



people have when they use a real mHealth device to share data with real people.

To study how new mHealth users share different types of personal information with different recipients over time, we conducted a user study with  $n=41$  participants. To the best of our knowledge, ours is the first study that explores users' privacy concerns by requiring them to *actually* share the information collected about them using mHealth devices; our subjects could decide whether to share the information and if so, how much information to share with others. The device we used for our study is one of the most popular devices, a fitness device called Fitbit. None of the participants in the user study had ever used a Fitbit prior to the study. At least ten participants had previously used a pedometer but had never uploaded its data to a website or online application.

A majority of the focus group participants considered fitness information to be less sensitive when compared to other types of health information; people might be more concerned about sharing health information like pregnancy status, onset of stress and HIV status, than fitness information like the number of steps they took and calories burned. We nonetheless expect our findings to shed light on the case of mobile devices and applications that collect any personal and activity information, not restricted to health or fitness information.

In this paper, we answer the following questions:

- Did participants share different types of personal or sensed information more or less frequently?
- Do participants' decisions about sharing health information differ across types of sharing partners (family members, friends, third parties, and the public)?
- Does sharing behavior change over time; are participants' privacy preferences dynamic?

We confirmed that people's sharing behavior depends on the type of information being shared and the sharing recipient. Our results showed that participants were generally less willing to share personal demographic information or context information collected by the mHealth device, than about sharing the health information that the device is meant to collect. Our results also showed something surprising – study participants were more willing to share some information with *strangers* than with their own family and friends; among strangers, they were more willing to share some information with specific third parties than with “the public” at large. We also confirmed that people's privacy behavior is dynamic; participants' sharing behavior changed during the course of our study. It is important to understand people's willingness to share, so that mHealth devices can provide patients with the controls to share their information in a manner such that they can enjoy the benefits provided by the device without disclosing more information than is necessary.

In this paper, we use the term “user” to denote the mHealth device user and “sharing partner” to denote the person(s) with whom the user shares her fitness information.

## 2. METHODS

We conducted a social experiment to examine users' decisions to share particular types of information with various types of information recipients (and requesters) over a 5-day period. During the study, users were asked to carry a Fitbit [8], a popular mHealth device that uses an accelerometer to estimate a user's calories burned, steps taken, distance

	Male	Female	Total
Students	8	13	21
Working	5	7	12
Retired	0	8	8
Total	<b>13</b>	<b>28</b>	<b>41</b>

Table 1: Participants

traveled, and sleep quality. During the five days of the study, each subject was asked to wear the Fitbit at all times, except when swimming, bathing or any time they felt uncomfortable wearing it. They were asked to upload the collected data at least daily to fitbit.com. Unfortunately, fitbit.com only provided users with limited coarse-grained data sharing options, and no mechanism for monitoring sharing behaviors. So, we developed a custom web interface that displayed both uploaded Fitbit data and personal traits and allowed users to share data with others. Participants used this interface (instead of fitbit.com) to view their data and make sharing decisions, throughout the study.

The goal of the user study was to understand people's willingness to share their personal health/fitness information with family, friends, third parties and the public. Previous work has shown that young and old people have different views about sharing health information [4, 7, 12]. So, we recruited a sample of college students, working adults, and retirees for the user study. We recruited 21 undergraduate students, 12 adult workers from the local area including Dartmouth employees, and 8 female elderly residents of a local retirement home, as shown in Table 1. It was not an aim of the study to understand the influence of gender and occupation on privacy concerns, so we did not focus on the distribution of participants among these categories. The recruitment flyer presented the study as a study of a new device to help individuals trying to lose weight and/or improve fitness and health. To avoid self-selection bias, the participants were not told that the study was about privacy. Subjects were required to own a computer, to be injury-free, to be able to walk and carry the device with them during the five days. Study participants were paid for their time.

We did not retain any sensitive information, like the participants' fitness information collected using the Fitbits, after the study; we stored only the sharing settings that they chose. Study participants were debriefed after the study to make them aware of the deception used in the study and to inform them that the goal of the study was to understand their privacy concerns and not just to collect their activity data, and that we shared their data only with the people they chose as their sharing partners. This study, including our use of deception and subsequent debriefing procedure, was approved by Dartmouth's Institutional Review Board.

To study participants' willingness to share secondary information, apart from the primary sensed information, we also collected other related personal information about each participant, including his or her age, gender, height, weight, health goals, overall activity level and academic major. Henceforth, we refer to these seven characteristics as “traits”.

To understand how participants share information with real people, we asked them to select family members and friends to receive their shared information. Throughout the study, participants also received requests to share data with specific third parties. These specific third parties represent

academic researchers, medical labs, private companies, and the government. The third parties were real, but the requests were fake. (For example, one of the email requests was from a fictitious group of students at Harvard University, requesting activity data for use in a machine-learning class.) Each participant also had a “public profile” available on the website with their Fitbit data; we told the participants that this profile was visible to anyone who had the url (unless they changed the setting, as described below).

The website provided opt-out sharing settings; by default, all the collected information was shared in the finest detail unless participants changed the settings to “opt-out” of sharing. We used an opt-out policy (instead of opt-in) to be consistent with the majority of online applications now available in which the default privacy setting is to share, with the option to “opt-out” of sharing. We understand, though, that people’s sharing decisions are influenced by default settings [24]. Although we agree that opt-in settings give more control to the user, for our study we used opt-out to provoke action (visit website and change settings) that we could observe among those with privacy preferences.

We wanted to study people’s willingness to share their information, and not how they adapted to the device and controls on the website. So, the study was divided into two phases; a learning phase in which participants were given two days to get used to the device and website, and the study phase, in which we observed participants’ sharing settings during days 3-5 of the study.

**The learning phase — day 1 and day 2.** The researchers met with participants, individually, on both days to answer any questions they had about using the device and the website. On the first day of the study, participants were asked to select at least one family member and two friends with whom to share their information. An email was sent to these sharing partners, informing them about the study and asking them to be a part of the study.

On the second day, we told the participants that their information would be shared with their family and friends from the next day onwards and that they could decide, by using the controls on the website, whether and what they wanted to share with their family members and friends. We also informed participants that over the next few days they might get requests from third parties to share their information but that they could use the controls on the website to limit sharing of their information; they were required to visit the website for each third-party request so that we could observe their sharing choices. Similarly we informed them that their data on the website would be open to the public but they could use controls to opt-out of sharing. We did not tell the participants that the third-party requests were fake or that their information was not actually exposed to the public.

**The sharing phase — day 3 to day 5.** On the third day of the study, an email was sent to the family members and friends with a link to a webpage where they could see the participant’s shared Fitbit information and traits. Throughout the study, participants could change the sharing settings for each type of information and for each sharing partner. The Fitbit-collected activity data (i.e., steps, calories and sleep) could be shared in 5-minute, hourly, 6-hourly, or daily summaries, or it could be not shared at all. By default, activity data was shared at the maximum setting, i.e., 5-minute granularity. Participants were also able to share or hide their personal traits (age, height, weight, gender, activity level,

health goals, academic major), independently by type and for each sharing partner. Participants also received emails everyday during the five-day period of the study, containing status information. From the third day of the study onwards, these mails explained who was receiving their information and what sharing settings they had chosen for each information type for each sharing partner. The message also contained a link to the site where they could change these settings at any time.

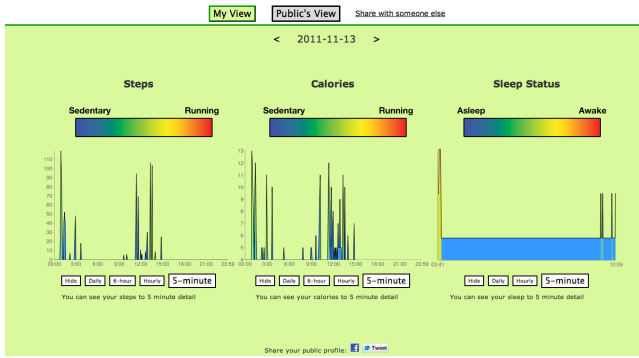
The web interface showed different views of the same information. The default view, shown in Figure 1(a), was the participant’s view. The interface also had views corresponding to each sharing partner; see Figure 1(b). On these views, the participant could decide what information she wanted to share with her sharing partner, make changes with the click of a button and observe exactly what her sharing partner will be able to see; the goal of our website design was to reduce the disconnect between sharing controls and the information being shared.

The emails from third parties, requesting to access participant information from the website, including all information and their email addresses, were sent by us as if from 6 different organizations/groups. Each third-party email explained who the group was and why they wanted the data. The groups identified were: college students, a research lab, a government agency, an engineering company, a wellness institute and a pharmaceutical company. Here we do not examine differences in willingness to share between these different types of third parties. Instead we analyze sharing behavior with all third parties as a category to compare to other sharing partner categories (family, friends, public). We told the participants that the data on the website to be shared with third parties was not anonymous (email address was shared), but that their email address would be used by the third parties only if the researchers needed to contact them. The email address was, however, not visible on their public profile and hence not shared with the “public”. The order of third-party email requests was randomized, and the number of requests varied across the days (three on the third day, one on the fourth and two on the last day). We did not actually share the participants’ data with any third-party organizations, but crafted the messages to be as believable as possible.

We monitored participants’ website activity through logs, that recorded when they logged in, when they looked at their own data view or the views of their various sharing partners and when they changed the sharing settings. To understand reasons behind participants’ sharing behavior, we conducted post-study interviews in which we asked the participants several questions: whether they ever took off the device to hide any information, whether they ever changed the sharing settings and why, and whether they fell for our deception (that the study was really about privacy-related behavior, and that no data was actually shared with the public or with third parties). After the interviews, we revealed the deception and explained to them that the goal of the study was not to collect their fitness information, but to observe their sharing behavior. We recorded the interviews, then transcribed and coded them manually.

## 2.1 Analytic methods

We conducted quantitative analysis of the website logs to understand the participants’ sharing settings and a qualita-



(a) Own view



(b) Sharing partner's view (Mom's view)

Figure 1: Screenshots of views on the study website

tive analysis of the post-study interviews to give us insight into the reasons for their sharing behavior. We use the analysis to answer the questions listed earlier.

## 2.2 Measuring sharing behavior.

To measure participants' willingness to share their information with a group of sharing partners, we defined a *sharing score* for each participant. The sharing score was computed as follows:

$s(u, p, t, d)$  is the setting chosen by user ( $u$ ) on day ( $d$ ) when sharing information type ( $t$ ) with sharing partner ( $p$ ). For steps, calories and sleep, the setting can be 0, 1, 2, 3 or 4, which corresponds to the five possible settings: hide, share daily summary, share 6-hourly summary, share hourly summary, and share 5-minute detail respectively. For the other information types, the setting can be either 0 (hide information) or 1 (share information). By default,  $s(u, p, t, d)=4$  for  $t=\{\text{steps, calories, sleep}\}$  and  $s(u, p, t, d)=1$  for  $t=\{\text{age, gender, health goals, height, activity level, academic major, weight}\}$ . That is, the default setting is the maximum sharing setting. We normalize each score by dividing by  $\max(t)$ .

$$\mathbf{max}(t) = \begin{cases} 4 & \forall t \in (\text{steps, calories, sleep}) \\ 1 & \text{otherwise} \end{cases} \quad (1)$$

To represent the amount of information shared with a category of sharing partners (i.e., family, friends, third parties, public), we computed an overall sharing score for that category, labeled as "group sharing score". Equation 2 defines the group sharing score for participant ( $u$ ) on day ( $d$ ) as the mean of all sharing scores for type ( $t$ ) for each member of the sharing partner group ( $g$ ). For example, if a participant identified two friends then the friend group sharing score for weight would be the mean of the sharing scores for weight for the two friends.

$$\mathbf{grpscore}(u, g, t, d) = \frac{1}{|g|} \sum_{p \in g} s(u, p, t, d) \quad (2)$$

To measure how personal traits were shared, we also computed a combined sharing score for traits, which is the average of the sharing scores of all the personal traits, labeled *traits*.

We focus most of our analysis on two snapshots: the initial sharing score,  $\mathbf{grpscore}(u, g, t, 2)$  for the last setting on day 2 for each type  $t$ , and the final sharing score,  $\mathbf{grpscore}(u, g, t, 5)$ . We normalize each score by dividing by  $\mathbf{max}(t)$ , on day 5. Unless mentioned otherwise, we used

t-tests to compute the difference between normalized sharing scores, to understand how different information types were shared with the different groups; paired sample t-tests were used to compare different behaviors of a single subject, whereas independent sample t-tests were used to compare one subject to another. We used *analysis of variance* (ANOVA) post-hoc testing using Bonferroni's method to compare the sharing behavior of students, employees and retirees; whereas the t-test is used to compare the means between two groups, ANOVA is a statistical procedure used to compare means between three or more groups.

## 3. RESULTS

We present below the results from the quantitative and qualitative analysis. We highlight only some of the important results in this paper; full results are available in the technical report [20].

### 3.1 Information sharing analysis

27 of the 41 participants identified both a family member and friends with whom to share information. Among the 27 participants, the mean number (and standard deviation) of family members and friends that the participants chose were 1.26 (0.44) and 2.11 (0.5) respectively. Two participants did not identify a family member, but did identify friends, while 11 participants did not select any friends or family members (the reasons given by these respondents included privacy concerns, not wanting to bother them, and expectation of lack of interest, as discussed in Section 3.2). These 11 participants, however, were not considered when we computed sharing scores for friends and family. Similarly, when comparing sharing scores between family and public, we used the scores of all participants who selected at least one family member.

*Did participants share different types of personal or sensed information more or less frequently?* Table 2 shows the final normalized sharing scores for family used to evaluate within-subject differences in sharing different types of information.

Table 2 shows that with family members, the study participants shared weight and health goals less than age, academic major, activity level and the sensed information (steps). It is not surprising that participants were willing to share obvious information known to their family members such as age and gender, but at least some participants seemed to consider information like weight and health goals as more private. In

**Table 2: Mean and (standard deviation) of final normalized sharing scores for family**

Information	Family
Steps	0.96 (0.19) <sup>†</sup>
Calories	0.96 (0.19)
Sleep	0.96 (0.19)
Activity	0.93 (0.26)
Age	0.96 (0.19)
Gender	0.93 (0.26)
Goals	0.86 (0.36) <sup>†</sup>
Height	0.93 (0.26)
Major	0.95 (0.21)
Weight	0.84 (0.36) <sup>†</sup>

<sup>†</sup> Sharing score for sensed information (steps, calories and sleep) is significantly higher than for weight and goals, sharing score for weight is significantly less than the score for most personal traits (age, gender, academic major, height and activity level) and sharing score for goals is significantly less than the score for age,  $p \leq 0.1$

the post-study interviews, some participants were reluctant to share this information because they worried that family members might judge them and even reprimand them. A web interface might influence sharing behavior, but in the web interface we built, however, sensed information was more prominent than personal traits, so the sensitivity of weight and health goals had nothing to do with the interface layout.

***Do participants’ decisions about sharing health information differ across types of sharing partners (family members, friends, third parties, and the public)?***

Table 3 shows the normalized final group sharing scores for sensed information (steps, calories, sleep), traits (the combined score), weight, goals and academic major across three comparison categories: family vs friends, family vs public, public vs third parties. Recall that a score of 1 implies that the information has been shared to its maximum with all the sharing recipients in that group.

**More information shared with family than friends.** Subjects shared weight with family significantly more often than with friends. The sharing scores for friends were marginally less than that for family members for all other information types as well, but the differences are not statistically significant. From the interviews, we learned that some participants were more concerned about sharing their information with their friends because they were worried of being judged by their friends more than by their family, especially in the case of students, since they see their friends everyday.

**More information shared with family than public.** Table 3 shows that participants shared more information about their steps, calories and sleep with family than with the public. Participants said that they felt uncomfortable sharing sensed information with strangers because it made them feel like they were “being watched”. Not surprisingly, participants also shared personal traits significantly more with family than with public.

**Less information shared with public than with third parties.** Table 3 shows that participants were generally more open to sharing weight and health goals with specific third

**Table 3: Final normalized sharing scores for family vs friends, family vs public, and public vs third parties (TP)**

	Family vs Friends		Family vs Public		Public vs TP	
	Family	Friends	Family	Public	Public	TP
Steps	0.96 (0.19)	0.94 (0.14)	0.96 (0.19)	0.91* (0.22)	0.89 (0.27)	0.89 (0.25)
Calories	0.96 (0.19)	0.94 (0.17)	0.96 (0.19)	0.89* (0.25)	0.89 (0.27)	0.89 (0.25)
Sleep	0.96 (0.19)	0.95 (0.14)	0.96 (0.19)	0.87* (0.30)	0.87 (0.30)	0.87 (0.29)
Traits	0.91 (0.23)	0.83 (0.26)	0.91 (0.23)	0.78* (0.29)	0.80 (0.29)	0.83 (0.32)
Goals	0.85 (0.36)	0.73 (0.42)	0.86 (0.36)	0.61* (0.50)	0.68 (0.48)	0.82* (0.38)
Major	0.94 (0.21)	0.91 (0.25)	0.95 (0.21)	1.00 (0.00)	0.98 (0.16)	0.82* (0.38)
Weight	0.83 (0.37)	0.64* (0.46)	0.84 (0.36)	0.54** (0.51)	0.59 (0.50)	0.74* (0.43)
	n=27		n=28		n=37	

<sup>†</sup> Final scores for the two groups are different,  $p \leq 0.1$

\* Final scores for the two groups are different,  $p \leq 0.05$

\*\* Final scores for the two groups are different,  $p \leq 0.01$

parties than with the public. In the post-study interviews, some participants said this was because they perceived some benefit in sharing information with specific third parties. Third-party request emails contained a reason for wanting the participants’ data and at least some participants apparently expected the third parties to use their data for the purposes mentioned in the email. In contrast, some participants expressed concern about who among the public would be accessing their information or how they might use it.

Surprisingly, participants were less willing to share academic major with the specific third parties than with the public. Some participants said during the post-study interview that they shared information with specific third parties because they thought that the information would be useful, based on the purpose stated in the third-party request. Some of them felt that academic major was not relevant to the request.

Given the comparison on personal traits, we were surprised to see no difference in sharing of sensed information between public and third parties.

***Does sharing behavior change over time; are participants’ privacy preferences dynamic?*** In Table 4, we show select traits and sensed information by the initial (end of day 2) and final (end of day 5) sharing scores for three sets of sharing partners: family, friends and public.

**Sharing sensed information.** For family, participants did not change sharing behavior of sensed information over the course of the study, while for friends, there was a slight (non-significant) change. For the public, however, we found that there was a statistically significant reduction in sharing scores for steps and sleep. Some participants felt uncomfortable sharing their steps and sleep, as the study progressed; they said that they felt like they were being watched.

**Sharing traits.** Similarly, in the case of the trait information, there was a slight (non-significant) reduction in sharing scores for family. However, there was a statistically signifi-

**Table 4: Initial and final normalized sharing scores for family, friends and public**

	Family		Friends		Public	
	Initial	Final	Initial	Final	Initial	Final
Steps	0.96 (0.19)	0.96 (0.19)	0.97 (0.10)	0.95 (0.14)	0.94 (0.22)	0.88* (0.26)
Calories	0.96 (0.19)	0.96 (0.19)	1.00 (0.02)	0.95 (0.17)	0.93 (0.23)	0.88 (0.27)
Sleep	0.96 (0.19)	0.96 (0.19)	0.98 (0.09)	0.95 (0.16)	0.93 (0.22)	0.87† (0.30)
Traits	0.95 (0.19)	0.91 (0.23)	0.93 (0.13)	0.84* (0.26)	0.92 (0.17)	0.80** (0.29)
Activity	0.96 (0.19)	0.93 (0.26)	0.95 (0.20)	0.84† (0.36)	0.98 (0.16)	0.85* (0.36)
Goals	0.93 (0.26)	0.86 (0.36)	0.88 (0.29)	0.75* (0.41)	0.83 (0.38)	0.68* (0.47)
Weight	0.89 (0.31)	0.84 (0.36)	0.80 (0.35)	0.63** (0.46)	0.83 (0.38)	0.61** (0.49)
	n=28		n=29		n=41	

† Initial and final scores are different,  $p \leq 0.1$

\* Initial and final scores are different,  $p \leq 0.05$

\*\* Initial and final scores are different,  $p \leq 0.01$

**Table 5: Final normalized sharing scores based on gender**

	Female	Male
grpscore(u, Friends, Traits, 5)	0.78	0.95 †
grpscore(u, Friends, Weight, 5)	0.47	0.95 **
grpscore(u, Friends, Goals, 5)	0.64	0.95 *

† Sharing scores of females and males are different,  $p \leq 0.1$

\* Sharing scores of females and males are different,  $p \leq 0.05$

\*\* Sharing scores of females and males are different,  $p \leq 0.01$

cant difference between the initial and final sharing scores for friends and for the public. We learned from the post-study interviews that some participants were embarrassed to share certain personal traits with friends and concerned about sharing their personal traits with strangers; they might have realized it only seeing their data over time. More details of the post-study interviews are given in Section 3.2.

**Demographic differences.** Though the study was not designed to examine differences in sharing by characteristics like occupational status or gender, we did find some differences across these characteristics and so present them as preliminary findings that are suggestive of future study. As shown in Table 5, independent sample t-tests revealed that females shared traits (the combined score), weight and goals with friends, significantly less than did male subjects. The table shows only the difference in sharing with friends, but there was a statistically significant difference in the extent personal traits, weight and activity level were shared with public and third parties as well [20].

As shown in Table 6, students shared their weight with family more than employees shared with family. In contrast, they shared their health goals less with the public, than employed adults did with the public. Some of the employees said they did not want to share weight information with

**Table 6: Final normalized sharing scores - students vs employees**

	Students	Employees
grpscore(u, Family, Weight, 5)	0.90	0.50 †
grpscore(u, Public, Goals, 5)	0.48	0.83 †

† Sharing scores of students and employees are different,  $p \leq 0.1$

**Table 7: Final normalized sharing scores - students vs retirees**

	Students	Retirees
grpscore(u, Public, Traits, 5)	0.74	1.00 †
grpscore(u, Public, Weight, 5)	0.48	1.00 *
grpscore(u, Public, Goals, 5)	0.48	1.00 *

† Sharing scores of students and retirees are different,  $p \leq 0.1$

\* Sharing scores of students and retirees are different,  $p \leq 0.05$

their family to avoid discussion about weight management. Students considered health goals to be sensitive and did not want to share this information with the public, but surprisingly, employees were more willing to share this information with the public.

Students were much more concerned than retirees about sharing their personal traits, weight and goals with the public, as shown in Table 7. (*Caveat:* Recall that we had only female retirees; given that females shared less than males with friends in Table 5, we are not sure why female retirees shared more with the public than students.) There are several possible reasons for this difference in behavior: we speculate that the retirees are not used to the technology or do not want to bother their family and friends by sharing with them information that the retirees feel will not be of interest; when they do share this information with others, they are less concerned about the information than students. We expect students, on the hand, to be used to the technology and used to sharing information electronically with others. We speculate that they might have changed the default settings either because they were curious about the different settings or because they were really concerned about what they were sharing with others. We expect students and employees to have more reasons to be worried about their activities and to hide it from their family and friends than retirees; either because they were embarrassed about some information, maybe their weight, or they wanted to hide some information, like partying or sexual activity. Students and employees were more engaged in the study than retirees. In Section 3.2, we present anecdotal evidence of such behavior and concerns.

To summarize, we found that weight and health goals appeared to be most sensitive among the information collected during the study. Participants exhibited disparate and dynamic sharing behavior of this information. We found some evidence that sharing behavior might vary with occupational status and gender. After observing the participants' sharing behavior, we wanted to understand the reasons for their sharing behavior, as discussed in the next section.

## 3.2 Post-study Interviews

To give us insight into the reasons for participants' sharing behavior, we conducted post-study interviews. We discuss the answers to the following questions: *“Did you change the sharing controls on the interface at any point? If so, what influenced that decision? Did you change the sharing controls for X? If so, why?”*. For answers to other questions, please refer to our technical report [20]. We recorded the interviews. We coded the interviews manually and grouped the statements into categories. We discuss below the reasons for the participants' sharing behavior.

**Amount of information collected.** Three participants mentioned that they felt the information was not sensitive because the device collected information only for five days. One female student said that the reason she shared the information with third parties was because *“it was a study and it wasn't very long.”*

**Context of data collection.** A few students were concerned about sharing information that was collected by devices while they were at parties or staying up late. A male employee asked, *“would [the device] be something you would keep on during sexual activity or when you go to the bathroom?”*

**Sensitivity of the data.** One male student shared all his information with others, but said that he might have had concerns about sharing *“if maybe I was someone who [was] trying to exercise more and I exercised less.”* A few female students did not want to share their activity information because they felt like they were being watched. One of them said, *“they can see every step I take, that was just a little weird.”*

**Information utility.** Some participants decided to share their information with others depending on how they would use the information. One female student said, *“I think I hid my weight from almost everybody, except for people who actually needed it for medical purposes.”* One employee, being a researcher, was open to sharing her health information with other researchers. Some participants considered the third parties differently, based on their reason for requesting the data. One female student said, *“I was fine with sharing things [with universities]; for some reason, they felt a lot more legitimate, you know what they would be doing, studying. It was random people that I didn't know what they were doing that I [did not want to share my information with].”* One male student said he would share information with everyone, as long as it would not affect him in the future when he was applying for insurance or jobs. He said, *“I would be fine with all of those, with the exception if that has an impact on the ability to apply for insurance or something of that nature, in which case I would start to worry.”*

**Anonymity.** A few students felt their information would not be linked to them (even though they were aware that their email was being shared), so they were comfortable sharing it with third parties. According to a female student, *“They don't know who I am, they are just doing research.”* One male employee felt that his identity is linked more to his name than his email. He said, *“It's my name, but I control [my email]. I control what I get, I can change my email address.”*

**Sharing partners.** Most participants said they would share information depending on who it was being shared with. One female student said *“If there was someone who was a lot heavier than me, I probably would have given them*

*the 5 minute calories, because they might feel bad that I used so many calories throughout the day. With friends who were less active than me, I would have shared less.”*

**Partner involvement.** One male student was happy to share his activity information with others; he said when you share activity information, *“you feel like other people are in this with you, it makes it easier to keep going”*, and he said encouraging feedback from his friends made him feel good about sharing his information. A female student shared her Fitbit information with third parties, because according to her, *“when I was wearing [the device] and getting data requests all the time, it felt like what I was doing was important”* and she was disappointed that the requests were fake, and to her, it meant that *“no one actually cares about your data and no one's going to use it, it was all for nothing kind of thing.”* Some students did not want to share certain information due to fear of being judged by others. A female student said, *“With my friends, I wasn't sure whether to share my height and weight, because sometimes especially if I am sharing with my girlfriends, oh they are like you are heavier than me, lighter than me.”* Another female student said, *“I don't mind articulating [my health information] in person, but on a website, I feel it is more easily judged in the wrong way that I can't fully explain what is going on.”*

**Negative experiences.** A female student was concerned about sharing information with the government, because she grew up in a country where *“everything was monitored by the government.”* One female employee was sharing all her information with the third parties during the study, until she noticed that she started getting spam about weight loss, which must have been coincidental.

**Relationships.** One elderly participant was not comfortable with sharing her activity information with her children. She said, *“I didn't want them to have to encourage me to walk more. They don't need to know. We are very very close but they don't need to know how much I walk.”* A male student said, *“I told [my mom] I would tell her of any results of any significance, but I told her that I was hiding the data and I wasn't going to let her see it. Honestly, my friends didn't care about the data.”*

Some students were more comfortable sharing their information with family than their friends. One said, *“If it was someone I didn't know I would share everything. Friends they know you, but you are not close enough to share everything with them. I shared everything with my mom.”*

Some female students were more comfortable sharing personal information with their family and third parties than with their friends. One student said *“I might have left height and weight with family, but friends don't need to know that. I shared it with companies and researchers, because I think it is pertinent.”*

Some students were more comfortable sharing their information with family and friends than third parties. Two of them said, *“I didn't share any information with the extra researchers. I don't know who they are and I have no affiliation with them.”* and *“I don't know [the requesters]. I don't think it is weird that they were asking for [the information], but it was weird sharing with them. From teammates, hid my weight and my health goals. From my mom, I didn't hide anything.”*

Some students wanted to share their information with third parties more than with people they knew, like their family and friends. They said they wanted to share less information

with family and friends: one participant said *“because I know them personally, whereas the third parties they seem, not that personal... so I felt like more of a pressure to hide more specific activity levels from [my family and friends].”* Another said *“Because my parents are people who are big on exercise. If I don’t do much exercise, they wouldn’t like that”,* while a third participant said *“A bunch of researchers looking at the data, I don’t care. But I might think twice about some people I know, depending on who they are.”* Another participant said *“People who don’t know me it would be fine. My age doesn’t bother me, it would be mostly my weight. It all depends on who gets it, what is the purpose. If it is somebody studying what is the better way to do things.”*

Some participants did not want to share information with private companies. A male student said, *“I’m against corporations. I probably wouldn’t want any of them [to have access to my information], except students.”* A male employee said, *“Oh yeah, I would share that info [with students]. With individuals, with family members or friends who are interested and people doing research I have no problem. It is just third-party companies [that I wouldn’t want to have the data].”*

One elderly retiree was not tech-savvy; her husband was helping her manage her Fitbit account and he might have had an influence on her sharing decisions.

**Information types.** We asked the participants whether they would use a mobile device that collects personal health information like their heart rate, breathing rate, pulse rate, medication, diet and exercise, location, social interactions, if it gave them similar sharing controls as the Fitbit in our study. Most students considered medications to be most sensitive. A male student said, *“Just like bodily functions, you can’t really use that against you, whereas medication you are taking, that’s something like, there are some medications people don’t want other people finding out that they are taking.”* Some of them were worried about sharing location and social interactions. Students who were athletes were concerned about sharing their vital signs and exercise information. One female employee was open to sharing any information *“as long as [she] could control who saw what”*.

## 4. DISCUSSION

The user study helped us identify the key factors that influence privacy decisions regarding health and fitness information collected using mHealth devices. The findings from this study can help guide mHealth device and application developers and privacy advocates to build flexible privacy controls for mHealth devices, with sensible defaults and expressive controls for users to change the settings thereafter.

Our study revealed three interesting findings about people’s privacy concerns regarding their sensed health information.

**1. Demographic information shared less than sensed information.** The study revealed that participants were less willing to share the demographic information we collected than the activity information that was sensed by the device. For example, initial sharing scores for weight and health goals were less than the initial sharing scores for other information types, including the sensed information. We expect users to be concerned about sharing certain context information, depending on how it might affect the value they perceive in the information being shared. For example, a user might share her location information when it is being collected by her asthma sensor and shared with her mother,

but she might not want to share her location, when it is collected as part of her activity information by her fitness device.

**2. Information shared more with strangers than their own family and friends.** We discovered that sharing scores for friends were lower than scores for family, while sharing scores for friends were lower than for third parties. However, sharing scores for public were mostly the same or less than friends. The post-study interviews revealed different reasons for participants’ sharing behavior, including their relationship with the sharing recipients. Participants were more willing to share if they perceived benefits in sharing, especially when it came to sharing with specific third parties, as opposed to the public.

**3. Dynamic sharing behavior.** We confirmed that privacy concerns are not static; mHealth device users may change their sharing decisions over time.

Given these findings, we elaborate on two recommendations that will help guide the development of flexible privacy controls that enable users to express their sharing preferences easily.

**Flexible controls need to support both fine- and coarse-grained approaches to sharing.** Throughout the study we saw a wide diversity in sharing behavior. Some participants used a very coarse-grained approach, while others took the time to fine-tune their privacy settings. Sensible default settings are required to support those users who never change their sharing settings, either because they are busy, lazy, not tech-savvy or want immediate benefits from the system. The availability of granular controls encouraged participants, who were averse to sharing everything, to share some information, instead of hiding all information. Participants expressed disparate sharing preferences and exhibited dynamic sharing behavior in our study, which implies that default “one size fits all” settings are not enough.

We observed contradicting behavior among participants; some participants shared more with their friends than with family, while others shared more with their family than with their friends. We also observed dynamic sharing behavior; some participants changed the amount of information they shared during the course of the study. Granular levels of sharing and expressive controls, which we discuss next, can help such users change their sharing setting easily to map their preferences.

**Reducing disconnect between information and granular controls.** Users make their sharing decisions based on what information they are sharing and who they are sharing it with. Since sharing decisions are dynamic, the information should be clearly presented and the controls flexible and easy to use, to allow the participant to map their privacy preferences easily. Narrowing the gap between settings and what is actually shared can help users change their behavior easily to suit their sharing preferences. For example, in our study, the website home page for every participant was divided into different views, one view corresponding to one sharing recipient, where the participant could decide what information she wanted to share with that recipient. View for sharing recipient “Mom” on the participant’s home page displayed exactly what Mom would see as the participant’s health information. By combining the information and the granular controls, the interface made it possible for the participant to observe what Mom would see for different choices of sharing settings and finally, choose the setting that best mapped her



privacy preferences. We did not test the usability of the system, so we cannot claim that the our design is the best way to provide granular controls for sharing health information. Designing an interface for an mHealth device and application that collects a large amount of sensed, personal demographic and context information and whose user has the option to choose a large number of sharing recipients is an interesting and challenging problem.

User studies, like ours, could benefit from a bigger sample size, better population sampling and longer duration. Nevertheless, the study helped us understand people’s willingness to share and their dynamic sharing behavior. We expect these findings to hold broadly for other mHealth devices and applications as well. A general privacy setting for all mHealth devices is not possible, given the disparate sharing behavior among users for even a single mHealth device. We recommend seeking a general approach to health information visualization: a flexible design that supports all mHealth devices and allows users to also visualize how they are sharing their health information with others.

## 5. RELATED WORK

Previous work has looked at people’s willingness to share information with others. Previous work suggests that users will change behavior when the context of information sharing varies [18]. For example, studies of location tracking show that at least some users will vary their willingness to share depending on place and social context [2], time since the start of information sharing [23], recipients [6], and their closeness to the recipients [27]. Other studies of context show that users are more likely to reveal information when the reward from the exchange increases, but less likely to do so when risk of identity theft increases [3]. Our findings confirm that these results hold even for people’s willingness to share their health information.

**Willingness to share.** Previous work has shown that people make privacy decisions based on the information being shared and the person they are sharing it with. It has been shown, through surveys and interviews, that users share location information based on the sharing recipient, why the recipients want the information, what would be useful to them and whether the users want to disclose that information with them; during the study, users received hypothetical sharing requests from family and friends [6]. (Our findings confirmed that people use the same sort of logic to make privacy decisions with health information). An online survey showed that participants do not understand the value of sharing location information and their privacy decisions depend on the sharing recipient [25]. The above two studies, however, never gave the participants an opportunity to actually share the information with real people, but just learned about their sharing preferences through interviews and surveys. People may not be aware of real privacy risks until they actually share the information with others and receive feedback about the sharing [5]; our study gave participants the opportunity to actually share the information with real people. Our study showed that people did have privacy concerns about sharing certain information types, but they changed their sharing settings during the course of the study. Our findings confirmed results from previous work, which showed that participants change their privacy policy decisions with time, but participants in that study knew the information was not being shared with real people [16]. The manner in which peo-

ple think they might share their information changes once their information is actually shared with real people; the findings from our study are more valid than previous work, because our study participants actually shared information with real people. Also the other studies were focused on understanding privacy concerns while sharing other types of information; we wanted to understand how sharing behavior changes when it comes to health and fitness information. Certain types of health information might be more sensitive than other types of personal information, like location, so it is important to study people’s privacy preferences regarding health information.

**Health information.** Maitland et al. conducted interviews to understand the role of peers in weight management and what information people are willing to disclose to their peers [14]. We, too, wanted to understand users’ willingness to share their fitness information, but we gave users an opportunity to actually share their own fitness information with family, friends and third parties. Olson et al. conducted surveys with employees (median age of 35) to study people’s willingness to share their personal information, including pregnancy and health status, with others and they identified similarities in what people wanted to share and who they wanted to share it with [19]. Our work is different from theirs in that we conducted a study with young students, employees and retirees, where subjects collect information about themselves and actually share that information with real people (or in some cases, believe that their data is being shared with actual people).

**Sensors and mobile devices.** Klasnja et al. study the privacy concerns of patients using a fitness device by conducting interviews [15]. We also study privacy concerns of patients using a fitness device, but we focus on their willingness to share the collected information. Rajj et al. showed that people are more aware of privacy risks once they receive feedback about their shared health information, collected using mobile sensors, and have a stake in the data, i.e., if the shared health data is their own [22]. The study participants (in this case, students) filled out a survey after seeing feedback about their information for 10-15 minutes. In our study, we allow the participants to share the collected information with real individuals chosen by them and study how willing they are to share their activity and sleep information with friends, family and third parties. To the best of our knowledge, ours is the first study that explores users’ privacy concerns by giving them the opportunity to actually share the information collected about them using mHealth devices.

## 6. SUMMARY

To provide flexible and expressive privacy controls, it is important to understand users’ willingness to share their personal health information collected using the device. Other researchers used interviews and surveys to understand people’s willingness to share; their results might not reflect real privacy concerns, since people remain unaware of real privacy risks until they actually share the information with others. We conducted a user study to understand how willing users were to share their personal health information that they collected using an mHealth device that they carried with them at all times for five days. We recommend a flexible design for sharing controls that narrows the gap between controls and the information being shared, allowing patients

to visualize how they are sharing their health information with others.

## Acknowledgements

This research results from a research program at the ISTS, supported by the NSF under Grant Award Number 0910842 (TISH) and Award Number 1143548 (PC3) and by HHS (SHARP program) under Award Number 90TR0003-01. We also thank the anonymous reviewers, undergraduate research interns Alexandra Della Pia and Tina Ma, and our colleagues in the Dartmouth TISH group, for their valuable feedback.

## 7. REFERENCES

- [1] Affectiva. <http://www.affectiva.com/>, January 2011.
- [2] D. Anthony, T. Henderson, and D. Kotz. Privacy in location-aware computing environments. *IEEE Pervasive Computing*, 6:64–72, 2007.
- [3] D. Baumer, J. Earp, and J. Poindexter. Quantifying privacy choices with experimental economics. In *Workshop on Economics of Information Security (WEIS)*, 2005.
- [4] M. T. Britto, T. L. Tivorsak, and G. B. Slap. Adolescents’ Needs for Health Care Privacy. *Pediatrics*, 126(6):e1469–1476, December 2010.
- [5] K. Connelly, A. Khalil, and Y. Liu. Do I Do What I Say?: Observed Versus Stated Privacy Preferences. In *Proceedings of Human-Computer Interaction (INTERACT)*, volume 4662 of *Lecture Notes in Computer Science*, chapter 61, pages 620–623. Springer, 2007.
- [6] S. Consolvo, I. E. Smith, T. Matthews, A. LaMarca, J. Tabert, and P. Powledge. Location Disclosure to Social Relations: Why, When, & What People Want to Share. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)*, pages 81–90. ACM, 2005.
- [7] G. Demiris, B. K. Hensel, M. Skubic, and M. Rantz. Senior Residents’ Perceived Need of and Preferences for Smart Home Sensor Technologies. *International Journal of Technology Assessment in Health Care*, 24(01):120–124, 2008.
- [8] Fitbit. <http://www.fitbit.com/>, January 2011.
- [9] J. H. Frost and M. P. Massagli. Social uses of personal health information within PatientsLikeMe, an online patient community: what can happen when patients have access to one another’s data. *Journal of Medical Internet Research*, 10(3):e15+, May 2008.
- [10] Google Health. <http://www.google.com/intl/en-US/health/about/>, January 2011.
- [11] Microsoft HealthVault. <http://www.healthvault.com/personal/index.aspx>, January 2011.
- [12] C. J. Hoofnagle, J. King, S. Li, and J. Turow. How Different are Young Adults from Older Adults when it comes to Information Privacy Attitudes and Policies? *Social Science Research Network Working Paper Series*, April 2010.
- [13] C. Jensen, C. Potts, and C. Jensen. Privacy practices of internet users: self-reports versus observed behavior. *International Journal of Human-Computer Studies*, 63(1-2):203–227, July 2005.
- [14] M. C. Julie Maitland. Designing for peer involvement in weight management. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)*. ACM, 2011.
- [15] P. Klasnja, S. Consolvo, T. Choudhury, and R. Beckwith. Exploring Privacy Concerns about Personal Sensing. In *Proceedings of the International Conference on Pervasive Computing (Pervasive)*. Springer-Verlag, May 2009.
- [16] M. L. Mazurek, P. F. Klemperer, R. Shay, H. Takabi, L. Bauer, and L. F. Cranor. Exploring reactive access control. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)*, pages 2085–2094. ACM, 2011.
- [17] Monica Healthcare. <http://www.monicahealthcare.com>, February 2011.
- [18] H. Nissenbaum. *Privacy in Context*. Stanford University Press, 2010.
- [19] J. S. Olson, J. Grudin, and E. Horvitz. A study of preferences for sharing and privacy. In *Extended Abstracts on Human Factors in Computing Systems (CHI EA)*, pages 1985–1988. ACM, 2005.
- [20] A. Prasad. Exposing Privacy Concerns in mHealth Sensing. Technical Report TR2012-711, Dartmouth College, February 2012. <http://www.cs.dartmouth.edu/reports/TR2012-711.pdf>.
- [21] A. Prasad, J. Sorber, T. Stablein, D. Anthony, and D. Kotz. Exposing Privacy Concerns in mHealth Sensing. In *USENIX Workshop on Health Security (HealthSec)*, August 2011. Position paper.
- [22] A. Raij, A. Ghosh, S. Kumar, and M. Srivastava. Privacy Risks Emerging from the Adoption of Innocuous Wearable Sensors in the Mobile Environment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)*, 2011.
- [23] N. Sadeh, J. Hong, L. Cranor, I. Fette, P. Kelley, M. Prabaker, and J. Rao. Understanding and capturing people’s privacy policies in a mobile social networking application. *Personal Ubiquitous Computing*, 13(6):401–412, Aug. 2009.
- [24] R. H. Thaler and C. R. Sunstein. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. Yale University Press, 2008.
- [25] J. Tsai, P. Kelley, L. Cranor, and N. Sadeh. Location-Sharing Technologies: Privacy Risks and Controls. In *Proceedings of the Research Conference on Communication, Information and Internet Policy (TPRC)*, 2009.
- [26] WakeMate. <http://wakemate.com/>, January 2011.
- [27] J. Wiese, P. G. Kelley, L. F. Cranor, L. Dabbish, J. I. Hong, and J. Zimmerman. Are you close with me? Are you nearby? Investigating social groups, closeness, and willingness to share. In *Proceedings of the International Conference on Ubiquitous Computing (UBICOMP)*, pages 197–206, 2011.
- [28] Withings blood pressure cuff. <http://www.withings.com/en/bloodpressuremonitor>, November 2011.

## APPENDIX

### A. FOCUS GROUP

This section is taken from the position paper presented at HealthSec 2011 [21], included herein for convenience to the reader.

#### A.1 Methods

We conducted exploratory focus-group discussions to gain a preliminary understanding about Patients' privacy preferences. We have thus far had eight focus-group sessions with 3-7 participants each, who were college students (aged 19-30), hospital outpatients (aged 80-85), or residents of a retirement community (aged 65-100). Each focus group lasted for not more than 90 minutes and all participants were paid for their time. We chose these groups since we wanted to talk to Patients who have some health experiences – some who have been recently hospitalized and others who are monitored continuously outside the hospital – and Patients who have limited healthcare-related experiences. In the next phase, we will conduct a user study to test individuals' sharing behavior.

Since mHealth devices are not yet common, the focus group participants were presented with hypothetical scenarios where mHealth devices were used. There were four scenarios, in which an mHealth device was used to collect a Patient's personal information (measuring medication intake, diet and exercise, location or social interactions); the collected data was uploaded to a private website and then shared with health providers, family or friends. The scenarios for the young and the old differed in the age of the protagonists and their medical condition, but were similar in every other aspect like the information collected and the manner in which it was collected, stored and shared.

We presented each scenario to the participants, after which they were asked about the advantages and disadvantages of using mHealth sensors in that scenario. They discussed their concerns regarding the collection of the particular health information in each scenario, and whether there were certain times and places when they did not want to collect that information. The participants talked about why they would want to share certain health information types with health providers and family members. They also raised some concerns regarding storage of the collected information.

We recorded the discussions. We coded the discussions manually, and grouped the statements into categories.

#### A.2 Results

We summarize comments by type of participant.

**Elderly and young participants** said that

- They are more comfortable sharing health information with doctors than with family members, since they trust doctors to keep the information confidential.
- Some of them recognized that information might get compromised during transmission and on a private website.
- Third parties, like insurance companies, advertisers and government agents, should not be allowed to access health information without the Patient's (or caregiver's) consent at the time of request.
- Health information can be shared with family members if the Patient is incapable of making decisions (like young children or elderly patients suffering from chronic illnesses).

- Most of them rated disclosure of diet and exercise information as a greater privacy threat over that of medication tracking, contradicting our expectations, and, all of them were concerned more about the disclosure of location and social interactions than that of medication and diet and exercise. This confirmed our assumption that Patients' privacy concerns may differ depending the information type being monitored.

**Elderly participants** said that

- They may not always want to share information with family members: a resident of the retirement home felt she wanted to be independent from her family, while one hospital outpatient felt some depressing news contained in the health information might worry his family.
- Doctors might be overwhelmed with irrelevant data (one participant).

**Young participants** said that

- They would be comfortable sharing their medical information, like information about their medication, with their doctor; privacy concerns arise, however, when devices capture other information about their lives.
- The devices can collect any information, as long as the Patient gives consent and it affects only the Patient. Privacy concerns arise when an individual's information gets recorded by devices carried by others around her, without her consent.
- Patients cannot understand complex privacy laws and they give consent without knowing how it will affect them later.
- Privacy laws keep changing and it is unclear whether the information collected now will still be protected by the laws later.

#### A.3 Discussion

The main challenge to building privacy controls for mHealth devices is the wide range of views held by different people. We suggest the following features for a privacy-control interface for an mHealth device:

- *usable*, so that Patients can understand what choices they have and what choices to make,
- *expressive*, so that Patients can express their preferences and decisions,
- *intuitive*, so that Patients can understand it quickly, without any training, and
- *transparent*, so that Patients are aware of what information is being collected, what is being shared, and with whom.

The interface could also provide recommended settings based on the Patient's background. From the focus groups, we discovered that there were some concerns that were specific to Patients who suffered from serious illnesses, distinct from those who were diagnosed with less serious medical conditions. Privacy concerns also vary with age and with the Patients' relationship with their doctor and family members. We expect that it is probable that Patients will have concerns similar to their peers among the focus-group participants. Since every Patient is different, we cannot set preferences for them, but we can provide recommended settings to help Patients make better sharing decisions about their health information.

The interface could provide sensible default settings for the privacy controls, which could be based on the results of the focus groups and with knowledge of the Patient's background and situation. Patients might select the default privacy settings, either because they are not sure of what settings to choose or because they need immediate benefits and have no time to spend on choosing what to share with whom.

The interface should provide access logs, so that the Patient can decide whether more information is being shared than required with a given data consumer, in which case she might want to restrict sharing or she might decide that more information needs to be shared.

The user interface can identify the sensitive content of the information (using annotations or icons), highlight the possi-

ble privacy risks that may result from sharing information, show changes occurring to the information, and reveal when the data is accessed and used, and by whom, to help Patients make sharing decisions.

The interface could provide hierarchical privacy controls – controls that allow patients to specify preferences at higher levels easily and drill down to specific details, if they want to.

Building interfaces that can be used easily by all kinds of Patients is a big challenge, especially since every Patient's needs and values are different. There must be a collaboration with privacy and HCI researchers to work towards building a usable privacy controls interface.