An Investigation of the Usability and Desirability of Health and Fitness-Tracking Devices

Ashton Pfannenstiel and Barbara S. Chaparro^(⊠)

Wichita State University, Wichita, KS, USA ashtonpfannenstiel@gmail.com, barbara.chaparro@wichita.edu

Abstract. This study investigated the usability and functionality of 6 different fitness tracking wristbands that have been suggested to improve and encourage healthy behaviors. While many previous studies assess the accuracy and behavioral effects of fitness tracking devices, limited research has been done to analyze the usability and desirability of these products. Participants were asked to rate their impressions of six fitness tracking devices - the Garmin Vivofit, Jawbone Up24, Fitbit One, Basis B1 Band, Misfit Shine, and the Tom Tom Multisport – before and after usage. Participants were also asked to describe the main factors contributing to their overall preference and likelihood to purchase and/or use each device. Results indicate that participants are initially more likely to favor, small, lightweight devices that have a display. After wearing the devices, the most valued features were attractiveness, long battery life, waterproof, and a heart rate monitor. The study suggest that a "one size fits all approach" to the design of fitness tracking devices may not be the most effective method to promote the actual use of the technology.

Keywords: Wearables \cdot Beauty technology \cdot Usability methods and tools Mobile products \cdot ISO \cdot Usability

1 Introduction

According to Comstock [1], 19 million wearable fitness tracking devices are owned this year, a figure that is expected to triple in the next 3 years. Despite this success, it is suggested that the early appeal and fascination with the devices does not last. This leaves companies eager to discover what factors appeal the most to individuals in the market for fitness tracking devices. However, Quinlan [3] suggests that this "one-size-fits all approach" is what limits the desirability of these products. In order to appeal to a greater number of potential users, Quinlan [3] states that fitness devices need to be versatile and considerate of people with different, or limited, abilities.

Versatility is very important in terms of fitness devices as one third of the population is considered obese. As a result, flexible material that can be size adjusted is crucial if the market is to expand farther than people who are already considered fit.

Kelly [2] reports 6 trends that are becoming more popular and may even be necessary for a device's popularity. The first trend is that "devices are getting smarter," meaning that fitness trackers do not just report activity, but they now instruct users on how active they should be. Second, smartphone extension is an important function in order for users to have quicker access to their data and additional features. Users also expect beauty from their fitness devices as many companies aim to make their products look more like jewelry, or offer products in a variety of colors. The integration of social networks increases the desirability of products as well, as individuals hope for products that will fit into other components of their lives. In addition, innovation, or new and unique features, encourage consumers to purchase or upgrade fitness tracking devices. The sixth and final trend, heart rate monitoring, is now an expected functionality. However, unlike previous devices, which required an extra chest strap to monitor heart rate, consumers now desire devices that utilize wrist sensors [2].

The current studies will explore available fitness tracking devices and their appeal towards consumers. The goal of the study is to discover how to make wearable fitness tracking technology more desirable to a greater number of individuals outside of the more athletic, fitness-focused community.

2 Method

2.1 Participants

Nineteen participants (10 male, 9 female) were recruited from a Midwestern university between June and July 2014. Participants ranged in age from 18 to 45 (m = 27.63). 17 of the 19 participants were unfamiliar with any fitness tracking device and two had experimented with a step counting device not used in this study.

2.2 Materials

Six fitness tracking devices, and their packaging were evaluated. The devices and packaging were labeled A through F. Participants rated their first impressions and final impressions using a paper 50 point scale. Other materials used during the study included: a tube sock and quarter, a Samsung Chromebook, and a "Feature Ratings" survey. Prior to the study, each participant was asked to complete a background questionnaire.

2.3 Procedure

Packaging. Before seeing any of the devices, each participant rated the packaging of the devices using the 50 point scale. For the first 2 questions, the participant was not allowed to touch the device. Instead, he/she was asked to instruct the researcher where to place each package on the scale. The participant was then allowed to touch and read the packaging. They were asked to explain what influenced their overall opinion of each package.

First Impressions. The packaging was removed and replaced with the 6 devices. Each participant was asked to rate the devices based on his/her first impressions. Again, He/she was not actually allowed to touch the devices and were asked to instruct the researcher where to place each device on the 50 point scale. They were asked to explain what factors influenced their overall first impressions of the devices.

Tasks. Each participant was asked to put on each device in a randomized order and complete the following representative tasks with each one: 1. Put on the device, 2. Walk around/Get a feel for the device, 3. Reach into a sock to pull out a coin (to simulate pocket), 4. Type on a laptop keyboard, and 5. Remove the device.

Final Impressions. After completing the tasks with all 6 devices, participants were asked to rate their final impressions of the devices using the 50 point scale. After rating each attribute, participants were asked to explain their ratings for the devices, specifically the highest and lowest rated device. To conclude the study, the participant was given a "Feature Ratings" survey. The survey consisted of a list of features that can be found on different fitness tracking devices. The participant indicated how important he/she think each feature is. The study took approximately 60–75 min to complete.

3 Results

Packaging. A one-way repeated measures ANOVA indicated significant differences for expense, F(5, 85) = 6.82, p < .01, eta2 = 0.28, quality, F(3.38, 57.54) = 6.30, p < .05, eta2 = 0.27, amount of features, F(5, 85) = 16.42, p < .01, eta2 = 0.49, and the likelihood to buy each device, F(5, 85) = 14.74, p < .01, eta2 = 0.46, when participants were asked their opinions about device packaging. Based on packaging, participants perceived Jawbone as being significantly cheaper than the Garmin, Fitbit, and Tom Tom. Overall, the Jawbone was believed to be lower quality than the other 5 devices. The Jawbone was also thought to have less features. The Jawbone was rated significantly lower than the Garmin, Fitbit, Basis, and Misfit, when asked which device the participant would buy.

First Impressions. A one-way repeated measures ANOVA indicated significant differences for weight, F(5, 85) = 58.91, p < .01, $eta^2 = 0.77$, and comfort, F(2.82, 48.82) = 7.91, p < .01, $eta^2 = 0.31$, when participants were asked about their first impressions of the devices. Overall the Basis and Tom Tom were perceived to be significantly heavier than all the other devices. The Basis was perceived as less comfortable than the Garmin and Misfit, while the Tom Tom was believed to be less comfortable than Garmin, Fitbit, and Misfit.

Final Impressions. A one-way repeated measures ANOVA indicated significant differences for comfort, F(5, 85) = 11.92, p < .01, eta2 = 0.41, masculinity/femininity of the device, F(2.55, 43.36) = 18.17, p < .01, eta2 = 0.51, and the likelihood to buy each device, F(5, 85) = 3.32, p < .05, eta2 = 0.16, when the participants were asked about their final impressions. After trying on and performing tasks with each device, the Basis and Tom Tom were considered less comfortable than the Garmin, Fitbit, and Misfit.

Overall, the Basis and Tom Tom were viewed as more masculine than the other 4 devices. Similarly the Misfit was rated more feminine than all the devices except the Jawbone. When asked which device they would buy, the Fitbit was rated significantly higher than the Basis (Fig. 1).



Fig. 1. Likelihood to buy each device based on packaging, before use, and after use

A one-way repeated measures ANOVA indicated significant differences for the likelihood to wear each device while exercising, F(5, 85) = 4.35, p < .01, eta2 = 0.20, during one's daily routine, F(3.55, 60.40) = 7.02, p < .01, eta2 = 0.29, and while sleeping, F(3.22, 54.84) = 5.66, p < .01, eta2 = 0.25. Participants indicated they were more likely to wear the Fitbit than the Basis while exercising. During their daily routine, they were more likely to wear the Garmin and Fitbit than the Basis. They also indicated that they would rather wear the Garmin than the Basis or Tom Tom while sleeping.

At the end of the study, participants were asked to rate the importance of the fitness tracking features. The features rated "important" or "very important" most often included the aesthetic/attractiveness of the device, the amount of battery life, being waterproof, containing a heart rate monitor, and having separate modes for different activities such as running, swimming, biking, etc. The features most often considered "not important" or "not important at all" included sleep tracking, smartwatch capabilities (i.e. smart alarms, social media notifications, etc.), having an accompanying phone app, GPS, food logging, and the ability to wear the device in different ways (versatility).

4 Discussion

The results indicate that participants are less willing to purchase a device if the packaging looks cheap. Although there were significant differences in the perceived weight and comfort of the devices, initially this did not create a significant difference in the likelihood to buy each device. After trying on the devices, participants believed they would be more likely to buy the more gender neutral, inconspicuous device (Fitbit or Garmin) over the more uncomfortable, "masculine," device that presented the most difficulty while completing the tasks (Basis or Tom Tom). Not only would they be more likely to buy this type of device, they would also be more likely to wear the device while exercising, during their daily routine, and while sleeping. The results imply that these particular individuals would be more drawn to devices that would not interrupt their daily lives or attract any outside attention. This is consistent with the participants' comments, as they were most likely to make remarks regarding the size and weight of the devices. Although the participants frequently commented that the small, lightweight devices, particularly the Fitbit, would be easy to lose, this did not seem to affect the overall preference. In addition, the most frequent positive comment made about the heavier, bulky devices was the fact that they had a display. As a result, the most favored devices were small, light devices that also had a digital display (Fitbit and Garmin). Typically, individuals indicated they would prefer a device that they could forget they were even wearing. Participants also indicated a desire for simplicity and rated the more complex features such as food logging, sleep tracking, smartwatch capabilities, etc. as "not important" or "not important at all".

The study was limited as the participants recruited were all college students. This may not accurately reflect the overall target population. In addition, participants may have been influenced by the brands of the devices. Multiple participants commented on the fact that they had seen or used other Garmin and Tom Tom products in the past (i.e., GPS). As a result, their past experiences with the brand may have influenced their opinion of the specific device.

Future studies should examine preferences among different populations. For example, understanding the preferences of serious athletes versus the preference of less active individuals would have very important marketing implications, especially if the goal is to appeal to the average consumer. Gender differences and age differences should also be explored. Longitudinal studies examining device preferences over a prolonged usage would also be useful in improving the likelihood of continued use of fitness tracking technology and overall consumer satisfaction.

References

- 1. Comstock, J.: Study: 19 M fitness wearables in use today, to triple by 2018, 25 November 2014. http://mobihealthnews.com
- Kelly, C.: Biggest fitness tracker trends from CES 2014, 14 January 2014. https://everymove. org
- 3. Quinlan, J.: Why fitness wearables can't be one-size-fits-all, 16 April 2014. http://venturebeat. com