Persuasive Technology in the Real World: A Study of Long-Term Use of Activity Sensing Devices for Fitness

Fritz, Thomas ; Huang, Elaine May ; Murphy, Gail C ; Zimmermann, Thomas

Abstract: Persuasive technology to motivate healthy behavior is a growing area of research within HCI and ubiquitous computing. The emergence of commercial wearable devices for tracking health- and fitness-related activities arguably represents the first widespread adoption of dedicated ubiquitous persuasive technology. The recent ubiquity of commercial systems allows us to learn about their value and use in truly "in the wild" contexts and understand how practices evolve over long-term, naturalistic use. We present a study with 30 participants who had adopted wearable activity-tracking devices of their own volition and had continued to use them for between 3 and 54 months. The findings, which both support and contrast with those of previous research, paint a picture of the evolving benefits and practices surrounding these emerging technologies over long periods of use. They also serve as the basis for design implications for personal informatics technologies for long-term health and fitness support.

DOI: https://doi.org/10.1145/2556288.2557383

Persuasive Technology in the Real World: A Study of Long-Term Use of Activity Sensing Devices for Fitness

Thomas Fritz, Elaine M. Huang
Univeristy of Zurich
{fritz,huang}@ifi.uzh.ch

Gail C. Murphy
University of British Columbia
murphy@cs.ubc.ca

Thomas Zimmermann
Microsoft Research
tzimmer@microsoft.com

ABSTRACT
Persuasive technology to motivate healthy behavior is a growing area of research within HCI and ubiquitous computing. The emergence of commercial wearable devices for tracking health- and fitness-related activities arguably represents the first widespread adoption of dedicated ubiquitous persuasive technology. The recent ubiquity of commercial systems allows us to learn about their value and use in truly "in the wild" contexts and understand how practices evolve over long-term, naturalistic use. We present a study with 30 participants who had adopted wearable activity-tracking devices of their own volition and had continued to use them for between 3 and 54 months. The findings, which both support and contrast with those of previous research, paint a picture of the evolving benefits and practices surrounding these emerging technologies over long periods of use. They also serve as the basis for design implications for personal informatics technologies for long-term health and fitness support.

Author Keywords
Personal informatics; persuasive technology; activity monitoring; wearable sensing; health; behavior change

ACM Classification Keywords
H.5.2 User Interfaces, H5.m Miscellaneous.

General Terms
Design, Human Factors

INTRODUCTION
Emerging persuasive technology and ubiquitous wearable sensors offer much promise for improving health and fitness practices. Commercial and research personal informatics systems that employ these sensors enable the automated tracking of personal information and activities, such as sleep and physical activity. Along with developments in the sensing technology itself, research has also made great strides in understanding other aspects of these technologies, such as how their design affects activity and behavior [9, 24], how to use visualization to motivate activities and provide awareness [7, 13], how feedback is understood and used [12], and how such technologies should be evaluated [11].

We aim to learn what value these systems provide even after months or years of use, whether and how this value changes over time, and how persuasive personal informatics technologies might be better designed to provide long-term support. We build upon and extend the findings of previous research in the area by contributing a study of in-the-wild use of activity monitoring devices by long-term users in a non-experimental context. Klasnja et al. have identified field studies as an important approach for understanding the impacts of these types of technologies [11]. The recent emergence of commercial wearable devices designed to track and motivate physical activity, such as the Fitbit (fitbit.com) and Nike+ FuelBand (nike.com/fuelband), provides a valuable opportunity to study the naturalistic use of these technologies, and the practices and perspectives that emerge over continued long-term use.

Although not everyone who tries a wearable activity monitor continues to use it in the long term, we focus on those who have done so and integrated it into their daily practices. Investigating the experiences of people who have adopted these technologies "organically" and continued to use them over time offers the opportunity to study certain contexts and aspects of use not possible in shorter-term experimental deployments. It also affords the opportunity to see how findings of previous shorter studies hold over longer-term use. In particular, we visit such issues as the value of system-provided metrics as people's fitness goals and activities change, the challenges of data sharing and identifying relevant social networks, and the ways in which the design of self-monitoring systems influence people's activities and conceptualizations of healthy behavior.

This work offers several contributions to the body of knowledge on personal informatics and persuasive technology. First, it provides a rich understanding of the influence and role of wearable persuasive technology for activity tracking. Second, it reveals perspectives and practices of long-term use that add depth to the findings of previous experimental and shorter term deployments of related technologies. Although in some cases our findings contrasted with those of previous work, the bulk of our findings confirm them, thereby adding richness to our understanding of such technologies as well as lending credence to the findings of shorter term studies. Finally, our

CHI 2014, April 26 - May 01 2014, Toronto, ON, Canada

ACM 978-1-4503-2473-1/14/04…$15.00.
http://dx.doi.org/10.1145/2556288.2557383

work offers a set of targeted implications for designing personal informatics technologies intended for long-term use.

RELATED WORK

Personal informatics systems to support health and fitness are a growing area of interest. Research and commercial systems have been developed to track and monitor information such as weight, steps, sleep, and overall activity. Munson proposes a spectrum of personal informatics applications ranging from *reflective* technologies intended to support insight into one's own behavior to *persuasive* technologies [15]. As Michie et. al. discuss, there is a wide range of approaches to achieve behavior change [14]. Persuasive technologies employ varied strategies for influencing behavior and activities, such as those described by Fogg, most notably *self-monitoring* and *conditioning* [10]. Self-monitoring is one of the most prevalent persuasive technology strategies [11], but technologies often employ multiple strategies.

A variety of monitoring devices have been studied and analyzed for their persuasive influence on practices and behavior. For example Tudor-Locke et al. looked at the use of simple pedometers for measuring and motivating activity [24]. Based on a quantitative synthesis of literature, Bravata et al. found that pedometers in combination with a step goal can significantly increase physical activity [5]. The average duration of the synthesized studies was 18 weeks and Bravata et al. point out that the long-term durability of these changes is unknown.

In recent years, HCI and ubiquitous computing research has produced more sophisticated devices that attempt to persuade using various representations of sensed activity data. The most prominent example is UbiFit, which combined activity sensing with a glanceable visualization of activity [7, 9]. In a field deployment of the system, the researchers found that the visualization helped participants maintain activity levels by providing positive feedback. Other systems attempt to persuade through coaching and advising metaphors. For example, Flowie, a persuasive virtual coach intended to motivate elderly people to walk more was deployed to two participants in an 11-day study and identified types of feedback that were most promising for motivation. Laura, a system with a similar goal, used an animated relational agent as an exercise advisor [4]. Participants increased their walking by 215% during the trial period. Li explored the use of contextual information as a supplement to performance data in the IMPACT system. A multi-week evaluation of the system revealed that contextual information is useful for retrospective interpretation of activities [12].

Researchers have also considered social aspects of persuasive monitoring technologies. Chick Clique, a system geared towards sharing health information among teenage girls, was evaluated in a 6-day deployment, which led to the finding that data sharing could be a powerful motivator [23]. Fish 'n' Steps compared shared and individual use of a fish tank visualization of activity, and found no significant differences in activity levels between the conditions during the deployment [12]. Houston, which combined journaling with automated activity tracking was deployed and evaluated, leading to design implications regarding social pressure and support [6]. Prasad et al. investigated related issues surrounding the use of Fitbit in a one-week study, specifically looking privacy concerns related to data sharing [19]. The findings suggested that people are less willing to share personal demographic information than information collected by the device. Such concerns are consistent with the tension Newman et. al note between sharing health and behavior goals with others in online forums and the need to maintain a positive impression of one's self to the social community [17]. It should be noted that the various trials and deployments of previous technologies described here were short to moderate in length, ranging from 6 days (Chick Clique) to 14 weeks (Houston) in duration. Our study builds upon and generally confirms the findings of these previous deployments, considering similar issues of motivation, practices and social use, with a particular focus they change or persist over long-term use.

Monitoring devices are often limited in the activity that can be sensed, leading to the need for integrating data from multiple sources to get a broader view on health and fitness. Systems that require more effort on the part of users for tracking activities are less likely to be successfully adopted. For example, research by Ahtinen et al. showed that manual entering of health data was burdensome and led to declining use of wellness applications [1]. Tollmar et al. looked at the effects of health information "mashups" that integrate data from multiple sensors and sources and discovered that these combinations allowed people to gain novel insights about their wellness [22].

In general, persuasive technologies intended to spur behavior change are challenging to evaluate because change can only be proven if it persists over a long period of time. Klasjna et al. therefore propose other approaches for evaluating the role of these technologies, including field deployments to understand their use [11]. As further motivation for our research, a recent survey of fifty-four people indicated a high value for long-term information that is often not considered by studies with limited timespan [3]. We therefore believe that studying the long-term use and influence of these devices is necessary for building a more complete understanding of the value of persuasive technology for motivating activity.

BACKGROUND

In addition to the research prototypes mentioned, several commercial wearable activity trackers have been released in recent years. Our study population included users of many such devices, including Fitbit, the Nike+ Fuelband, Jawbone UP (jawbone.com), Striiv (striiv.com), and Bodybugg (bodybugg.com), and many of the participants had tried multiple types of devices or multiple versions of

Figure 1. Nike+ Fuelband (left) and Fitbit Ultra (right)

the same device. Although there are variations in what these devices sense, record, and display, they share certain commonalities.

The devices vary in form factor, including arm bands, wrist bands, and or clip-on models (Fig. 1). Some have passive or interactive displays that can show limited representations of the wearer's data. Several of the systems also provide support for other aspects of health and wellness, such as sleep tracking and food logging. In this work, however, we focus solely on the activity tracking functions.

The devices make use of different sensing technologies, such as accelerometers and altimeters to track movement or activity. They also provide ways of viewing the information through various visualizations on websites, mobile apps, or on the devices themselves. Most systems provide multiple representations of the activity, including concrete measures such as step count, distance traveled, or flights of stairs climbed, as well as abstracted compound representations such as Nike's "fuel" points or FitBit's "activity score."

With respect to persuasion strategies identified by Fogg [10], all of these technologies can be classified primarily as *self-monitoring* technologies. Nearly all of them also employ Fogg's notion of *conditioning* and Michie et al's notion of contingent rewards [14] to some extent in the form of rewards and motivational messages displayed to the user. These features vary in their degree of explicit persuasion, ranging from badges that reflect achievements to fitness challenges and competitions offered to the wearer. Most of them also allow the wearer to set explicit goals, an important strategy identified by Consolvo et al. [8] Nearly all of the systems support data sharing through online communities and social networking features on websites or apps, another potentially valuable strategy identified by Munson et al. [16] and Michie et al. [14].

METHOD AND PARTICIPANTS

To learn about the use and influence of these technologies on people's activity, we conducted a study using in-depth semi-structured interviews. We recruited 30 participants (16 female, 14 male) in various cities in North America (23), Europe (6), and Asia (1) between their early 20s and mid 60s who had been using such a device for at least three months (Table 1.) We chose three months as the minimum length of use as previous experimental wearable activity sensors for fitness had typically considered use periods of three months or less. Most participants had been using their devices for substantially longer; our longest-use participant had been using activity tracking devices for 54 months at the time of interview, and the overall mean length of use across participants was 14.8 months.

Participants were recruited through a variety of approaches including snowball sampling, recruiting emails, and posts to online forums such as the Fitbit community. Participants came from a variety of professions, including a business development expert, a teacher's aid, an attorney, and students, however, overall our interview population slanted towards people in technical professions, such as software developers, and software project managers (Table 1). This proportion may be partly attributable to the snowball sampling approach through which we recruited participants, but also to the fact that the early adopters of these types of devices are likely to be tech-interested individuals.

Twenty-four participants were using Fitbit at the time of interview, four participants used a FuelBand, one participant used a Jawbone and one participant used a Striiv. Five participants had previously used other types of activity monitoring technologies (see Table 1.) It is important to note that our participants' experiences are not necessarily representative of experiences with wearable activity sensing in general, as we are focusing specifically on people who have continued to use the technology for long periods of time. We did not include people who had stopped using these devices and this work is therefore limited in the insight it can offer about why these technologies might fail to be adopted. In addition, as some of our participants were recruited from online forums focused on these technologies, it is likely that our population overall is more active and enthusiastic about their devices than the general population of wearers.

Interviews lasted between 25 and 45 minutes in duration, and were conducted in person when possible (18), and otherwise over the phone or Skype (12). The general interview format commenced with basic questions about the participant's reasons for acquiring or using their device, and the length of time that they had had it. This was followed by more open questions geared towards eliciting anecdotes and experiences. We explored topics such as individuals' daily practices with the technologies, how they used the data the devices provided, their experiences with social aspects of the system, and how their activities and use of the devices had changed over time.

Interviews were audio recorded and transcribed. To analyze the data, we used a combination of inductive open coding and closed coding with codes derived from concepts in related literature to identify important themes. In this work, we present some of the most prominent emerging themes,

Table 1. Summary of participant information.

Parti-cipant	Gender	Activity Sensing Devices Used (Current **Bold**)	Months of Use	Occupation
P1	male	**Fitbit**	3	Usability
P2	female	**Fitbit**	3	Professor
P3	female	**Fitbit**	6	Operations Specialist
P4	male	pedometer,3**Fitbits**	10	Software Developer
P5	female	bodybugg, 2**Fitbits**	18	Student
P6	female	pedometer, **Fitbit**	7	Director
P7	female	2 **Fitbits**	18	Professor
P8	female	**FuelBand**	4.5	Student
P9	female	3 **Fitbits**	11	Student
P10	male	2 **Fitbits**	11	Program Manager
P11	male	5 **Fitbits**	11	Director
P12	male	**FuelBand**	3.5	Researcher
P13	female	5 Fitbits, **Striiv**	9	Attorney
P14	male	**Fitbit**	16	Director
P15	male	**FuelBand**	5	Director
P16	male	**Fitbit**	7	Researcher
P17	male	**Fitbit**	6	Service Engineer
P18	female	4 **Fitbits**	12	Researcher
P19	male	3 bodybuggs, 1 Jawbone Up, 5 **Fitbits**	54	Patent litigater
P20	male	**FuelBand**	8	Retail
P21	female	Fitbug, Striiv, 3 **Fitbits**	16	Psychologist
P22	male	4 **Fitbits**	21	Military Personnel
P23	female	**Fitbit**	12	Teacher's Aid
P24	male	**Jawbone UP**	5	Retail
P25	female	**Fitbit**	21	Web Designer
P26	female	3 **Fitbits**	42	Business Developer
P27	female	2 **Fitbits**	42	Volunteer Coordinator
P28	male	**Fitbit**	8	Finance Counselor
P29	female	2 **Fitbits**	18	Health Professional
P30	female	2 **Fitbits**	22	Loss Mitigation Manager

most notably those that illustrate the dynamic use and influence of these the technologies over time.

FINDINGS

In this section we present the primary findings of our study. We begin by providing a general portrayal of people's use and perspectives of their devices. Subsequently we examine some of the specific ways in which these technologies influence fitness and activity among long-term users. We unpack several phenomena that illustrate how people derive motivation and value from these systems, despite changes in practices and needs that arise over long-term use. In particular we focus on the ways in which metrics, data, and social networking features provided by the system influence people's engagement with their personal fitness and the technologies.

General effects

Most of the participants had integrated the devices deeply into their routines and daily practices, wearing them either all of the time or putting them on first thing in the morning and taking them off just before bed. They described strong **attachments** to them:

"I'm a little obsessed with it. I look at it all the time … I'm always curious, like where I am at what point of the day." (P30, 22 months)

"I feel I find it hard not to wear it." (P29, 18 months)

"There's also I think, some degree of psychological attachment to the thing, [losing it is] kind of like when you lose your cell phone. It's weird to have something that you have with you all the time and you're constantly sort of playing with disappear, like you've lost something significant." (P19, 4.5 years)

Continued and routinized use of the devices did not necessarily mean that wearers stayed consistently enthusiastic about them. Although some participants remained excited about them, others indicated that early novelty had worn off and given way to more moderate attitudes despite continued use:

"I liked it a lot in the beginning and I think now I've sort of fallen out of love with it. You know it's like you've been married for a long time and… it's all right but the excitement has gone." (P21, 16 months)

"That might be something that's changed since I've had it a while. I don't really feel like an urgent impulse to meet my goals and have it validate my activity. I do wear it and I do look at the numbers, but it doesn't necessarily affect my feelings day to day that much at this point." (P27, 3.5 years)

Long-term users developed a high degree of **awareness** of the "value" of their activities in the context of the measurements provided by the systems. Participants revealed a detailed understanding of both routine and non-routine activities:

"I know that if I stay home and work all day … I'm only going to hit like, you know, 3,000 or 4,000 in the apartment. And I know if I'm [at work] … I could get to 8000 really easily. So I'd like to be over 10,000, maybe around 12,000. You know, and then I feel really good about myself if I hit 15,000 or 20,000. But that doesn't really happen; those are the days I don't do any work." (P5, 18 months)

"When I went skiing I knew that I was definitely going to be exceeding my goal, so... I manipulated it a litte bit so that I could say, okay if a regular day of skiing, let's say, is 4000 points, I'm going to set my goal to maybe 4,300. So that way I knew that if I

felt exhausted by the end of the day, I should have reached my goal." (P20, 8 months)

Participants described the **immediate impacts** of the devices on their activities, such as seeing a low step count and then walking to a shop to pick up coffee rather than to a nearby kitchen (P12). It was clear that the numerical feedback for most participants could reinforce and motivate activities, as in one particularly extreme example:

"I was at I think 17 flights, and I thought, oh, I could get to 25. I just started walking up and down the steps while I was reading my Kindle. I hit 25 and I was just getting warmed up and I thought, well, maybe I'll see if I can't hit 50. At 50 I was getting warm, but I wasn't tired … so I kept going. After I hit 100 floors, I decided okay I probably ought to cool down now because it was 11:00 at night … and I just walked at a slower pace while I was reading for the next hour, and I ended up having my steps up over 20,000 and my stairs at 100 that day." (P17, 6 months)

Aside from looking at immediate effects on activity, much persuasive technology research has focused on supporting and evaluating longer term behavior change and health effects, a challenge that has been identified as difficult to measure and assess in less than a year's time [11]. Although explicitly evaluating behavior change was not the goal of this research, we were able to gain some insight as to how long-term users viewed their own changes in practice. Most participants reported that the use of the device had motivated or helped them make **durable changes**, such as walking more, taking the stairs, or standing while working rather than sitting. Twenty-five participants explicitly mentioned changes in their routines or behavior that they felt were lasting and profound:

"So initially it was, oh, okay. Your practice needs to change a little bit if you account to get this extra health benefit in. I've figured out ways to do that, and now they're part of my daily practice." (P7, 18 months)

"For 15 years I had an unhealthy BMI. I could not figure out how to lose weight. I was just doing the same old thing, eating the same old crappy fast food. And once I got this FitBit…it just really opened my eyes to a healthy lifestyle. I've done like 180 degree turn to where I was before." (P22, 21 months)

Although we also were not evaluating direct effects on physical or mental health (e.g., changes in weight, blood pressure, self-esteem), a few participants also did attribute **changes in health and well-being** to using such devices:

"I like that [the Fitbit is] driving me to stay active enough that I'm consistently losing weight. I have not gained anything, and I've lost at least half a pound every week since I started doing this. And that for me is pretty much almost a direct result of having the – it's not really accountability, but it is." (P3, 6 months)

"I'm much more active now I never used to do the walking thing or - I even started running so, something that I'd never thought I'd ever do... I added that because I started becoming more active because of the Fitbit... I kind of built the confidence after - 'cause I had lost a significant amount of weight so I felt more confident to get into those types of activities." (P25, 24 months)

We cannot necessarily assume a causal relationship between device use and health and behavior effects, even though the participants often credited the devices with motivating these changes. The fact that people were willing to acquire and use the devices suggests that some participants may already have been motivated to make changes to their practices. What can be surmised, however, is that these devices and their associated services were able to influence and provide support for activity awareness and increased activity in general.

Motivation and reflection

Munson draws a distinction within personal informatics between technologies that are intended to be *persuasive* and technologies that are intended to be *reflective* [15]. Either type can motivate behavior change, though the former *explicitly* promotes certain behaviors through its design. The commercial technologies have been designed with some features geared towards persuasion in mind, such as the inclusion of goals, rewards for achievements, and challenges to engage the wearer in specific activities and behaviors. Motivation and reflection were both clearly objectives in people's decisions to acquire the devices:

"I got [my FuelBand] precisely to motivate me to be a bit more active" (P15, 5 months)

"In my case, what I wanted to understand really was, well, how much I was actually walking per day. I had been in rehab since an auto accident... this was part of... trying to learn and get a better feel for just how much I was generally being active during the day." (P11, 11 months)

We expected that initial motivational effects might eventually diminish in the longer term as the novelty of the devices wore off, and that people who continued to use them might do so primarily because they found value in their reflective benefits, or simply in having the data recorded. This change was noted by some participants:

"I pay less attention to it now than I did when I first got it. And maybe that's because by now, I sort of know my patterns. I know that if I go to work and I walk from the train station and I have a normal day at work, and then I walk home from the train station, I know that I'm at about somewhere between like 8,000 and 9,000 steps or something like that." (P2, 3 months)

However, it was interesting to find that even over months or years of use, participants still experienced what they perceived to be motivating effects of the devices.

"I really only had one period [of disuse] and I think it was because I knew I was being pretty sedentary at that time. I wasn't feeling well and was really fatigued and maybe I just didn't want to see the numbers. So … part of the reason I continue to wear it is it does motivate me to maintain a certain activity level. (P27, 42 months)

"I go through between, let's say, 10,000 and 11,000 steps [daily], and I aim for 10 flights of stairs. So I know that if I'm gonna fall short of that... I do work a little harder. I'll run up and down the stairs in my house a few times. So it just motivates me to reach that level. (P23, 12 months)

Accounting and getting credit

The numerical feedback provided by the devices was generally motivating for participants, manifested in their desire to meet step goals, or their pride in achieving high numbers. This appeared to be the case regardless of whether the numbers were concretely understandable, such as counts of steps or flights of stairs, or abstracted and not fully understood, such as the FuelBand's fuel points:

"[My goal is] 2,000 Fuel points... I don't really know what that means, but that's roughly what I try to reach. I get it 80, 90 percent of the time." (P12, 4 months)

"I definitely find myself, 'Oh I'm not going to go to bed until I hit my 3,000.' I stay up and do some Xbox sports or something like that; just to do something until I can hit my goal… I've never seen [what fuel stands for] – And people ask that a lot. They're like, 'What does that mean?' I'm like, 'Yeah it's some weird Nike proprietary thing that takes the other activity, not just the steps that you take.' But no, I don't know what it means. So I'm working in ignorance on that point." (P15, 5 months)

Many participants expressed a strong emphasis on the accounting aspect of the devices in a way that suggested that getting system credit for one's activities was an immediate objective that in some cases even overshadowed the underlying motivation to be fit or active:

"I hate when we go out for the day and [I forget my Fitbit]; we'll talk about how it's wasted steps and it's wasted stairs. And she's like, she doesn't get it, 'cause she's like, 'It's not wasted. Your body still knows that you're moving.' But for me, if it's not on I might as well just be sitting down." (P3, 6 months)

"It's like I'm missing all these opportunities [because I forgot my Fitbit today]. So I have to rush home. I'm supposed to meet someone at 5:00... I'm actually gonna go home and get my Fitbit and put it on [first]." (P18, 12 months)

Participants described instances in which they forgot their devices as "annoying" or "irritating" because they felt they were not getting proper credit for their activities. In some cases, the fact that their activities were not being accounted for substantially affected their enjoyment of their activities, despite the fact that they were still getting exercise:

"We went to the Space Needle and we were walking up these stairs and then we were in Pike's Place and we were doing those hills and it's just, it almost puts a damper on my day, where I'll just say lots of times, 'Damn it, I wish I had my Fitbit. Damn it, I wish I had my Fitbit.' My daughter is like, 'Who cares? We're having a good day' and I'm like, 'I don't have my Fitbit.' You know, it's like awful. And my nine-year-old is like, 'Hello, it's still fun.' But I don't feel rewarded for all the work I'm doing, which is kind of – it's kind of a negative in that respect, is that I let it impact my day too much if I forget it." (P3, 6 months)

Although the numerical feedback is intended to provide a representation of the wearer's physical activity, participants did not treat it purely as such, and the discrepancy between overall activity and what the systems could account for often had an influence on people's activities. Some participants described favoring activities that would be accounted, and avoiding activities that they knew would not be accounted:

"I've noticed it doesn't pick up when I go backwards on my elliptical; it only picks up when I go forward, and then that makes me not want to go backwards at all on the elliptical." (P5, 18 months)

"Mine doesn't measure staircases which would be nice, and I think for the only reason that it would probably get me to [take the stairs] … if I knew it was being measured … once you know that there's a device that's monitoring your steps and things, you're a little bit more conscious about actually using it and triggering it." (P11, 11 months)

This discrepancy between activity and representation often became greater in the long term as people's fitness practices evolved. Ironically, devices that initially helped foster engagement in fitness sometimes became too naïve to support increasingly sophisticated fitness priorities. For example, P21 found that her Fitbit better served her needs early in her use, stating that, "It's very much something that's very good for people who don't do any exercise to get them moving." After adding new activities as her fitness increased, however, it failed to measure the activities that had become important to her:

"I'm putting too much time getting up to 15,000 steps a day when I'd be better off going to the gym and actually working out with weights… It's all focused on walking. So what I think it's really good is for people who aren't very fit to get them moving, but I think actually it had a negative impact on somebody like me who's going to the gym doing serious weights five days a week." (P21, 16 months)

Goals and rewards

Goals and rewards within the context of technologies to motivate physical activity are highly complex [8]. Our findings illustrate this complexity how system rewards become part of an ecosystem of rewards over use.

The devices used by our participants all offered some form of explicit rewards or goals to be met, which we refer to as "system rewards." For example, in addition to the customizable goals, the Fitbit website incorporates a system of achievement badges, such as one for walking a "lifetime distance" of 500 miles (i.e., 500 miles since starting to use Fitbit), or for climbing 200 flights of stairs on a single day.

Eighteen participants explicitly stated that system goals and rewards influenced on their personal activity and fitness goals. Goals could provide a benchmark that allowed people to contextualize their own activity or they could serve as a target level of activity, and most said they did not edit the systems' default goals:

"I've just left [the system goal set] at 10,000 because it seems kind of reasonable like your worst case scenario is you're just under 10,000 and you should go do something, so I thought 10,000 worked pretty well." (P18, 12 months)

"The 10,000 steps is just a rule of thumb I've been using 'cause I read somewhere if you do 10,000 steps you can maintain a pretty -- I don't even remember where I read it…. I don't know, but it

sounded reasonable to me so… I haven't changed that in the time I've been using it." (P14, 16 months)

System-based rewards also exerted a strong influence on people's personal goals, often to such an extent that the system goals seemed to supplant underlying goals of improved health or increased fitness. In particular we uncovered a game-like phenomenon which we termed "number fishing," in which participants reportedly engaged in activities explicitly for the system rewards:

"There was one time that I came home, and it had been a particularly active day, and I think I was at like 33,000 steps or something like that, and I was just like, 'I'm just going to walk around the house until I get 35.' So I did that. I just literally walked in circles around the house picking up things and moving things around." (P2, 3 months)

"I think the badges are a great way to get, you know, 'cause you're – if I did like 60 flights of stairs and I know at 75 I'm going to get something for it, you know, do the extra steps, which is completely ridiculous, but it works." (P4, 10 months)

The quote from P4 above indicates a perspective that rewards such as badges are a somewhat trivial, albeit effective, manifestation of the fundamental health-oriented benefits of the system. This acknowledgement that achieving numbers or badges was an artificial goal was common among our participants, even though nearly all of them indicated that they perceived the system goals or rewards as valuable. In many cases, system rewards such as badges and numerical goals were desirable enough that people wanted them even when they knew they were not accurate representations of what they had accomplished, such as in the case of false positives or system gaming. One participant described an incident in which he discovered that construction in his building was registering as steps on his Fitbit, and how the resulting incorrectly high step count and potential to achieve a reward still served as motivation:

"I left my Fitbit at home, and I got this email [at work] … saying, 'Congratulations, you just got a 30,000 step badge!' I'm like, 'I left it at home. What happened?' And I went home, and it was at like 32,000 because of this nailing into the wall. Now I will tell you that that caused me to go put it in my pocket and then go walk out and do 3,000 more steps so I would get the 35,000 badge, so I admit to that, okay?" (P11, 11 months)

Over time, participants also developed "reward ecologies" for themselves that went beyond the explicit system goals and rewards by incorporating real-world rewards or other tools. For example, one participant used her Fitbit in conjunction with an app for tracking and managing calorie intake information to give herself rewards that were entirely external to the system:

"I like using my Fitbit 'cause I want to walk a lot and do the stairs so that when I get home I can have a snack. Like it kind of drives my getting a snack in the evening, and if I don't walk enough and I don't earn enough extra calories I don't get a snack in the evening" (P5, 6 months)

Participants also found other ways to convert their activities into external rewards. For example, P1 used a Striiv feature that allowed her to make donations to charities based on meeting fitness challenges. Another participant used a service called EarnedIt to amplify her rewards:

"[EarnedIt] will take all your Fitbit activity and give you prizes from it... it will look at how much activity came from like mild activity versus rigorous activity and it will give you points, and then you can use those points to buy stuff. So you get free stuff. It's pretty cool." (P6)

Sharing of data and social effects

Activity monitoring technologies often support various ways of sharing data and viewing others' shared data. The study of widely adopted commercial technologies allows us to examine how these practices play out among organically occurring social networks in which people opt in or out of sharing based on their needs or motivations.

Not all participants made use of the social features of the system or shared their data. Some simply lacked interest in sharing data with others, and a few participants expressed concerns the personal nature of their data, although this was not common. There were two general concerns about privacy, the first about having personal data uploaded to a server, and the second that health data, and in particular weight, would be visible to others:

"For me it's very much about losing weight, and that for me is very personal. So I don't post a lot of things or share a lot of things." (P3, 6 months)

However, among the 15 participants who explicitly mentioned using social features of the systems, the motivating effects often proved to be highly durable, offering another set of potential goals:

"I think the key motivation was to lose weight in the beginning. Now, as we move on, I think the greater motivation is competition with friends in terms of how many steps or how many active minutes in the FitBit world, that kind of thing." (P19, 4.5 years)

"When I'm on the website and I see who's stacking where, if I'm close to one of them… I'll make an extra push to take it a little bit further every day so that I can try to get my average up and beat them." (P17, 6 months)

One issue that arose in the long-term use of these devices and their associated social features was that of **finding the right community** to motivate oneself or otherwise enhance the experience. Previous systems have proposed on leveraging a person's existing social network for motivation [6, 23] Among our participants, however, it was rare that sharing data with a person's real-life friends and family proved to be highly motivating over time. Some participants had one or a few friends with whom they regularly shared data but this was generally only motivating when the friends had similar patterns of activity. It was more common that people described the attrition of their friends or colleagues over time, or ways in which they failed to motivate as a community:

"Most of [my co-workers] got [a Fuelband]. And now it's only a select few that still use it on a day to day basis." (P20, 8 months)

"Firstly, I don't know anyone [who has a Fitbit]. I mean, [my son and daughter-in-law], they're the only two I know who have them, and their numbers are like off the charts. So you know, so I wouldn't want to compare my numbers with them." (P23, 12 months)

People who made extensive use of social and sharing features of the system to compare data were more likely to do so with people they had met through the online communities. Interactions with friends made through the online social networks often proved more motivating or engaging because participants perceived them as better "peers" for comparing activity data, a finding that relates to Peng and Hsieh's studies of online gaming friends [18]:

"I have a few [real world] friends that have them and we don't compete much because I'm generally in the very high step category and they're more in the reasonable step categories, but the FitBit website has a community that has all kinds of groups that show leader boards. So it's competition among FitBit friends from all over the world that I will never meet in real life." (P19, 4.5 years)

"You know, there's some of 'em [Fitbit friends] that I've never ever talked to, and then there's a few that we've talked a little bit [online]… But, you know, having even the ones we've never talked on there… it's like seeing somebody at the gym working out. And they're over there working hard. So, just seeing that, you know, 'Hey, if they can do it, and maybe I can do it a little bit better.'" (P28, 8 months)

Even though the online communities allowed people to find peers with whom to share data, finding the "right" community could still prove challenging:

"It's hard to figure out who to compare against, or who to be a right set with. I mean, there's some sort of ridiculous people on there, and they're like they're walking … 50,000 steps by the time they get to work… And then there's people who are … trying to lose 50 pounds or more, which is pretty dramatic as well… So, I don't know how it can filter better so that you can find your peer group in a more meaningful way." (P26, 3.5 years)

Although real-life friends and family were often less motivating than online contacts in terms of comparing or competing, connecting with them through the systems' social networking features proved to be valuable in other ways. Participants described making use of shared data to provide support or maintain awareness, much as they might with general purpose social networking sites. For example, one participant talked about the awareness and peace of mind he gained from sharing data with his elderly mother:

"It's actually kind of cool having my mom on it, right, because she's up in her late seventies now. I kind of worry about that." (P10, 11 months)

P10 described calling to check on his mother when he saw that she had not uploaded data, and discovering that she had hurt her knee. They subsequently used her data to make sure she did not go over 5,000 steps per day while healing.

P18 talked about being able to use other people's data to know when to offer encouragement and support:

"[My sister]'s just finally now getting back to walking after a year of surgery and everything, and when I see she's down at 3,000 steps I'll send her an email and say, "C'mon, you need to get out there. Go take a walk." (P18, 12 months)

DISCUSSION

Our findings contribute new knowledge about the use of wearable activity sensing technologies and evolution of practices surrounding them among long-term users in naturalistic contexts. In this section, we further contribute a set of considerations for the design of such technologies for long-term support, and situate our conclusions within the landscape of previous findings.

Designing for long-term support

As persuasive technologies become increasingly common some key questions remain: What constitutes the success of a persuasive technology? Does a successful persuasive technology continue to offer value by supporting and scaffolding the behavior over time? Or does it serve as a 'gateway' technology [21], training and routinizing new practices to the point that the technology itself is no longer necessary? We believe that both approaches have merit as design goals. Technologies can be intentionally designed such that their success obviates them, or designed to provide continued support and persuasion. Although many people who use wearable activity monitors eventually stop using them, our findings suggest that some segment of the population continue to derive value and motivation from them even after many months or years of use. Participants reported achieving goals such as weight loss or increased fitness, and talked about changes to their routine and activity levels that they believed were long term and durable, but still relied upon recording and viewing their activity data, and sharing data with social networks. These findings yield several implications for the design of persuasive technologies intended to provide *long-term support in addition to prompting behavior change*:

Motivating maintenance as well as change: Persuasive technologies often focus on prompts or cues to introduce changes or improvements in behavior. Although these aspects have proven to be valuable, tools should also offer explicit motivation for the *maintenance* of practices or achievements, as maintenance is critical for durable behavior change [20]. As our findings indicate, long-term users of wearable sensing technologies often rely upon their devices to motivate a consistent level of activity or provide awareness of a baseline, although the system's rewards for doing so are less explicit. Participants found the continued awareness of their data valuable and motivating. However, it is worthwhile to consider how designs can provide more explicit support for behavior maintenance. For example, a system might reward the wearer not only on how many steps she has achieved in a given day, but also on how many days in a row she has managed to reach that goal.

Supporting the identification and evolution of appropriate social networks: Among long-term users, engagement with social networks often provided a source of

motivation and kept interaction dynamic. However, finding the right people or communities with whom to share data often proved a challenge. Relying on real-life friends and family alone may be too naïve an approach when people's goals are as individual and dynamic as those of our participants. Existing systems support the formation of groups with specific interests, attributes or goals, which offers a step in the right direction. Future designs should consider adding more support for finding relevant and motivating communities, and for transitioning between communities or dynamically evolving communities as people's needs and practices change. For example, a system might analyze the activity of a person and find and suggest groups that have proven to be motivating (i.e., have resulted in improved fitness) for people with similar patterns.

Supporting changes in activity and metrics: The type of tracking and measurement provided by the tools influenced how people thought about their activity, apparent in their fluent translation of activity to metrics such as steps or points. Even so, people's practices sometimes veered away from system metrics as they became more engaged in fitness. Metrics and data that had been valuable early in use could become insufficient over time; activities that became important to people as their practices changed, such as weightlifting and yoga, were no longer reflected by the system's metrics. This resulted in a mismatch between the wearer's priorities and the system's support. Designers of systems should consider ways in which measurements and tracked activities can be augmented or evolved over time, or ways in which different sensing technologies can be integrated into the overall ecosystem of support tools. For example, systems might be designed to collect and track more general data, such as steps, early in use. Subsequently they might adapt to track more specific types of activities, perhaps even integrating dynamically with sensors in the environment or technology-augmented fitness equipment.

Supporting the evolution of rewards: The commercial technologies used by our participants all included a built-in system of rewards. It was, however, interesting to see how these system rewards fit into a larger ecosystem of rewards that people created for themselves, including "real world" rewards. Although system rewards support reflection and offer motivation particularly early in usage, providing a variety of rewards and changing rewards over time seem to be valuable in motivation. It may be worthwhile to consider a greater scope of reward types, and ways of adding concrete 'real world' rewards in addition to system-specific symbolic rewards, including dynamic or evolving approaches to rewards to maintain engagement. For example, a system might allow a person to specify a desired future gift to himself, and then integrate with personal budgeting software that would allow for savings to be put aside towards the gift based on the person's fitness activity.

Comparisons with previous research findings

Our findings build upon and extend previous work, supporting and contrasting with findings of experimental deployments of health and fitness sensing technologies. The general value that people attributed to these devices confirmed the most fundamental findings of several previous deployments, such as those of Fish 'n' Steps [12], UbiFit [7], and Chick Clique [23], as well as research by Fogg [10], about how providing data about people's physical activity raises consciousness of increases activity at least in the short term. Most of these studies do not make claims about long-term behavioral changes as trials generally lasted from a few days to a few months. Their findings held true for our participants, and many of them also attributed what they believed to be long-term behavior changes to the devices. This should not be taken, however, as an indicator of their potential for prompting such changes for the general population as we did not study use by people for whom the technology failed to motivate activity. Additionally, some of the shorter durations of use represented in our study (as few as 3 months) may be too short to inspect for real behavior change [20].

Our study also revealed participants' extreme focus on numerical goals and numerical data. These findings echo Li's claim that people focus on numbers rather than the activities the numbers represent [12], as well as Albaina et al.'s findings which indicated that numbers were motivating. Our findings regarding the importance of numbers, accounting, and accuracy lend credence to Consolvo et al.'s design requirement that appropriate credit for activities is essential for encouraging physical activity [6]. Our work extends this knowledge by illustrating how important and embedded these representations can become in daily practice, even after very long use.

Although our findings regarding the practices of long-term wearers generally confirmed those found in previous shorter-term studies, some of our findings contrasted with them, for example in how wearers responded to system goals and rewards. Consolvo et al. conducted a survey asking people what kinds of fitness goal-setting strategies they would want to use, and determined that self-set goals or goals chosen with the help of the user were viewed more positively than those determined entirely by others [8]. Although we did see some customization of system goals, our participants overwhelmingly left the default goals unchanged, and many used them as personal daily targets. This contrast may be due to the fact that the proposed goals in the survey were formulated in terms of workout activities, whereas the system goals offered by the devices in our study are achievements expressed in terms of data collected by the system and therefore seem a natural fit.

The system rewards, such as achievement badges, also proved to be motivating and satisfying for many participants. In contrast, Munson et al.'s evaluation of an activity tracking system that incorporated rewards and journaling found that the rewards were not deemed valuable and did not affect activity [16]. Again, however, our findings need to be considered in light of the fact that we were learning largely from the experiences of people who

had adopted and continued to use these devices, and their attitudes towards these rewards may not generalize.

Regarding the social features of these technologies, our study uncovered varied practices and benefits. Previous experimental deployments have also touched upon aspects of data sharing, such as with whom people would want to share data [15], and the social interaction that arises from sharing among friends [6]. In our examination of long-term users of these technologies, we found evidence that early practices of sharing data with friends and family are likely to change over time, and that people go through a process of identifying relevant people with whom to share data. People's evolving goals and practices further complicate this challenge. From a motivational standpoint, finding peers with similar activities or goals, whether in one's real world or online social network, is more valuable than simply having a personal connection. We believe that aspects of social use can be more fully addressed going forward by looking at how practices emerge within populations of in-the-wild adopters and their naturally evolving social networks.

CONCLUSION

The recent widespread use of commercial technologies for activity sensing and monitoring provides a valuable opportunity to gain insight into the naturalistic use and effects of these ubiquitous computing systems. In this work, we examine the impact and influence of tracking devices on long term wearers' activities and attitudes. Our findings reveal that despite changing goals and practices over time, some population of wearers continue to derive value and motivation from the technologies. This suggests that it is worthwhile for designers of personal informatics and persuasive technologies to consider how to provide long term support for healthy behaviors, in addition to offering initial motivation for change. We believe that long-term wearers have needs and practices that are different from those in the initial weeks of use. By looking at the changing use of metrics, data sharing, and rewards, we provide grounded insights about how systems can evolve alongside wearers and provide continued benefits.

ACKNOWLEDGMENTS

We thank the study participants and the CHI reviewers. This work was partially funded by the Natural Sciences and Engineering Research Council of Canada (NSERC).

REFERENCES

1. Ahtinen A., Mattila, E., Väätänen, A., Hynninen, L., Salminen, J., Koskinen, E., Laine, K. User experiences of mobile wellness applications in health promotion. *Proc. PervasiveHealth'09*.
2. Albaina, I., Visser, T., van der Mast, C., Vastenburg, M. Flowie. A persuasive virtual coach to motivate elderly individuals to walk. *Proc. PervasiveHealth'09*.
3. Barua, D., Kay, J. and Paris, C. Viewing and controlling personal sensor data: what do users want? In PERSUASIVE 2013.
4. Bickmore, T., Caruso, L., Clough-Gorr, K. Acceptance and usability of a relational agent interface by urban older adults. *Ext. Abstracts CHI'05*.
5. Bravata, D., Smith-Spangler, C., Sundaram, V., Gienger, A., Lin, N., Lewis, R., Stave, C., Olkin, I., Sirard, J. Using pedometers to increase physical activity and improve health. *JAMA 2007, 298 (19)*.
6. Consolvo, S., Everitt, K., Smith, I., Landay, J. Design requirements for technologies that encourage physical activity. *Proc. CHI'06.*
7. Consolvo, S., Klasnja, P., McDonald, D., Avrahami, D., Froehlich, J., LeGrand, L., Libby, R., Mosher, K., Landay, J. Flowers or a robot army? Encouraging awareness & activity with personal, mobile displays. *Proc. UbiComp'08.*
8. Consolvo, S., Klasnja, P., McDonald, D., Landay, J. Goal considerations for persuasive technologies that encourage physical activity. *Proc. Persuasive 2009.*
9. Consolvo, S., McDonald, D., Toscos, T., Chen, M., Froehlich, J., Harrison, B., Klasnja, P., LaMarca, A., LeGrand, L., Libby, R., Smith, I., Landay, J. Activity sensing in the wild: a field trial of UbiFit garden. *Proc. CHI'08.*
10. Fogg, B. *Persuasive Technology: Using computers to change what we think and do*. Morgan Kaufmann, 2003
11. Klasnja, P., Consolvo, S., Pratt, W. How to evaluate technologies for health behavior change in HCI research. *Proc. CHI'11*.
12. Li, I. Beyond counting steps: using context to improving monitoring of physical activity. *Proc. UbiComp'09*.
13. Lin, J., Mamykina, L., Lindtner, S., Delajoux, G., Strub, H. B. Fish'n'Steps: encouraging physical activity with an interactive computer game. *Proc. UbiComp'06*.
14. Michie, S., Ashford, S., Sniehotta, F. F., Dombrowski, S. U., Bishop, A., French, D. P. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: The CALO-RE taxonomy. Psychology & Health, 26(11)..
15. Munson, S. Mindfulness, reflection, and persuasion in personal informatics. *Ext. Abstract CHI'12*.
16. Munson, S., Consolvo, S. Exploring goal-setting, rewards, self-monitoring, and sharing to motivate physical activity. *Proc. PervasiveHealth'12*.
17. Newman, M.W., Lauterbach, D., Munson, S., Resnick, P. and Morris, M.E. It's not that I don't have problems, I'm just not putting them on Facebook: Challenges and opportunities in using online social networks for health. CSCW 2011.
18. Peng, W., Hsieh, G. The influence of competition, cooperation, and player relationship in a motor performance centered computer game. In *Computers in Human Behavior 28(12)*.
19. Prasad, A., Sorber, J., Stablein, T., Anthony, D., Kotz, D. Understanding sharing preferences and behavior for mHealth devices. *Proc. WPES'12*.
20. Prochaska, J., Johnson, S., and Lee, P. The transtheoretical model of behavior change. In S.A. Shumaker, E.B. Schron, J.K. Ockene and W.L. McBee, eds. The handbook of health behavior change. 1998.
21. Schwanda, V., Ibara, S., Reynolds, S., Cosley, D. Side Effects and 'Gateway' Tools: Advocating a Broader Look at Evaluating Persuasive Systems. *Proc. CHI'11.*
22. Tollmar, K., Bentley, F., Viedma, C. Mobile health mashups. *Proc. PervasiveHealth'12*.
23. Toscos, T., Faber, A., An, S., Gandhi, M. Chick Clique: persuasive technology to motivate teenage girls to exercise. *Ext. Abstracts CHI'06*.
24. Tudor-Locke, C. Taking steps toward increased physical activity: using pedometers to measure and motivate. *President's Council on Physical Fitness and Sports: Research Digest, Series 3, No. 17, June 2002.*