration while driving because of the vibrations being muffled under their jacket and because of the motorbike's vibrations. Most participants mentioned that the recognition was harder when they were in the middle of conversations.

Social Scenarios
Participants mentioned pressing the "yes – no time" or even not pressing any button when it felt socially awkward to respond. Most of the time, those were situations when they were discussing something with a person of authority or a customer. Participants mentioned that answering the survey on the watch generated conversation and interest amongst their friends and colleagues.

63% of the people mentioned that they could recognize the value while having a conversation while the remaining 37% mentioned that it was difficult for them and that they had to either focus on the conversation or on the counting. Participants had similar feelings when in a meeting.

Design Implications
This section presents our findings in terms of design implications.

ActiVibe
The designed ActiVibe pattern works with a very high accuracy both in the laboratory (96%) and in-situ (89%). We believe the in-situ rate will increase when the icons are presented in an increasing rather than a random order. The strategy of using both duration and the number of pulses helps limit the total vibration count and the discrimination between one long and one short vibration is easy to perform. The different durations of pulses and pauses will need to be adjusted depending on the hardware used. The intensity of the vibration also has its role to play.

Pre-signal
The accuracy rate does not change much between the two conditions, one with a pre-signal and one without. However there is a strong user preference in having a pre-signal vibration. Therefore we advise that the pre-signal is useful in letting users know when to pay attention to the incoming vibrations.

Sensors
We find that apart from running, other activities do not affect the recognition rate much. However, we find that when the person is in a discussion or giving/listening to a presentation, it is much more complex for them to attend to the vibration. Cognitive psychology research shows that it is difficult to count and use numbers while getting verbal information (e.g., listening to a person talk) [19]. This shows that rather than focusing on activity sensing as has been suggested, we should also look at using speech detection to understand how many people are with the users and who is talking at a given time.

Frequency
The frequency of the vibration did not disturb the participants even over the length of a full month. This shows that participants could receive 10 values for progress within a day without much disturbance. This is important for our usage case of getting feedback about progress towards a goal, for instance a daily step count.

FUTURE WORK
In the future, we will implement ActiVibe as feedback for activity tracking and monitoring and compare it with traditional feedback. We hope that ActiVibe can provide as much vital information to the user as a traditional activity tracker using a visual or audio display while being less invasive. We are also planning to track other goals than just activities, such as learning a new skill, spending time playing a musical instrument, or even taking medications.

We are also interested in tactile ambient displays [20]. They "provide context and awareness for an ongoing situation, available for automatic monitoring or processing." Comparing a notification approach and an ambient approach will highlight situations for which each type of display is most appropriate.

CONCLUSION
We presented the design and the evaluation of ActiVibe both in a laboratory setting and in a longitudinal study over 28 days. ActiVibe is a pattern of vibrotactile icons that represents progress through the digits 1 to 10. We have shown that ActiVibe can accurately be identified with little learning using existing off-the-shelf wearable hardware. During the 28-day study, we have learned that the vibrations did not disturb the users and that social context had more impact on recognition rate and user comfort than the activity being performed, which is what is traditionally tracked by many wearable devices. ActiVibe is the first step towards building vibrotactile displays that can help users become more aware of their progress towards their goals and thus stay on track towards goal achievement.

ACKNOWLEGMENTS
We would like to thank Sunny Consolvo for her advice on running a longitudinal study, Nicolas Roussel for his help on starting this project, as well as Kat Gregory, Edwin Park, and Jessica Zhao on their efforts on the first version of the software.

REFERENCES
1. Eve Hoggan, Andrew Crossan, Stephen A. Brewster, and Topi Kaaresoja. 2009. Audio or Tactile Feedback: Which Modality When?. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI'09), 2253-2256.
2. Martin Pielot and Rodrigo de Oliveira. 2013. Peripheral Vibro-Tactile Displays. In *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services* (MobileHCI'13), 1-10.
3. Diane Tam, Karon E. MacLean, Joanna McGrenere, and Katherine J. Kuchenbecker. 2013. The Design and Field Observation of a Haptic Notification System for Timing Awareness During Oral Presentations. In

Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'13), 1689-1698.

4. Ringly. *Ringly*. 2015. Retrieved 09/15/2015 from: https://ringly.com.
5. Bruce JP Mortimer, Gary A Zets, and Roger W Cholewiak. 2007. Vibrotactile Transduction and Transducers. *The Journal of the Acoustical Society of America* 121, 5: 2970-2977.
6. Stephen Brewster and Lorna M. Brown. 2004. Tactons: Structured Tactile Messages for Non-Visual Information Display. In *Proceedings of the fifth conference on Australasian user interface - Vol. 28*, Australian Computer Society, Inc. 15-23.
7. Sabrina Paneels, Margarita Anastassova, Steven Strachan, Sophie Pham Van, Saranya Sivacoumarane, and Christian Bolzmacher. 2013. What's around Me? Multi-Actuator Haptic Feedback on the Wrist. In *World Haptics Conference* (WHC'13), 407-412.
8. Thomas Pietrzak, Andrew Crossan, Stephen Brewster, Benoit Martin, and Isabelle Pecci. 2009. Creating Usable Pin Array Tactons for Nonvisual Information. *IEEE Transactions on Haptics* 2, 2: 61-72.
9. Jerome Pasquero, Scott J. Stobbe, and Noel Stonehouse. 2011. A Haptic Wristwatch for Eyes-Free Interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI'11), 3257-3266.
10. Hong Z. Tan and Alex Pentland. 1997. *Tactual Displays for Wearable Computing. Personal Technologies* 1, 4: 225-230.
11. Ali Israr and Ivan Poupyrev. 2011. Tactile Brush: Drawing on Skin with a Tactile Grid Display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI'11), 2019-2028.
12. Seungyon "Claire" Lee and Thad Starner. 2010. Buzzwear: Alert Perception in Wearable Tactile Displays on the Wrist. In *SIGCHI Conference on Human Factors in Computing Systems* (CHI'10), 433-442.
13. Koji Yatani and Khai Nhut Truong. 2009. Semfeel: A User Interface with Semantic Tactile Feedback for Mobile Touch-Screen Devices. In *Proceedings of the 22nd annual ACM symposium on User interface software and technology* (UIST'09), 111-120.
14. Kevin A. Li, Timothy Y. Sohn, Steven Huang, and William G. Griswold. 2008. Peopletones: A System for the Detection and Notification of Buddy Proximity on Mobile Phones. In *Proceedings of the 6th international conference on Mobile systems, applications, and services* (MobiSys'08), 160-173.
15. Sonja Rümelin, Enrico Rukzio, and Robert Hardy. 2011. Naviradar: A Novel Tactile Information Display for Pedestrian Navigation. In *Proceedings of the 24th annual ACM symposium on User interface software and technology* (UIST'11), 293-302.
16. Sunny Consolvo, David W. McDonald, Tammy Toscos, Mike Y. Chen, Jon Froehlich, Beverly Harrison, Predrag Klasnja, Anthony LaMarca, Louis LeGrand, Ryan Libby, Ian Smith, and James A. Landay. 2008. Activity Sensing in the Wild: A Field Trial of Ubifit Garden. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI'08), 1797-1806.
17. Predrag Klasnja, Beverly L. Harrison, Louis LeGrand, Anthony LaMarca, Jon Froehlich, and Scott E. Hudson. 2008. Using Wearable Sensors and Real Time Inference to Understand Human Recall of Routine Activities. In *Proceedings of the 10th international conference on Ubiquitous computing* (Ubicomp'08), 154-163.
18. Nicholas D Lane, Mashfiqui Mohammod, Mu Lin, Xiaochao Yang, Hong Lu, Shahid Ali, Afsaneh Doryab, Ethan Berke, Tanzeem Choudhury, and Andrew Campbell. 2011. Bewell: A Smartphone Application to Monitor, Model and Promote Wellbeing. In *Proceedings of the 5th international ICST conference on pervasive computing technologies for healthcare*. (Pervasive Health'11), 23-26.
19. Michael C Frank, Evelina Fedorenko, Peter Lai, Rebecca Saxe, and Edward Gibson. 2012. Verbal Interference Suppresses Exact Numerical Representation. *Cognitive psychology* 64, 1: 74-92.
20. Karon E. MacLean. 2009. Putting Haptics into the Ambience. *IEEE Transactions on Haptics* 2, 3: 123-135.