Flowers or a Robot Army? Encouraging Awareness & Activity with Personal, Mobile Displays

Sunny Consolvo[1, 2], Predrag Klasnja[2], David W. McDonald[2], Daniel Avrahami[1], Jon Froehlich[3], Louis LeGrand[1], Ryan Libby[3], Keith Mosher[3], & James A. Landay[1, 3]

[1]Intel Research Seattle
Seattle, WA 98105, USA
[sunny.consolvo, daniel.avrahami, louis.l.legrand]@intel.com

[2]The Information School
University of Washington
Seattle, WA 98195, USA
[consolvo, klasnja, dwmc]@u.washington.edu

[3]Computer Science & Engineering
University of Washington
Seattle, WA 98195, USA
[jfroehli, libby, kemosher, landay]@cs.washington.edu

ABSTRACT

Personal, mobile displays, such as those on mobile phones, are ubiquitous, yet for the most part, underutilized. We present results from a field experiment that investigated the effectiveness of these displays as a means for improving awareness of daily life (in our case, self-monitoring of physical activity). Twenty-eight participants in three experimental conditions used our UbiFit system for a period of three months in their day-to-day lives over the winter holiday season. Our results show, for example, that participants who had an awareness display were able to maintain their physical activity level (even during the holidays), while the level of physical activity for participants who did not have an awareness display dropped significantly. We discuss our results and their general implications for the use of everyday mobile devices as awareness displays.

Author Keywords

ambient display, calm technology, personal mobile display, persuasive technology, mobile phone, awareness, physical activity, fitness, field experiment, glanceable display

ACM Classification Keywords

H.5.2 User Interfaces, H5.m Miscellaneous.

INTRODUCTION

From the screens on mobile phones that are already prevalent, to the growing area of dynamic displays on jewelry and personal accessories (e.g., [1] and [13]), it is becoming increasingly common for individuals to carry or wear small, personal displays. However, these displays—particularly the background screens and screen savers on mobile phones—are underutilized. Mobile phone displays typically show the date, time, battery status, signal strength, missed calls, voicemail, and text message icons overlaid on a stock image or personal photo; displays of some newer phones also provide quick access to the phone's applications and information sources. While that may seem like a lot of useful information already, there is an opportunity to provide even more.

At the core of the work presented in this paper is the notion that these always-available displays could be used to increase an individual's awareness about various elements of daily life. For example, these personal, mobile displays could provide awareness of one's own behavior (e.g., how physically active one is) or of the activities of friends and family. Awareness about one's own behavior is particularly useful if one is trying to change behaviors or habits (e.g., lead a healthy lifestyle).

Yet presenting awareness information on a personal, mobile display poses important design challenges. The information is often private and may be sensitive (e.g., one's weight or simply that one is self-monitoring behavior). While the displays are personal, they are often not private as it is not unusual for people to see the displays of each other's mobile phones. Furthermore, the design of effective awareness displays must account for the fact that individuals will not attend to the information *all* the time—just because individuals are self-monitoring to change behavior does not mean that awareness of the behavior must be their constant focus.

We previously proposed using stylized representations of behavior on personal, mobile displays as a technique well suited to address these design challenges [5]. When information is represented in a stylized, abstract way, a sense of privacy can be provided while important information is conveyed to the individual in a readily accessible way. Our current investigation examines the long-term effectiveness of using a stylized representation of behavior on a personal, mobile display to encourage regular and varied physical activity. We conducted a three-month field experiment of our system, UbiFit, with 28 participants over the winter holiday season—a time of year known to be particularly challenging for maintaining physical activity [14]. We provide results that describe participants' activities and experiences throughout the study, focusing on the effectiveness of the display on awareness and behavior.

We begin with a discussion of systems that encourage self-monitoring and behavior change through (1) the use of stylized representations of behavior and (2) mobile phones. We follow with a description of UbiFit and the three-month field experiment, then present key results. Finally, we discuss implications of the findings and conclude.

RELATED WORK

In this section, we discuss recent projects that examined the use of (1) stylized representations of behavior in non-mobile settings to encourage individuals to change their behavior, and (2) mobile phone applications that encourage self-monitoring and increase of physical activity.

Stylized displays to encourage behavior change

Two projects that use stylized displays to represent behavior in non-mobile settings are of particular relevance to our work: Breakaway and Fish'n'Steps. *Breakaway* [10] encourages individuals who have sedentary desk jobs to take breaks. A small sculpture on the user's desk slumps over the longer she remains in her chair without getting up. Constant, peripheral awareness is provided to the user when she is at her desk. However, if ignored, Breakaway takes no additional action. A pilot evaluation with a single user who collected one week of baseline data and one week with Breakaway confirmed that the ability to ignore Breakaway during busy times was appreciated.

Fish'n'Steps [11] links an individual's step count to the emotional state, growth, and activity of a virtual fish in a shared virtual fish tank as a means to encourage individuals to take more steps each day. The fish tank is displayed on a kiosk in a common area of an office and on a personal web site. Fish'n'Steps was evaluated over six weeks by 19 participants with sedentary jobs who worked in the same office (then compared to pre- and post-intervention baselines from the same participants using pedometers only over eight additional weeks). A finding of particular relevance to our work is that some participants ignored the display when their fish was not happy—punishment seemed to result in avoiding the system rather than encouraging the desired behavior. However, several participants who wanted to increase their step count did so, and the appealing design garnered interest from other office inhabitants.

Mobile phone applications that encourage self-monitoring & increase of physical activity

Several recent research projects use mobile phones to encourage individuals to self-monitor and increase physical activity (e.g., [4], [7], [12], and [16]). However, two projects are of particular relevance to the work presented in this paper: Houston and the Healthy lifestyle coach. Similar to Fish'n'Steps, *Houston* [4] encourages individuals to take more steps each day. Small groups of friends use their mobile phones to share their step counts and performance toward a daily goal. Prompts on the phone encourage self-monitoring. Simple rewards are provided upon goal attainment (e.g., an '*' next to the step count). A three-week long, in situ pilot study (one week of baseline data collection and two weeks with Houston; n=13) found that participants increased awareness of their activity levels and were motivated by the simple rewards for achieving their goal. However, several participants complained that step count did not always represent their activity levels well, as pedometers do not detect activities such as bicycling and rock climbing. This resulted in some participants not doing those types of healthy behaviors. Participants suggested that *all* relevant physical activities should be included.

The *Healthy lifestyle coach* [7] encourages individuals to be more active and eat more fruits and vegetables. Individuals self-report daily duration of physical activity and fruit/vegetable consumption. Similar to Houston, a simple reward (a "smiley" face) is provided upon goal attainment, and prompts encourage self-monitoring. A four-week long in situ study (n=40) compared a mobile phone version of the system with a functionally equivalent desktop computer-based application (participants were further divided into "individual" and "team" usage modes). Results revealed no significant difference in goal attainment based on the version of the application (mobile vs. desktop), nor between participants acting as individuals versus those in teams. However, the mobile application was used more consistently throughout the day, suggesting that its use was better integrated with everyday life.

UbiFit draws from these projects, for example, by prompting the user to self-monitor, providing positive reinforcement rather than punishment, providing simple rewards for goal attainment and for performing the desired behavior, providing frequent opportunities for self-reflection by using the individual's mobile phone, accounting for a range of physical activities, and using stylized, abstract representations of behavior. UbiFit was also designed for occasional ignorability.

The work we present in this paper is novel as it examines the use of stylized representations of behavior *on* personal, mobile displays with the aim of encouraging self-monitoring and behavior change. Additionally, our three-month field experiment is one of the longer-term deployments where an early-stage, functional prototype was deployed for multiple months in the field with representative users who were not affiliated with the research team or their company/institution.

THE UBIFIT SYSTEM

UbiFit is a mobile, persuasive technology that we have developed to encourage individuals to self-monitor their physical activity and incorporate regular and varied activity

Figure 1. UbiFit Garden's Glanceable Display. a) at the beginning of the week; b) after one cardio workout; c) a full garden with variety; and d) a full garden on the background screen of a mobile phone. Butterflies indicate met goals.

into everyday life. Because the underlying technology has been presented elsewhere [5], we offer a brief overview here.

UbiFit targets individuals who have recognized the need to incorporate regular physical activity into their everyday lives but have not yet done so, at least not consistently[1]. It consists of three components: (1) a glanceable display, (2) an interactive application, and (3) a fitness device. The *glanceable display* uses a stylized, aesthetic representation of physical activities and goal attainment to keep the individual focused on the act of self-monitoring and her commitment to fitness. It resides on the background screen of the individual's mobile phone to provide a subtle reminder whenever and wherever the phone is used. The *interactive application*, which resides on the mobile phone, includes detailed information about the individual's physical activities and a journal in which activities can be added, edited, and deleted. Finally, the *fitness device* (currently a separate device worn on the waistband above the hip like a pedometer), automatically infers and transmits information about several types of activities—walking, running, cycling, use of the elliptical trainer, and use of the stair machine—to the glanceable display and interactive application.

The UbiFit system has been designed using an iterative design process. In addition to drawing from prior work (including our own [4]), the design was informed by a survey we conducted with 75 respondents from 13 states across the

Figure 2. Garden mappings and two sample gardens.

U.S. which examined a range of attitudes and behaviors with mobile devices and physical activity. The survey tested assumptions about the glanceable display and elicited general feedback. We then conducted a three-week field trial of the full system with 12 participants [5]. In the overview that follows, we highlight elements of the system that were redesigned based on results from the three-week field trial.

UbiFit's Glanceable Display

The glanceable display is a stylized, aesthetic image that resides on the background screen, or "wallpaper," of the individual's mobile phone. It presents key information about physical activity behavior and goal-attainment status, as well as a subtle but persistent and easily accessible reminder of commitment to physical activity and self-monitoring.

The display we have implemented for our investigations uses the metaphor of a garden that blooms throughout the week as physical activities are performed, thus we refer to this version of UbiFit as "UbiFit Garden" (Figure 1). Different types of flowers represent the types of activities that the American College of Sports Medicine (ACSM) suggests are important to a well-balanced routine[2]: Cardiovascular training (cardio), Strength training, Flexibility training, and Walking (Figure 2). Based on results from the three-week field trial, a fifth flower was added to represent activities from the "Other" category (e.g., where activities such as *housework* or *chopping wood* could be included). Each flower represents an individual event (*e.g.*, a 40-minute run and a 3-hour bicycle ride are each represented by one pink cardio flower; a 22-minute walk by one sunflower; a yoga class by one white flexibility flower; and a weight lifting session by one blue strength training flower). Walking and cardio activities must be at least 10 minutes in duration to receive a flower; a flower's height has no relation to the activity's duration and

---

[1] UbiFit targets the contemplation, preparation, and action stages of change of the *Transtheoretical Model*, which describes the stages through which individuals progress to intentionally modify problematic behaviors [15].

[2] The ACSM specifies three types of activities: cardio, strength, and flexibility [18]. Based on our prior work [4], UbiFit maintains walking as a separate category to distinguish it from more vigorous cardio activities such as running and cycling.

varies simply for aesthetics. With UbiFit Garden's glanceable display, a healthy garden represents healthy behavior.

Since most agencies that promote physical activity guidelines (e.g., the ACSM and American Heart Association [8][18]) use a weekly time frame, UbiFit's display represents one week's worth of activities, and goals are set and attained on a per-week basis. Upon meeting a weekly goal, a large butterfly appears in the upper right of the garden display. Up to three smaller butterflies represent recent goal attainments, serving to reward and remind the individual of successes over the past month. Yellow butterflies indicate that the *primary weekly goal* was met. Based on results from the three-week field trial, an optional, alternate weekly goal was added to help the individual meet goals during difficult periods, such as a deadline at work or a mild illness. White butterflies indicate that the *alternate weekly goal* was met (Figure 2). At the end of each week, the display resets to an empty garden (Figure 1a).

Drawing on lessons from Fish'n'Steps [11], Houston [4], and the Healthy lifestyle coach [7], UbiFit only uses positive reinforcement—the individual is not punished for inactivity (e.g., the garden has no wilting flowers, weeds, or stormy sky). Instead, inactivity results in a sparse (or empty) garden with a blue sky and green grass (e.g., Figures 1a & 1b). If the individual does not meet a goal, the garden will simply not have a butterfly.

UbiFit's Interactive Application

UbiFit includes an interactive application that runs on the individual's mobile phone. It was built using the MyExperience framework [6], a scripting environment for mobile phones. The interactive application includes details about activities inferred by the fitness device (discussed next) and a journal to add, edit, or delete information about activities, including those *not* inferred by the fitness device. Through the application, the individual can:

- View a daily list of activities performed today and any prior day;
- Add, edit, or delete activities for today and yesterday;
- View progress toward the weekly goal(s); and
- Add a comment to the daily activity list (e.g., "sick," "hiked Gyeryongsan," or "UbiComp paper deadline").

If the individual has not manually journaled for two days, a prompt asks if she has anything to add. In the event of certain technology failures, a troubleshooting survey appears on the phone. For example, if the Bluetooth connection between the phone and fitness device drops, a dialog alerts the individual and attempts to help her fix it. UbiFit avoids disrupting normal phone usage by delaying prompts and other actions when the individual is on a call.

UbiFit's Fitness Device

To automatically infer and transmit information about several types of physical activities to the glanceable display and interactive application, UbiFit employs the *Mobile Sensing*

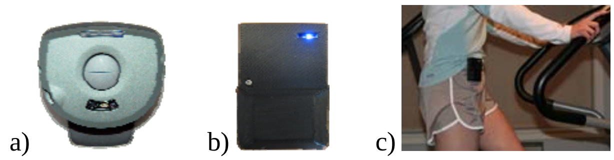

Figure 3. The fitness device—the MSP prototype. a) in the original gray case, b) in the new black case, and c) as worn by a woman during exercise.

Platform (MSP) [3] (Figure 3), a research platform for mobile sensing and inference systems.

The MSP is a pager-sized, battery powered computer with sensors chosen to facilitate a wide range of mobile sensing applications[3]. Though the MSP contains several sensors (3-d accelerometer, barometer, humidity, visible and infrared light, temperature, microphone, and compass), UbiFit only uses two to infer physical activities in real time: the 3-d accelerometer and barometer. The MSP has been trained to infer walking, running, cycling, using an elliptical trainer, and using a stair machine. It transmits a list of activities and their predicted likelihoods to the phone four times per second over Bluetooth. The phone application aggregates and "smoothes" these fine-grain, noisy data resulting in "human scale" activities such as a 22-minute walk or a 35-minute run. Additional details about the MSP can be found in [3][5].

THE THREE-MONTH FIELD EXPERIMENT

To evaluate the effectiveness of mobile, ambient displays on awareness and behavior, we conducted a three-month field experiment that examined participants' activities and experiences throughout the study both quantitatively and qualitatively. The intent of this experiment was to get beyond potential novelty effects that may have been present in our previous three-week field trial [5] and to systematically explore the effectiveness of the glanceable display and fitness device components through the use of experimental conditions. Additionally, whereas our three-week study [5] focused on reactions to activity inference and the overall concept of UbiFit Garden, this field experiment specifically investigated the effectiveness of the glanceable display and fitness device as a means of encouraging awareness and behavior of physical activity.

The field experiment was conducted over the winter holiday season which included Thanksgiving, Christmas, and New Year's—a time that is notorious for physical *in*activity [14]. We note that the goal of this study was to encourage self-monitoring and physical activity, *not* weight loss. However, we provide a brief account of weight change over the course of the study in the next section, as the holidays are known for both physical *in*activity *and* weight gain [19].

[3] One month into our field experiment, we upgraded all fitness devices in the relevant experimental conditions to the latest MSP prototype that included an additional battery—bringing battery life from 11.25 to 16 hours—and a more discrete case (Figure 3).

Study Design

The field experiment included three in-person sessions and 12 weeks of in situ use of three versions of UbiFit—(1) *Full System* (which included the glanceable display, interactive application, and fitness device), (2) *No Fitness Device* (glanceable display and interactive application only), and (3) *No Glanceable Display* (interactive application and fitness device only). The three conditions[4] were used to help assess the impact of the glanceable display and fitness device.

Prior to the first in-person session, participants completed a consent form and a questionnaire about their demographics, familiarity with and use of technology, and physical activity attitudes and routines. They also prepared their phone contacts for transfer to the study phone. During the first session, participants completed questionnaires about their barriers to physical activity [17], stage of change [17], and last seven days of physical activity [9], and they set a weekly physical activity goal of their own choosing that had to include at least one session of cardio, walking, strength training, *or* flexibility training. Each participant had their height and weight measurements taken, received the study equipment, and instructions on how to use the equipment.

In the second and third in-person sessions (conducted approximately 4 and 12 weeks after the first session), participants were interviewed about their experiences in the study (interviews were audio recorded and transcribed), had their weight measurement retaken, and repeated two of the questionnaires. In the second session, participants were also able to revise their weekly goal and add the optional alternate weekly goal. Participants in conditions *with* the fitness device also had the option of upgrading to the fitness device in the new, more discrete case (Figure 3).

Participants were encouraged to carry the phone and wear the fitness device as often as possible. Participants received compensation at the end of the field experiment based on their use of the system; compensation was *not* based on performing activities or meeting weekly goals.

Participants

Twenty-eight individuals who were recruited by a market research agency participated in the three-month field experiment of UbiFit (15 female/13 male, aged 25 to 54)[5]. All were regular mobile phone users who wanted to increase their physical activity. Participants agreed to put their SIM cards in a phone we provided which they used as their personal phone throughout the study.

The participants represented a range of occupations including real estate agent, personal care assistant, public relations specialist, retail manager, psychologist, event laborer, project manager, human resources specialist, teacher, business developer, comedian, and others. Seventeen were employed full-time (one was also a student), eight were employed part-time (one was also self-employed), two were homemakers, and one was a full-time student. The highest level of education completed for four participants was "some college," for 19 was a Bachelor's Degree, one had course work at the Master's level, three had a Master's Degree, and one had a PhD. Twelve were classified as *normal weight*, nine as *overweight*, and seven as *obese* according to Body Mass Index (BMI) calculations performed on their height and weight measurements.

Participants were randomly assigned to one of the three experimental conditions as follows: 10 participants were assigned to the *Full System* condition (5 female/5 male), nine to the *No Fitness Device* condition (6 female/3 male), and nine to the *No Glanceable Display* condition (4 female/5 male). While we could not balance for BMI in assignment to condition as BMI was not known when assignment took place, each condition included participants from all three BMI categories (incidentally, our analysis found no effect of BMI on any of our measures; this is briefly discussed below).

All 19 participants who used the fitness device chose to switch from the old to the new case in the second interview session. Eighteen participants set an alternate weekly goal at the second session (seven from the *Full System* condition, six from the *No Fitness Device* condition, and five from the *No Glanceable Display* condition).

RESULTS

To evaluate the effectiveness of the UbiFit System, and in particular the effect of the glanceable display on participants' ability to maintain activity levels as well as awareness of their behavior over the three months of the field experiment, we performed both quantitative analyses of activity levels over the course of the study and a qualitative analysis of interview data. Qualitative analysis was done using open coding, a standard method for analyzing interview data.

Statistical Analysis Method

Our analysis of activity level data used two primary outcome variables: the total weekly duration of cardio and walking activities (Activity Duration) and the total number of weekly activities, including cardio, walking, flexibility, and strength training (Activity Count). Since the ACSM makes specific duration recommendations only for cardio and walking activities, Activity Duration included only these two activity types. Flexibility and strength activities, representing a major portion of exercise for some participants, were captured in our Activity Count measure.

We used a Mixed Model analysis, which allowed us to account for the fact that each individual contributed

---

[4] A potential fourth condition—interactive application only—has been available commercially on mobile phones for years (e.g., Athletix or MySportTraining on the Palm Treo), and thus we chose to investigate more novel approaches in our three conditions.

[5] Thirty individuals were originally recruited, however, two dropped out of the study.

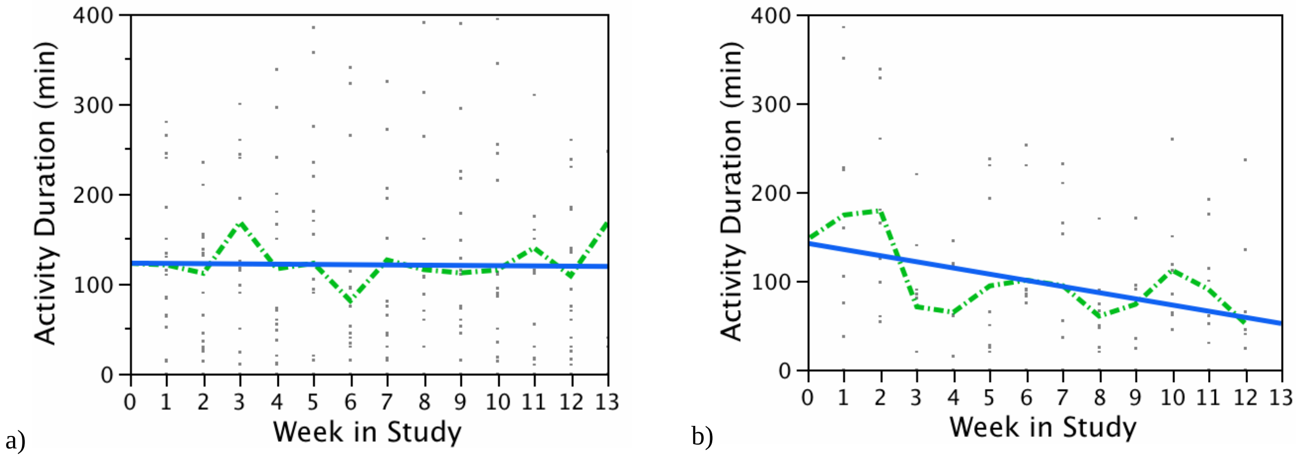

Figure 4. Activity Duration (in minutes per week) over the course of the study for a) participants *with* the Glanceable Display, and b) participants *without* the Glanceable Display. The dashed lines are best fit lines, and the solid lines are linear trend lines.

multiple data points and control for individual differences. The following variables were modeled as fixed effects: the availability of the Glanceable Display (Yes or No), the availability of the Fitness Device (Yes or No), Age, Gender, BMI (Normal, Overweight, or Obese), Week in Study, and Holiday Week (Yes or No) for weeks that included Thanksgiving, Christmas, and New Year's. Two-way interactions between both Glanceable Display and Fitness Device by Age, Gender, Week in Study, and Holiday Week were also included in the model. We excluded from the analysis weeks from the start and end of the study with fewer than 4 days of participation.

Although our study did not include setting a weight loss goal, the winter holiday season often involves weight gain [2][19]. Because our study explicitly targeted the holiday season, we tested participants' weight change over the course of the study to investigate the potential effectiveness of our system to counter holiday weight gain. Here too, we used a Mixed Model analysis, with the ParticipantID modeled as a random effect, and Gender, BMI, Age, Time, and conditions with the Glanceable Display and Fitness Device modeled as fixed effects.

Effectiveness of the glanceable display

We started out by examining the effect of the availability of the Glanceable Display on Activity Duration and Activity Count over the course of the study (our *Week in Study* measure). In accordance with prior literature which shows seasonal differences in individuals' activity levels (i.e., activity goes down in winter months when weather tends to make exercising outdoors less appealing [14]), our analysis showed a decrease over time in both Activity Duration ($F[1, 311]=6.29$, $p=.0127$) and Activity Count ($F[1, 311]=7.0$, $p=.0086$). Our interview data provides support from participants who explicitly noted that they were less inclined to perform outdoor activities in winter, particularly when it was cold, rainy, and/or dark:

> *"In the winter I kind of go into hibernation or something"* <participant F4[6]>

> *"…and if it wasn't so cold. I mean there were times where, you know, it's really cold"* <N5>

> *"I go* [for walks] *at night and so I don't want to go by myself when it's dark out. And if nobody else could go, I'm like, 'oh, I don't want to go.'"* <F3>

Despite the difficulties of being physically active in winter, our analysis revealed a significant effect of the interaction between the availability of the Glanceable Display and Week in Study on Activity Duration ($F[1, 312]=6.51$, $p=.0112$). A closer examination of this interaction shows that it was the average Activity Duration for participants *without* the glanceable display that *decreased* over time (Figure 4b). While participants who *had* the glanceable display *maintained*, on average, their activity duration (Figure 4a). A similar trend was found for the effect of this interaction on Activity Count, although it did not reach statistical significance ($p = 0.095$).

This significant effect of the glanceable display on participants' ability to maintain activity resonates well with comments from participants who stated that the garden helped keep them aware of their activities:

> *"I could see my progress if I was—how much more I needed to do to get to my goal…And I could see the butterfly and think, 'I did it last week, you can do it again this time'."* <F10>

> *"Every time I see my phone, it's a friendly reminder, not a pressing reminder, that I should be doing*

[6] The participant ID reveals the participant's condition. IDs beginning with an 'F' indicate the *Full System* condition, 'N' the *No Glanceable Display* condition, and 'S' the *No Fitness Device* condition.

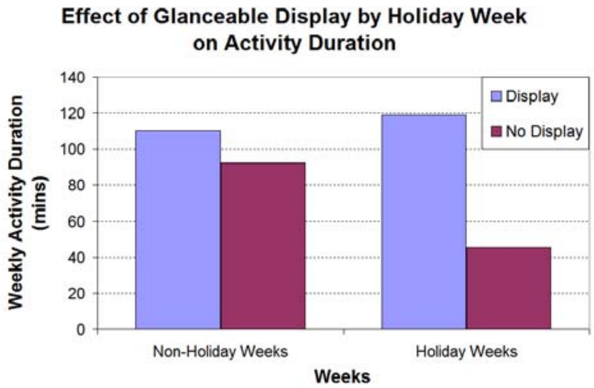

Figure 5. A statistically significant interaction between the availability of the Glanceable Display and Holiday Weeks on Activity Duration.

something. That was great. That was really cool." <S8>

Further, all participants who used the glanceable display thought that it was an essential component of the system, and many who did not use it, wished they had been in one of the other conditions. Toward the end of the exit interview, we described the three versions of the system to each participant, and then asked the participant to speculate on which version would work best for him or her. Of the 28 participants, 25 would want a glanceable display (though not necessarily with a garden metaphor)—including all 19 participants who used the glanceable display and another six who did not. A participant from the No Glanceable Display condition commented:

> *"I love that idea…if it's not flowers and butterflies…I mean it's just something that you don't have to pull up, you know, it's right there. And a lot of times you'd probably just flip on your phone and make a call and not think about it…but it'd be pretty obvious, you know, just having it show up there…That would be a much quicker reminder because you're looking at the phone so many times during the day anyway."* <N10>

Two participants said it would have been a nice but not essential component to have, and one commented that it was *"so cheesy"* <N6>.

The glanceable display vs. turkey, pie, and football

While the above evidence suggests that the glanceable display was effective, further evidence emerged from our analysis of the effects of activity levels over weeks that contained a holiday. Thanksgiving, Christmas, and New Year's occurred during the field experiment, and with them, the visits to/from family and friends, thorough house cleaning, shopping, football game watching, overeating, and other factors that can make the maintenance of physical activity difficult during the holiday season. For example:

> *"Once November comes, I don't know what it is, but there's something inside me—I want to bake. I want to, you know, eat sweets. You anticipate eating. You get a little more lazy. Holidays everybody's around."* <N5>

> *"I need to clean instead of going on my walk or doing something more active."* <F3>

Again, despite the barriers to activity that come with the holidays, our analysis revealed that the interaction between the availability of the Glanceable Display and Holiday Weeks was significant ($F[1, 311]=4.11$, $p<.05$). To understand the meaning of this interaction, we performed a post-hock Tukey HSD pair-wise comparison that showed a significant difference between the weekly Activity Duration over Holiday Weeks for participants with and without the Glanceable Display. While Activity Duration significantly *decreased* over Holiday Weeks for participants who *did not have* the Glanceable Display, Activity Duration over Holiday Weeks *remained as high* as during non-holiday weeks for participants who had the Glanceable Display (Figure 5). This is of particular note as holiday weeks were interspersed throughout the study and the system did not change between weeks with and without holidays (e.g., there was no additional encouragement from the system to keep participants focused on activity during holiday weeks).

Weight change over the winter holiday season

Individuals tend to gain weight over the winter holiday season, and the weight that is gained over that time is thought to be an important contributor to the gradual weight gain that often occurs in adulthood; that is, individuals do not tend to lose all of the weight that they gain over the holidays [19]. Therefore, losing or simply maintaining weight over the holidays would be a positive result. Over the course of the study, participants lost weight on average, although that loss was not statistically significant. This finding is consistent with the health sciences literature which reports that self-monitoring over the holidays helps prevent or at least minimize weight gain [2].

Frequency and variety of activity

UbiFit aims to encourage regular *and* varied physical activity to support the ACSM's notion of a well-balanced routine (i.e., one that includes cardio, walking, strength, *and* flexibility training). Additionally, we had a goal of not repeating one of Houston's results where some participants refrained from performing healthy activities that were not supported by the system [4]. UbiFit supports variety and accounts for a range of activities in two primary ways: (1) different images in the glanceable display represent different types of activities (Figure 2), and (2) in the interactive application, individuals may manually journal any type of physical activity, as well as correct inference mistakes made by the fitness device.

Over the three months of the field experiment, 1792 cardio, walking, strength, and flexibility training activities were recorded. The total number of activities per participant ranged from 12 to 176, with a mean of 61.1 activities (SD=38.29, median=54.5). The average number of activities for each participant per week ranged from 2.83 to 8, with a mean of 5.04 activities (SD=1.2). Not surprisingly, walking was the most frequently recorded activity, followed by cardio, flexibility, and then strength training.

Considering activities that were performed *at least once* during the three-month field experiment, 21 participants (71.43%) performed all four main types of activities. All 28 participants both walked *and* did cardio at least once, 25 (86%) did strength training, and 23 (82%) did flexibility training. However, if we consider activities that were performed *at least once per month*, only five participants (18%) performed all four main types of activities monthly. Fifteen participants (54%) both walked *and* did cardio at least once per month, 13 (46%) did flexibility training, 12 (43%) did strength training, six walked but did not do cardio, and six did cardio but did not walk.

In addition, 17 of the 28 participants chose to journal "other" physical activities (i.e., activities that the participant did not consider to be cardio, walking, strength, or flexibility), for a total of 61 activities that were recorded as "other." Among the journaled "other" activities were housework, climbing the stairs, chopping wood, cleaning the basement, shampooing the rugs, gardening, and crabbing. Housework and climbing the stairs were the two most commonly journaled "other" activities.

The diversity of reported activities suggests that UbiFit does in fact support variety. Participants also confirmed a finding from our prior three-week field trial that emphasized the importance of being able to include in the system the range of activities that were performed [5].

Effect of other contextual variables

We found no significant main effect for any of the other independent variables (BMI, Gender, etc.). Interestingly, the fitness device did not have the same effect as the glanceable display. We found neither a main effect for the fitness device, nor any interaction effects. Similarly, when we repeated the analysis using study condition as an independent variable in place of separate variables for the glanceable display and fitness device, no effect was found. This suggests that it was indeed the glanceable display that most contributed to helping participants maintain activity over holiday weeks and over time in general.

Metaphors for the glanceable display

While we implemented a garden metaphor in the glanceable display for our investigations, we have always believed that such a system should offer the individual a choice of metaphors. As it turned out, so did many of the participants.

Overall, participants appreciated how the garden worked, that is, they could quickly glance at it to determine if they had been active that week, if their routine contained variety, and if they had met their goal this week or any of the past three weeks. If they looked more carefully, they could count the number of activities they had done so far this week and how many of each type. If they looked closer still, they could determine if the activities they had done counted toward their weekly goal or not (e.g., if a flower's stem has leaves, the activity it represents counts toward the weekly goal, if the stem does not have leaves, the flower does not count toward the weekly goal).

However, even those who liked the garden thought that over time, they would want the ability to switch to other metaphors to prevent boredom. Common suggestions included fish- or animal-related themes, sports themes, car themes, celestial themes, and forest-related themes. One of the more unique suggestions was for a robot theme, *"…it would be nice to have a more like dudely kind of motif for it* [the glanceable display]*…Like a robot. If you could amass a robot army…I'd run a lot!"* <F7>. A unique animal-related suggestion came from a participant's friend, "[a friend who does small animal rescue] *was very interested…Instead of flowers she wanted to get rats…so that would be her garden. She'd have a little garden of rats. Little bobo and—"* <S9>.

Not surprising to us, the suggestion of a stripper-related theme was raised. We have encountered this suggestion before (in the survey *and* three-week field trial). The general idea is that the display begins with an attractive, clothed individual, and as activities are performed, items of clothing are removed. We have been concerned that such a metaphor may not be appropriate for the background screen or screen saver of one's mobile phone, given that as we argued earlier, while the screens are personal, they are often not private. When reflecting on this idea (*without* our suggesting that it might not be appropriate), one participant came to this conclusion on his own:

> *"I have* [my mobile phone] *in a business environment. What if somebody saw something like that?...So, though that would probably be entertaining to see that and work toward my goals that way, it's probably—probably wouldn't necessarily work out…Cause I have my phone sitting on my desk and if the—I don't know. If it rings or something like that. What if somebody saw that? That would be horrible."* <F6>

DISCUSSION

We opened this paper by arguing that stylized, abstract representations of information on personal, mobile displays could be used to increase an individual's awareness about various elements of daily life, for example, his or her own behavior. The quantitative and qualitative results presented above indicate that such displays can indeed be effective at raising awareness and potentially influencing behavior. We provided evidence that activity levels were maintained over the course of the study, even during holiday weeks, for participants who *did* have the glanceable display, but that activity significantly declined for participants who *did not* have the glanceable display. Further, the participants commented in interviews on the effectiveness of the display on their awareness, and most participants, including all who used the glanceable display for the three months of the study, thought that a glanceable display would be an essential component to such a system. Participants agreed that the mobile phone was the right place to include both the glanceable display *and* interactive application (several thought that an accompanying web site for viewing longer-term trending would be nice). It seems that the glanceable display on the mobile phone background screen was effective at offsetting well-known negative effects of holiday-related and seasonal changes on physical activity.

Additionally, the glanceable display on the mobile phone was effective at providing awareness in a persistent yet subtle way. Even participants who had the display but were relatively inactive during the study appreciated the persistent awareness that it provided, as it kept them focused on their long term commitment to becoming more active, even if something else (e.g., school deadlines) had temporarily taken priority over fitness, *"It was just, as I say, keeping my eyes on the prize. It's just a reminder that this is something I should be about…it's not annoying. I mean it's annoying in a good way"* <S8>.

Returning to the issue of supporting a *variety* of physical activities, our data suggests that when participants are given the opportunity to track a wider range of activities, they do so. Over the course of the study, participants recorded 60 distinct kinds of activity across the five activity types (cardio, walking, strength, flexibility, and "other")—26 kinds of cardio activities alone were recorded. This suggests that systems that target domains where a variety of activity types is important to the end goal (e.g., being fit and healthy) should not artificially limit recorded activities to those that the system can automatically sense (e.g., as Houston [4] found when some participants refrained from healthy activities that the pedometer did not detect). The lack of support for the full range of relevant activities can be a major source of frustration for users and, consequently, a potential cause of system abandonment. No participant mentioned refraining from an activity because of UbiFit.

It is important to note that while the fitness device did not show the same type of significant effects on activity levels as the glanceable display, participants were generally positive about the idea of the fitness device. The problems that participants noted—its size, discomfort, bright LED, relatively short battery life (even at 16 hours), need to be in constant communication with the phone, and occasional technical difficulties—had to do with the particular research prototype and not with the idea of activity inference or a fitness device. Continuing to improve activity inference, as well as the form factor of inference devices, thus continues to be a valuable research aim. When speculating on the future form factor of such devices—for example, should the sensors be integrated with the phone itself, or should there be a much smaller fitness device that could operate without a constant connection to the phone—participants were undecided. While the phone-only solution was initially appealing, it became much less so if it would require wearing the phone clipped to one's waistband or in a front pant pocket. Both trajectories should be explored further.

One surprising issue related to awareness came up during the exit interviews. While we expected that awareness would be raised somewhat simply by participating in a study about physical activity, and that the glanceable display and fitness device would further raise activity awareness, the phones used in our study produced an unanticipated confound. As mentioned above, participants put their SIM cards into a study phone that we provided. We made this choice to maximize the naturalness of phone use (e.g., having participants carry a second phone that ran UbiFit, in addition to their personal phones, would not have been as natural, and the alternative of only recruiting people who already had the model of phone that ran UbiFit would have been too restrictive). However, for some participants in the *No Glanceable Display* condition, the study phone itself became a trigger for awareness about activity as it was associated specifically with the fitness study. In retrospect, this makes sense; after all, the phone was not participants' normal phone, and they were returning it at the end of the study. We should note, however, that this unexpected confound itself does not change the results we presented; in fact, one should expect the effect of the glanceable display to be even stronger without this confound.

A related issue has to do with the inevitable question of a novelty effect. Although our study is reasonably long when compared to much of the related work, it is not clear whether we succeeded in overcoming the novelty effect, even after three full months of system use. This caused us to wonder, how do we, as ubicomp researchers, confidently know when and if we have overcome novelty in our studies of early-stage, novel prototypes? Is this even possible with an often limited number of prototypes that are usually not ready for truly long term deployments of natural use? If not, what does that mean for the way we design our studies and how do we, as a community, evaluate work that tries to demonstrate effectiveness of early prototype systems? These issues, we believe, are worth further consideration.

CONCLUSION

In this paper, we described a system that we built, UbiFit, which uses a stylized, abstract representation of physical activity behavior and goal attainment on a personal, mobile display to encourage individuals to self-monitor their physical activity and incorporate regular and varied activity into everyday life. We reported results of a field experiment in which 28 participants recruited from the general population used one of three versions of the UbiFit system for three months over the winter holiday season. UbiFit, the glanceable display in particular, was well-received. We presented results that showed that participants who had the glanceable display were able to maintain their physical activity level over time and on holiday weeks, while the level of physical activity for participants in the condition who did not have the glanceable display dropped significantly. We provided qualitative results that confirmed that participants appreciated the glanceable display and thought it was an essential component of the system.

The work we presented is novel as it examined the use of a stylized, aesthetic representation of behavior *on* a personal, mobile display with the aim of encouraging self-monitoring and behavior change. Additionally, the three-month field experiment is one of the longer-term deployments in the ubiquitous computing community where an early-stage, functional prototype was deployed in the field with representative users who were not affiliated with the research team or their company/institution.

ACKNOWLEDGMENTS

Thanks to Gaetano Borriello, Mike Chen, Tanzeem Choudhury, Cherie Collins, Kieran Del Pasqua, Beverly Harrison, Anthony LaMarca, Jonathan Lester, Jean Moran, Wanda Pratt, Adam Rea, Christine Riley, Ian Smith, Tammy Toscos, David Wetherall, Alex Wilkie, and many other family, friends, colleagues, study participants, and paper reviewers who have contributed to this work.

REFERENCES

1. Ana-Digi watch by PHOSPHOR. http://www.phosphorwatches.com/ {verified 1 Apr 08}
2. Boutelle, K.N., Kirschenbaum, D.S., Baker, R.C., & Mitchell, M.E., "How Can Obese Weight Controllers Minimize Weight Gain During the High Risk Holiday Season? By Self-Monitoring Very Consistently," *Health Psychology 18(4)*, (Jul 1999), pp.364-8.
3. Choudhury, T., Borriello, G., Consolvo, S., Haehnel, D., Harrison, B., Hemingway, B., Hightower, J., Klasnja, P., Koscher, K., LaMarca, A., Lester, J., Landay, J.A., LeGrand, L., Rahimi, A., Rea, A., & Wyatt, D., "The Mobile Sensing Platform: An Embedded System for Capturing and Recognizing Human Activities," *IEEE Pervasive Computing, 7(2)*, (Apr-Jun 2008).
4. Consolvo, S., Everitt, K., Smith, I., & Landay, J.A., "Design Requirements for Technologies that Encourage Physical Activity," *Proceedings of CHI '06*, Montreal, Canada, (Apr 2006), pp.457-66.
5. Consolvo, S., McDonald, D.W., Toscos, T., Chen, M.Y., Froehlich, J., Harrison, B., Klasnja, P., LaMarca, A., LeGrand, L., Libby, R., Smith, I., & Landay, J.A. "Activity Sensing in the Wild: A Field Trial of UbiFit Garden," *Proc of CHI 2008*, (Apr 2008), pp.1797-806.
6. Froehlich, J., Chen, M.Y., Consolvo, S., Harrison, B., & Landay, J.A., "MyExperience: A System for *In Situ* Tracing and Capturing of User Feedback on Mobile Phones," *Proceedings of Mobisys '07*, San Juan, Puerto Rico, USA, (June 2007), pp. 57-70.
7. Gasser, R., Brodbeck, D., Degen, M., Luthiger, J., Wyss, R., & Reichlin, S., "Persuasiveness of a Mobile Lifestyle Coaching Application Using Social Facilitation," *Proceedings of Persuasive 2006*, Eindhoven, The Netherlands, (May 2006), pp.27-38.
8. Haskell, W.L., Lee, I-M, Pate, R.R., Powell, K.E., Blair, S.N., Franklin, B.A., Macera, C.A., Heath, G.W., Thompson, P.D., & Bauman, A., "Physical Activity and Public Health: Updated Recommendations for Adults from the ACSM and the AHA," *Circulation*, (Aug 2007), *116*, pp.1081-93.
9. *International Physical Activity Questionnaire: Long Last 7 Days Self-Administered Format*, (Oct 2002). http://www.ipaq.ki.se/ipaq.htm. {verified 1 Apr 08}
10. Jafarinaimi, N. Forlizzi, J., Hurst, A., & Zimmerman, J., "Breakaway: an ambient display designed to change human behavior," In *CHI '05 Extended Abstracts*, Portland, OR, (Apr 2005), pp.1945-8.
11. Lin, J.J., Mamykina, L., Lindtner, S., Delajoux, G., & Strub, H.B., "Fish'n'Steps: Encouraging Physical Activity with an Interactive Computing Game," *Proceedings of UbiComp '06*, (Sep 2006), pp.261-78.
12. Maitland, J., Sherwood, S., Barkhuus, L., Anderson, I., Hall, M., Brown, B., Chalmers, M., & Muller, H., "Increasing the Awareness of Daily Activity Levels with Pervasive Computing," *Proceedings of Pervasive Healthcare '06*, Innsbruck, Austria, (Nov/Dec 2006).
13. Orth, M. & Berzowska, J.'s Animated Fashion Module.
14. Pivarnik, J.M., Reeves, M.H., & Rafferty, A.P., "Seasonal Variation in Adult Leisure-Time Physical Activity," *Med & Sci in Sp & Exer*, (2003), pp.1004-8.
15. Prochaska, J.O., DiClemente, C.C., & Norcross, J.C., "In search of how people change: Applications to addictive behaviors," *American Psychology, 47(9)*, (1992), pp.1102-14.
16. Toscos, T., Faber, A., Connelly, K., Upoma, A.M., "Encouraging Physical Activity in Teens: Can technology help reduce barriers to physical activity in adolescent girls," In *Proceedings of Pervasive Healthcare '08*, Tampere, Finland, (Jan/Feb 2008).
17. USDHHS, PHS, CDC, NCCDPHP, & DNPA, *Promoting Physical Activity: A Guide for Community Action*, Champaign, IL: Human Kinetics (1999).
18. Whaley, M.H., Brubaker, P.H., & Otto, R.M. (Eds). "General Principles of Exercise Prescription," *ACSM's Guidelines for Exercise Testing and Prescription*, 7th Ed, Balti, MD: Lippincott Williams, & Wilkins, (2006).
19. Yanovski, J.A., Yanovski, S.Z., Sovik, K.N., Nguyen, T.T., O'Neil, P.M., & Sebring, N.G., "A Prospective Study of Holiday Weight Gain," *The New England Journal of Medicine, 342(12)*, (Mar 2000), pp.861-7.