

One Step Forward, Two Steps Back: The Key to Wearables in the Field is the App

Bliss Altenhoff, Haley Vaigneur and Kelly Caine

Department of Psychology, Department of Industrial Engineering, School of Computing
Clemson University
Clemson, United States
{blissw, hvagine, caine} @clemson.edu

Abstract—Wearable trackers and sensors are becoming an increasingly popular option for people to manage their health and fitness and for physicians to monitor patients with chronic illnesses. Consumer wearables such as the Jawbone UP and Fitbit Flex empower people to change habits that may help prevent health problems. The success of wearables depends on their perceived ease of use and how successfully users are able to incorporate them into their lives over the long term. We conducted a usability test and field study with fourteen users comparing two consumer wearable devices. While participants rated the aesthetic design of the Jawbone UP24 higher, they rated app characteristics such as usefulness and trust in data generated of the Fitbit Flex app higher. These findings suggest that while wearable technologies are advancing in capabilities, their acceptance and satisfaction may depend on the quality of the app, in addition to the qualities of the wearable device itself.

Keywords—wearable computing; quantified self; preventive healthcare; mobile app design

I. INTRODUCTION

As health care initiatives and societal trends have begun to encourage healthier lifestyles, technology has met the demand in the form of “wearables”. Wearables aim to seamlessly incorporate technology into everyday life in an easy to wear format. Consumers can use the devices to look at trends to identify what makes them feel good, and to identify when they are deviating from specific fitness goals. These devices have become popular, in part, because they offer solutions for those wishing to prevent, diagnose, and manage diseases. Although the possible applications for wearable tracking devices are immense, examples include occasional use by wellness or fitness tracking users to continuous use by chronically ill inpatients or outpatients and their physicians. In the healthcare setting, wearables can be used to continuously sense a patient’s physiological data, such as heart rate, blood pressure, etc., to quickly alert a physician when necessary. For example, diabetic patients can wear a blood glucose sensor that communicates readings directly to the patient or physician [4, 12, 13].

With a recent increased focus on improvement of electronic health records (EHRs), efficient data collection is necessary for the success of meeting government-set initiatives and improving patient safety. Data from wearables

could be used to populate an individual’s EHR [14], supply rich information from a large sample to decision support systems to guide physicians, and predict or alert physicians or users when aid is needed. Rather than monitoring all patients, wearable sensors could allow health care providers to offer more effective and efficient care, providing real-time feedback, and reduce costs of services [3, 17].

More recently, wearables have evolved to attract a healthier population. Factors such as social networking, personal fitness awareness, smartphones, and gamification, have created a market for wearables as activity trackers [10]. It has been estimated that 80 million fitness monitoring devices will be sold by 2016 [18]. As observed at the 2014 Consumer Electronics Show, wearables are quickly growing in popularity [1], with one in ten Americans now owning an activity tracker of their own [9].

Wearables also have the potential to significantly reduce medical costs and improve healthy habits of its users as a tool for preventative care. In 2008, the estimated annual medical costs for people who are obese was \$147 billion, \$1,429 higher per person than those of normal weight. With one-third of the population qualifying as obese, and about two-thirds as overweight [2], consumer wearables such as the Jawbone UP, Fitbit, and Nike FuelBand are helping individuals track and quantify their physical activity to reach healthier lifestyle goals. By measuring diet, sleep, and fitness performance, users can more easily identify and quantify areas for improvement, using the wearable as a teaching aid [6].

A. Purpose and Goals

The need for healthier lifestyles combined with the quantified-self movement, which uses technologies to track and measure personal health metrics (sleep, diet, exercise, stress, etc.) [15], has motivated many companies to develop wearable trackers. However, while research on development and design of wearable devices has been increasing in recent years, the emphasis has been on functionalities [8], design of the form factor, [16] and increasing accessibility [11] rather than better understanding consumers’ abilities to correctly and effortlessly use such meaningful technologies.

Many activity tracking wearables available to consumers require the use of a mobile app for viewing data, additional tracking, and personalization. Perhaps because of the recent emphasis on development of wearable devices, resources may have not been adequately allocated to the research and design of the accompanying mobile app. In the current study we examine the usability of the Fitbit Flex and Jawbone UP24 wearable bands and accompanying mobile apps, which were the two of the most popular and similar devices on the market at the time of this study, with the most units sold in 2013 [5]. By selecting two devices quite homogeneous in price, device design, and functionalities, any differences in the usability or user experience could then likely be attributed to app design differences rather than functionality. With one third of wearable consumers forgoing the device within six months of purchase [9], we sought to investigate how the app may influence initial impressions, which may possibly influence future behavior.

B. Technologies

The Jawbone UP24 is a wristband activity tracker that uses accelerometers to automatically track steps walked during the day and amount and quality of sleep at night. Using the Jawbone UP smartphone app, users can view their tracked data, which is synced wirelessly via Bluetooth, and record additional information about diet, mood, workouts, and trends. They can also use the device to provide an idle alert if they want to be reminded to remain physically active throughout the day, wake them up in the morning based on the optimal time in their sleep cycle, or even set a power nap alarm. The only display on the band is the sleep/awake status lights. As seen in Figure 1, a moon icon will light up to indicate sleep mode, and a sun icon will light up to indicate day mode.



Fig. 1. Jawbone UP24 band and app home screen.

The Fitbit Flex is an activity and sleep tracking wristband. Data collected can be synced wirelessly to an iOS or Android smartphone or a computer (via a dongle) with Bluetooth 4.0. It tracks steps taken, distance traveled, calories burned, active minutes, hours slept, and quality of sleep. The Flex shows daily progress on the wristband with LED lights and vibration, and shows statistics through charts and tables once the data is synced. Additional features include silent alarms, waterproof design, and sharing with other health and social apps. A picture of the device and the app software is shown in Figure 2.

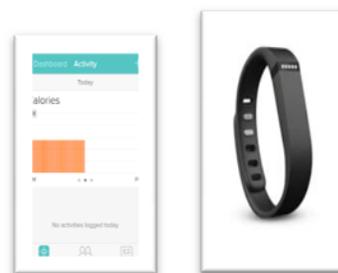


Fig. 2. Fitbit Flex band and app

II. METHODS

A mixed-methods study was conducted to evaluate wearable trackers and their apps. We engaged users in both a think-aloud usability test and a three day field study.

A. Participants

Participants were eight women and six men ($N = 14$), between the ages of 22 and 54 years old, and were randomly assigned to either the Fitbit Flex group (mean age 29.8 years, 4 females, 3 males) or the Jawbone UP group (mean age 24.2, 4 females, 3 males). All participants were English speaking and iPhone, Android, or Windows smartphone users. Participants had no previous experience with wearables.

B. Procedure

Participants were first introduced to the band and app through a think aloud usability test, interacting with main features such as setting up the app for the first time, setting the alarm, and logging diet and workout. Participants then completed a field study where they wore the Fitbit or Jawbone for three days, incorporating it into their everyday life as if they bought it for themselves. They were asked to behave as they wished, using the device as much or as little as they preferred. Finally, after wearing the device and using the mobile app for three days, participants completed a post-test survey that queried their overall experience and an unstructured interview where they discussed their experiences while using the technology in the field.

III. RESULTS

Results are organized by usability study and field study.

A. Usability Evaluation

During the think-aloud usability test time-on-task and errors were recorded (see Table 1), as well as participant comments. Findings are organized by task.

1) *Time on Task*: It took participants more time overall to complete the tasks while using the Jawbone UP ($M = 3.67$ $SD = 3.15$) than the Fitbit Flex ($M = 2.13$; $SD = 2.25$), $t(36) = 2.36$, $p = .022$.

TABLE I. TIME ON TASK, NUMBER OF STEPS, AND SECONDS TAKEN PER STEP FOR EACH TASK IN THE MOBILE APP USABILITY TEST

Metric	Device	Task			
		1. Setup	2. Diet	3. Alarm	4. Workout
Average Errors	Fitbit Flex	0.3	0.4	0	0.1
	Jawbone Up	1	1.1	3	2.3
Average Time (min)	Fitbit Flex	6.1	1.5	0.6	1.2
	Jawbone Up	9.2	3.7	2.5	2
Seconds/Step	Fitbit Flex	52	23	13	18
	Jawbone Up	26	37	25	20

2) *Number of Steps*: It took participants the same number of steps overall to complete the tasks while using the Jawbone UP (M =9.2; SD = 6.8) and the Fitbit Flex (M = 4.2; SD = 1.6), $t(36) = 1.59$, $p = .187$.

3) *Errors*: Participants experienced more errors with the Jawbone tasks (M=1.5, SD=1.6) than with the Fitbit Flex (M = 0.4, SD = 0.7), $t(36) = 3.42$, $p = 0.002$.

B. Observations and Participant Comments

1) *Setup*: This task had the highest number of steps and time on task for both the Jawbone Up and the Fitbit Flex. While the longer time can be attributed to taking more steps, this was also the participants’ first interaction with the device, and the participants worked faster as they became more familiar with their respective devices. Participants also had trouble physically activating the devices. A Jawbone participant stated “I’m not sure what button to push to wake up the band”. Fitbit users found that the instructions did not accurately depict where or how hard to tap the band and multiple participants expressed “I don’t think this is the right spot, and I don’t want to push too hard and break it.”

2) *Diet*: Jawbone UP participants struggled more than Fitbit Flex participants to log diet. Nine out of fourteen of the participants had experience with some kind of health logging app, possibly making this task seem more intuitive. However, a Fitbit participant realized that there was not an option to input decimal amounts (e.g. 10.5 ounces water), only whole numbers. One of the Jawbone errors was due to the participant losing their log: “I accidentally closed the food menu completely when trying to reverse a single food.”

3) *Alarm*: Several of the Jawbone Up participants had trouble locating and understanding the alarm. Although accessible through the band icon in the top right corner, two wanted to go to sleep information, and one expressed they thought it would be under settings. Another participant was also confused about the “smart sleep” setting: “So, it says a smart alarm wakes me up when I’m in light sleep within 30 minutes, but is that before or after?” While the Fitbit Flex alarm function was also under the accounts tab, which could

have presented confusion like the goals, participants did not have trouble with this task because they saw the function when setting the goals, and knew exactly where to go.

4) *Workout*: This task was more straightforward for those with the Fitbit. A Jawbone participant stated “I would go to the right menu because it looks like where you input data”. The Fitbit participants easily went to the “active minutes” section and searched for their exercise and selected the amount.

C. Field Study

After using the wearable device in the field for three days, participants were asked to rate their satisfaction with the band and app. For example, “How useful did you perceive the Jawbone UP/Fitbit Flex to be on a scale of 1 to 5 with 1 being not at all useful and 5 being extremely useful?”

Overall, participants assigned to the Jawbone Up rated it higher than those assigned to the Fitbit in terms of aesthetics $t(10) = 2.34$, $p = 0.04$ and lower in terms of usefulness, burden and trust (see Figure 3), though these tests were non-significant (p 's >.05).

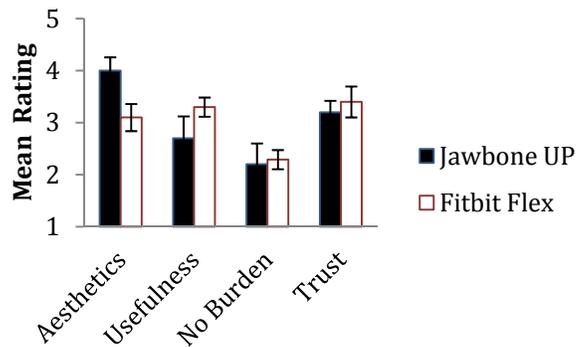


Fig. 3. Mean ratings after using device for three days

IV. DISCUSSION

Overall, participants using the Fitbit were able to complete the usability test faster and with fewer errors and had more positive statements about their experience at the end of the three days relative to those using the Jawbone, despite its lower aesthetic ratings. Issues encountered during the usability

test as well as participant statements tended to focus on interactions with the app rather than the band. It is possible that if users do not believe the wearable device and app to be useful enough, trust the data presented, and find it to be little burden, they may decrease usage over time [9]. Interviews at the end of the three days indicate that participant trust in the data provided by the wearable was influenced by the first experiences with the app, and those beliefs rarely seemed to change across the three days. For example, one Jawbone participant reported not trusting sleep data after the first night when it did not accurately report the time he fell asleep. Another participant was surprised when upon first syncing the band and app, the app displayed about 80 steps before she had taken any actual steps, commenting on day 3, “I feel like I would just use it when working out to figure out what I’d actually done and for sleep but not walking because it’s not accurate.” Another Jawbone participant mentioned that “Certain aspects were a burden, so if they were then I just wouldn’t do them – like the food.” Although the current study only spanned three days, some Jawbone participants reported decreased usage over the short period of time, suggesting that initial impressions of the app may have lasting effects on user engagement.

V. RECOMMENDATIONS AND CONCLUSION

As mobile phones continue to be a popular and promising platform for users to access health tracking data because of their ability to support Internet access, automated sensing, text messaging, and journaling [7], wearables’ accompanying mobile apps should be used to maximize benefits of the wearable. Though the Fitbit Flex and Jawbone UP offer the same features through a wearable wristband and phone app, the physical comfort and superior usability of the Fitbit Flex app appear to contribute to higher subjective ratings of usefulness and trust in data generated, despite the higher rated aesthetic design of the Jawbone UP. This finding suggests that the app design can strongly influence the overall satisfaction and acceptance of the wearable, as a whole. Although one in ten Americans owns a wearable device, half no longer use it and one-third stop using the device within the first 6 months [9]. It is possible that these user-interface issues may contribute to the declining user interest over time, if they feel that tracking their data is burdensome or inaccurate. Especially if first impressions of the app influence overall opinions of the wearable, future designs should ensure they support the user, from the first interaction with a focus on effort of use, and perceived accuracy of data and how they contribute to this lack of long-term engagement. These findings are important because they demonstrate the importance of not only creating easy to use wearables, but also the importance of supporting technologies such as apps. Future research should further investigate change in satisfaction and engagement over a longer period of time to fully understand adoption.

ACKNOWLEDGMENT

This material is based upon work supported by the National Science Foundation under Grant No. 1314342. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

- [1] T. Bradshaw, “CES 2014: Wearables emerge as top trend,” *Financial Times*, <http://www.ft.com/cms/s/0/50264574-7485-11e3-9125-00144feabdc0.html#axzz2wKofTlB>
- [2] Center for Disease Control and Prevention. Adult obesity facts, 2012. <http://www.cdc.gov/obesity/data/adult.html>
- [3] Center for Disease Control and Prevention. “CDC’s vision for public health surveillance in the 21st century,” *Morbidity and Mortality Weekly Report*, vol. 61, 2012.
- [4] T. Fritz, E. M. Huang, G. C. Murphy, and G. C. Zimmermann, “Persuasive technology in the real world: A study of long-term use of activity sensing devices for fitness,” *Proc CHI’14*, pp. 487-496, 2014.
- [5] A. Hsiao. “Terapeak trends: Fitbit, Nike top the market for activity trackers,” 2013. <http://www.terapeak.com/blog/2013/07/31/fitbit-nike-top-the-market-for-activity-trackers/#.VKsTJUvUpqa>
- [6] E. B. Hekler, P. Klasnja, V. Traver, and M. Hendriks, “Realizing effective behavioral management of health,” *IEEE Pulse*, vol. 4, pp. 29-45, 2013.
- [7] J. Knight, H. Bristow, S. Anastopoulou, C. Baber, A. Schwirtz, and T. Arvanitis, “Uses of accelerometer data collected from a wearable system,” *Personal and Ubiquitous Computing*, vol. 11, 2007.
- [8] O. D. Lara, and M. A. Labrador, “A survey on human activity recognition using wearable sensors,” *IEEE Comm Surveys & Tutorials*, vol. 15, pp. 1192-1209, 2013.
- [9] D. Ledger, and D. McCaffrey, “Inside wearables: How the science of human behavior change offers the secret to long-term engagement,” *Endeavour Partners*, 2014.
- [10] M. J. McGrath, and C. N. Scanalli, “Wellness, fitness, and lifestyle sensing applications,” *Sensor Technologies*, pp. 217-248, 2013.
- [11] G. Ngai, S. C. F. Chan, V. T. Y. Ng, J. C. Y. Cheung, S. S. S. Choy, W. W. Y. Lau, and J. T. P. Tse, “i•CATch: a scalable plug-n-play wearable computing framework for novices and children,” *Pro CHI* pp. 443-452, 2010.
- [12] M. L. Resnick, “Ubiquitous computing: UX when there is no UI,” *Proc HFES*, pp.1007-1011, 2013.
- [13] J. Rooksby, M. Rost, A. Morrison, and M. Chalmers, “Personal tracking as lived informatics,” *Proc. CHI ’14*, ACM, pp. 1163-1172, 2014.
- [14] J. Sorber, M. Shin, R. Peterson, and D. Kotz, “Plug-n-Trust: Practical Trusted Sensing for mHealth,” *Proceedings MobiSys*, pp. 309-322, 2012.
- [15] M. Swan, “Emerging patient-driven health care models: An examination of health social networks, consumer personalized medicine and quantified self-tracking,” *International Journal of Environmental Research and Public health*, vol. 6, pp. 492-525, 2009.
- [16] A. Teller, “A platform for wearable physiological computing,” *Interacting with Computers*, vol. 16, pp. 917-937, 2004.
- [17] A. Pantelopoulou, and N. G. A. Bourbakis, “Survey on wearable sensor-based systems for health monitoring prognosis,” *IEEE Trans on Systems, Man, and Cybernetics – Part C: Applications and Reviews*, vol. 40, pp. 1-12, 2010.
- [18] J. Zheng, Y. Shen, Z. Zhang, T. Wu, G. Zhang, and H. Lu, “Emerging wearable medical devices towards personalized healthcare,” *Proc Body Area Networks*, pp. 427-431, 2013.