Risk Perceptions for Wearable Devices

Anonymous

Some Place

ABSTRACT

Along with great benefits, wearable devices, or "wearables," bring new potential privacy and security risks which expose users' activities without their awareness or consent. With the additional capabilities of wearable devices and their increasing popularity, people have expressed interest in being notified before data capture [11], but human attention is a finite resource [6]. Therefore, user concerns should be investigated to to warn users only about situations they are likely to care about. Informed, select notifications make for a better user experience and prevents habituation to such notifications while avoiding scandalous breaches of privacy. To this end, we conducted the first large-scale study to investigate user security and privacy concerns for wearable devices. We surveyed 1,784 Internet users for their perceptions of wearable devices and contribute: relevant perceived risks for wearables, effects of data type and data recipient on perceived risk, users' self reported concerns, and an assessment of how wearable device capabilities compare to familiar technologies. We conclude with a discussion on future research directions for wearable devices.

Categories and Subject Descriptors

K.6.5. [**Management of Computing and Information Systems**]: Security and protection—*Unauthorized access*

Keywords

Privacy, Security, User Studies, Risk Perception, Ubiquitous Computing, Wearable Devices

1. INTRODUCTION

Wearable technologies, or "wearables," are a $700 million industry [2] of electronically enhanced clothing items and accessories that interweaves technological interaction with everyday life. A top 25 market research company estimates that 52% of technology consumers are aware of wearables and 33% said they were likely to buy one [5]. With 20% of the general population owning at least one wearable and 10%

using it daily [10], wearables are transforming ubiquitous computing into a part of every day life. Forbes has named 2014 the "Year of Wearable Technology [42]."

The constantly captured data from these devices has many benefits, ranging from a more natural, human-centered interface experience to a healthier, fitness-data inspired lifestyle. There will likely be many more applications in the future which take advantage of such data. It is clear why wearable devices are becoming even more popular, especially as they have more capabilities and benefits over traditional devices.

Along with these benefits, wearable devices bring new potential privacy and security risks which expose users' activities without their awareness or consent. Fitbit allowed sex to be tracked as exercise while fitness profiles were public by default [22], resulting in the inadvertent disclosure of sensitive information. Public discomfort toward facial recognition prevented the capability from being deployed to Google Glass [30]. Google Glass, the iconic wearable of its time, has since disappeared [21]. Most suspect that the reason for its disappearance was because it was the "straw the broke the back of the privacy camel's back" [13]. Some Glass wearers faced assault [38, 29, 12], while bystanders felt uncomfortable personal details, conversations, and photos possibly recorded.

We have seen similar privacy [26, 39, 41] and security issues [14, 16] related to data capture with respect to smartphones. Mobile platforms have tried to address this by communicating data capture to users as the data is captured. However, many users are habituated to these notifications, because they see them all the time, often for things that they don't care about [17].

With the additional capabilities of wearable devices and their increasing popularity, people have expressed interest in being notified before data capture [11], but human attention is a finite resource [6]. Therefore, user concerns should be investigated to to warn users only about situations they are likely to care about. Informed, select notifications make for a better user experience and prevents habituation to such notifications while avoiding scandalous breaches of privacy.

The goal of this paper is to gain a better sense of user concerns for wearables. To our knowledge, this is the first large-scale study to investigate user security and privacy concerns for wearable devices. We surveyed 1,784 Internet users for

their perceptions of wearable devices and contribute the following:

- Comparisons of users' perceptions of a range of privacy and security risks of wearables. We found that users care much more about the type of data than the recipient of the data.
- Insight into how users feel about various data recipients. We observed that users make less distinction between sharing data with friends, co-workers, and the general public, comparatively to sharing with an application's servers.
- A taxonomy and report of users' self-reported top concerns for wearable devices. Privacy is by far the top concern, along with security, health risks, financial risk, and social impact.
- Rankings of how data-collection capabilities of wearable devices compared to more familiar technologies. Most saw new capabilities as benign, but we suspect that this may be due to a lack of exposure to these newer technologies.

2. RELATED WORK
We discuss related work that has examined users' perceptions surrounding security and privacy risks.

2.1 Wearables Concerns
A small-scale interview of how bystanders feel about wearable devices [11] found that bystanders were predominantly split between having indifferent and negative reactions to the device. A variety of factors that make recording more or less acceptable, including what they are doing when the recording is being taken. Additionally, bystanders are expressed interest in being able to give permissions for the data being captured. We also investigate how the type and mechanism of data capture affects privacy concerns, but we examine the privacy concerns of wearables owners at a large-scale while their research examined the privacy concerns of wearables bystanders at a small-scale.

2.2 Ubiquitous Sensing Concerns
Many authors have emphasized that we are rapidly moving towards a world of ubiquitous sensing and data capture, with ensuing privacy challenges [1, 33, 7]. Many researchers have worked to study how privacy can be preserved in such a future. Examples of such efforts include frameworks to design for privacy [4, 8, 27] or evaluate privacy [40] in ubiquitous computing applications. Others have suggested various models for understanding privacy in ubiquitous computing systems [23, 25]. However, none of these works attempted to quantify or rank user concern over different privacy risks.

2.3 Smartphones Concerns
Many researchers have attempted to study end-user concerns about security or privacy issues associated with their smartphones [9, 34, 17].

2.4 User Perceptions and Behaviors
In this paper we focus on measuring people's perceptions of security and privacy risks. One limitation of user perceptions is that people don't always have enough information to make privacy-sensitive decisions; even if they do, they often trade off long-term privacy for short-term benefits [3]. Also, actual behavior may deviate from self-reported behaviors [24] and privacy preferences [43].

3. METHODOLOGY
To obtain a comprehensive list of possible risks that wearable devices might present in the future, we examined the sensors, capabilities, permissions, and applications of the most popular wearable devices on the market at the time of this study. At the time of this study, August 2014, the most popular wearable devices included the fitbit fitness tracker which performs continuously monitors heartbeat, steps taken, and sleep patterns [19, 44], the pebble smartwatch which can take pictures, send texts, show notifications from online, and push notifications to services [35, 47, 37], and google glass [48, 45]. These wearable devices, along with other comparable wearable devices on the market, were researched as inspiration for the survey questions.

Our survey contained two main sections. In one section, we presented participants with several scenarios—something undesirable that might happen with their wearable device—and asked them to rate their level of concern if each scenario were to happen. This was intended to elicit their perception of the severity and impact of the risk. In the other section, we asked participants to compare the risks and benefits of wearable technologies to better understood technologies, following the same methodology as a seminal study in risk perception by Fischhoff *et al.* [18]. Our survey design is based on two prior perception studies, as we describe next.

3.1 Motivation
3.1.1 Smartphone Risk Scenarios
Felt *et al.* previously studied the security concerns of smartphone users by conducting a large-scale online survey [15]. Their survey asked 3,115 smartphone users about 99 risk scenarios. Participants were asked how upset they would be if a certain action had occurred without permission. Participants rated each situation on a Likert scale ranging from "indifferent (1)" to "very upset (5)." Our methodology closely follows that study, but with different scenarios chosen to shed light on security and privacy risks of wearable devices.

3.1.2 Technology Risk Perception
Fischhoff *et al.* performed a seminal study of perceived risks with 30 widely used technologies [18]. In their study, participants were asked to separately rate the risks and benefits for those technologies. They were told to think about all people affected by the technology, and to think about long-term vs. short-term risks and benefits. Then, the participants rated these technologies with respect to each other on a numerical scale, being instructed to rate the least risky or least beneficial technology a 10 and scaling the ratings linearly (e.g., a technology with risk rating 20 is considered twice as risky compared to a technology with a risk rating of 10). We apply their methodology to evaluate perceived risks and benefits of several technologies related to wearable computing.

3.2 Survey Questions
In our survey, each participant answered 27 questions, across five different sections:

> Every once in a while, an app might do something on your *Cubetastic3000* without asking you first. Depending on what the app does, your feelings could range from indifference (you don't care) to being very upset.

5. How would you feel if an app on your Cubetastic3000 learned what medical conditions you have and shared that with your friends, without asking you first?

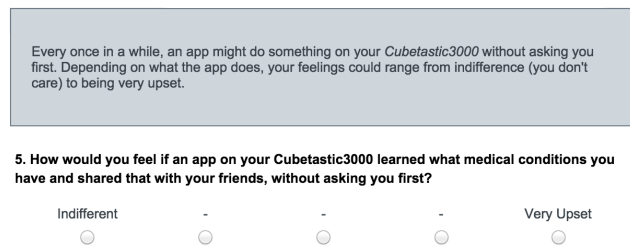| Indifferent	-	-	-	Very Upset
○	○	○	○	○

Figure 1: An example of a wearable scenario question participants saw while taking the survey.

- 2 comprehension questions
- 6 questions about wearable computing scenarios
- 2 questions about smartphone scenarios
- 2 risk/benefit questions
- 15 demographic questions

We randomized the order participants saw sections of the survey (with the exception of the comprehension and demographic questions, which were always first and last, respectively), as well as the order of questions in each section.

3.2.1 Comprehension Questions

Because participants might be biased to specific companies (e.g., visceral reactions to Google Glass based on popular media stories), we based our questions on a fictitious wearable. Thus, the beginning of the survey introduced participants to the "Cubetastic3000," which was the basis for all questions on wearables risks. We highlighted the capabilities of this device and described use cases. To ensure that participants had read and understood this device's capabilities, we ask them two multiple-choice comprehension questions.

3.2.2 Wearables Scenarios

We presented scenarios involving data capture using the Cubetastic3000 and asked them to rate how upset they would be if a particular data type (e.g., video, audio, gestures, etc.) were shared with a particular data recipient without asking first (see Figure 1). Responses were collected on a 5-point Likert scale (from "indifferent" to "very upset"), which was modeled after Felt et al.'s study of smartphone users' risk perceptions [15]. Our questions were of the format:

"How would you feel if an app on your Cubetastic3000 learned <data> and shared it with <recipient>, without asking you first?".

We created an initial pool of 288 questions by combining 72 data types (<data>) with 4 data recipients (<recipient>). The 4 possible data recipients were:

- Your work contacts
- Your friends
- The public
- The app's server (but didn't share it with anyone else)

The purpose of these questions was to determine the extent data types and data recipients play a role in upsetting participants when data is inappropriately shared. Additionally, we added 16 questions about other misbehaviors that

did not follow this format, lacking either <data> or a <recipient>, but we found relevant nonetheless. An example of one of these questions was, "*How would you feel if an app on your Cubetastic3000 turned your device off, without asking you first?*" There were a total of 304 questions in this set, from which we randomly 6 questions for each participant.

3.2.3 Smartphone Scenarios

We presented participants with a second set of scenarios to control for the type of device being used. These questions followed the format of the previous question set, but substituted "smartphone" for "Cubetastic3000." Rather than using the previous pool of 288 <data> and <recipient> combinations, we selected 5 of the scenarios that Felt et al. found least and most concerning to their participants [15]. We randomly presented each participant with 2 of these 5 questions:

1. *How would you feel if an app on your smartphone vibrated your phone without asking you first?*
2. *How would you feel if an app on your smartphone connected to a Bluetooth device (like a headset) without asking you first?*
3. *How would you feel if an app on your smartphone unmuted a phone call without asking you first?*
4. *How would you feel if an app on your smartphone took screenshots when you were using other apps, without asking you first?*
5. *How would you feel if an app on your smartphone sent premium (they cost money) calls or text messages, without asking you first?*

3.2.4 Risk and Benefit Assessment

In addition to investigating reactions to particular scenarios, we examined broad perceptions of new technologies and how those compared to perceptions of other understood technologies. We modeled this section after a seminal risk perception study by Fischhoff et al. [18], in which participants ranked technologies by their relative risk and benefit to society. We asked participants to perform this exercise for 4 technologies previously examined by Fischhoff et al.: handguns, motorcycles, lawnmowers, and electricity, which were chosen to span varying levels of risks and benefits.

Alongside the 4 studied technologies, we asked participants to evaluate one of 20 technologies relevant to wearables: internet, email, laptops, smartphones, smart watches, fitness trackers, Google Glass, Cubetastic3000, discrete camera, discrete microphone, facial recognition, facial detection, voice recognition, voice-based emotion detection, location tracking, speech-to-text, language detection, heart rate detection, age detection, and gender detection. We asked about familiar technologies such as the internet, general and specific wearable artifacts, and a range of new capabilities.

To parallel Fischhoff et al.'s risk perception study, we gave our participants a similar prompt to numerically express the perceived gross risk/gross benefit over a long period of time for all parties involved. We randomized whether they performed the ranking for risks or benefits first. The prompt is listed in Appendix A. The question format was as follows:

Fill in your <risk/benefit> numbers for the following:

Handguns: _____
Motorcycles: _____
Lawnmowers: _____
<Wearable Technology>: _____
Electricity: _____

3.2.5 Additional Questions

The exit portion of the survey firstly consisted of questions asking for age, gender, and education. Then, we asked participants if they owned a wearable device so we could control for prior exposure, and included an open-ended question on what would be the most likely risks associated with wearable devices. We end with the 10-question Internet Users' Information Privacy Concerns (IUIPC) index [28], so we could control for participants' general privacy attitudes.

3.3 Focus Group

We conducted a one-hour focus group to validate our design, gauge comprehension, and measure fatigue. The focus group began with participants taking the survey. Afterward, we asked participants to give feedback on the format and the content, noting any instructions or questions that were unclear. The focus group concluded with a discussion of possible benefits and risks of wearable devices, in order to brainstorm any additional scenarios to include. Finally, we compensated participants with $30 in cash. We recruited all of our focus group participants from Craigslist. Of the 13 participants, 54% were female, and ages ranged from 18 to 64 (mu = 36.1, sigma = 15.3). Education backgrounds ranged from high school to doctorate degrees, and professions included student, artist, marketer, and court psychologist.

3.4 Recruitment and Analysis Method

We recruited 2,250 participants August 7th-13th 2014 via Amazon's Mechanical Turk. We restricted participants to those over 18, living in the United States, and having a successful HIT completion rate of 95% or above. Based on incorrect responses to either of the two comprehension questions, we filtered out 366 (16% of 2,250) participants. We filtered out an additional 99 participants (4% of 2,250) due to incomplete responses, and three participants who were under 18, leaving us with a total sample size of 1,782. Of these, 57.9% were male (1,031), 41.0% were female (731), and 20 participants declined to state their genders. Ages ranged from 18 to 73, with a mean of 32.1 ($\sigma = 10.37$). Almost half of our participants had completed a college degree or more (49.2% of 1,782), which includes the 219 (12.3% of 1,782) who reported graduate degrees. While our sample was younger and more educated than the U.S. population as a whole, we believe it is still consistent with the U.S. Internet-using population.

In performing our analysis in the next section, we chose to focus on the very upset rate (VUR) of each scenario. The VUR is defined as the percentage of participants who reported a '5' on the Likert scales. We use the VURs rather than the average of all Likert scores for the same reasons as Felt *et al.*: the VUR does not presume that the ratings, ranging from "indifferent" to "very upset," are linearly spaced. Additionally, most were be upset, at least a little, in all scenarios

when a device takes action without permission (rating distribution: "1"= 455, "2" = 523, "3" = 902, "4"' = 1,746, "5" = 6,654). Thus, the main distinguishing factor of a participant reacting to a given scenario is whether they were maximally upset or not, rather than how upset they were.

We followed Fischhoff *et al.*'s methodology and did not normalize the numerical responses. Rather, we report medians and quartiles, which are not impacted by outliers. For the open-ended question at the end (i.e., additional privacy concerns), two researchers independently coded 1,784 responses, with an initial agreement rate of 89.7%. The researchers discussed and resolved any disagreements so that the final codings reflect unanimous agreement.

4. RESULTS

We present our survey results and provide analyses of the data. We first discuss participants' responses to the various data-sharing scenarios, and how data type, data recipient, and device contributed to how threatening a situation was perceived. Next, we discuss participants' risk/benefit assessment of various new technologies relative to well-established technologies. We conclude the section with participants' self-reported concerns about the biggest risks in owning wearable devices.

4.1 Concern Factors

Many factors impact participants' concern levels for each scenario: the data recipient, the data type, and whether or not the scenario occurred on a wearable or a smartphone. We analyze each factor individually, as well as present a statistical model of participants' concerns as a function of all of factors, including demographic traits.

4.1.1 Data Type

Based on our data, we observed that the largest effect stemmed from the data being shared. We present various statistical models in section 4.1.6 to support this conclusion. The 10 top and bottom concerning data can be seen in Table 1, and the full list of data can be seen in Appendix [?].

Participants were most concerned about photos and videos, especially if they contained embarrassing content, nudity, or financial information. As seen in Table 1, photos and videos accounted for 5 of the top 10 concerns, and are almost unanimously considered to be concerning. Information that could be used to impersonate someone (e.g., usernames/passwords for websites) or invade privacy (photos of someone at home) were also among the most concerning data types.

Participants were least concerned about data that could be observed through observations of public behavior, such as demographic information (e.g., age, gender, language spoken) and information available to advertisers (e.g. TV shows watched, music on device). As seen in Table 1, participants had spread distributions in perceptions regarding such information. These may have appeared as unconcerning because of unfamiliarity in what applications would use this data for, or because there does not seem to be any immediate financial, social, or physical consequences from having this information shared.

For the complete ranked list of data considered in this study, see Appendix [?]. Although certain data is considered unanimously upsetting to have shared, it is interesting to note that no data was considered unanimously non-upsetting to have shared, nor any data which evoked strong disagreement on how upsetting it was. Generally, the rank of the data being shared is negatively correlated with the standard deviation of the answers.

The data types examined span several existing and future use cases. We do not believe they are comprehensive in coverage of how wearable devices might capture data nor for all possible future use cases of the data. Therefore, we coded each data type in two ways: firstly in terms in the data and how it was captured (e.g., video, audio, text, etc.), and secondly in terms of the type of risks presented to the user. This examines why participants viewed certain data collected as more/less risky than others, letting us generalize our results. Two researchers agreed on a codebook and independently coded each of the 72 data types[1].

The data fell into the following six possible categories:

1. Photo
2. Video
3. Audio
4. Behavioral Information
5. Biometric Information
6. Demographic Information

The associated risks fell into the following five categories:

1. **Financial:** The risk involves the loss of money or property.
2. **Image:** The risk involves loss of control over one's self-image (e.g., publicizing something embarrassing).
3. **Medical:** The risk involves disclosure of medical information.
4. **Physical:** The risk involves physical harm to the user.
5. **Relationships:** The risk involves damage to the user's inter-personal relationships.

After independently coding all data types with respect to these categories, the researchers met to resolve any disagreements, such that the resulting codings reflected unanimity. Prior to this resolution, there was 83% agreement. Cohen's κ is 0.81 for the data capture and 0.75 for the associated risks, both indicating "excellent" agreement [20].

4.1.2 Data Recipient

Across all scenarios, 42.3% of participants stated that they would be "very upset" if their data was shared with only the app's servers, whereas the VURs for friends (69.5%), work contacts (75.2%), and the public (72.4%) were much higher. A chi-square test indicated that these differences were statistically significant (Table **??**). However, these effect sizes were small: the largest effect was between work contacts and an app's server ($\phi = 0.11$); while the VUR for sharing with work contacts was significantly higher than sharing with friends, the effect size was negligible ($\phi = 0.004$).

---

[1] We excluded the data types that did not feature a data recipient.

Rank	Recipient	VUR	sigma	Distribution
1	Work Contacts	75.16%	0.94	
2	Public	72.41%	0.98	
3	Friends	69.47%	1.02	
4	App's Server	42.28%	1.15	

Table 2: The overall upset rate for all recipients.

We note that this chi-square test violates the assumption of independent observations, since participants responded to multiple scenarios. But based on the randomization of treatments and large sample size, we do not believe that this significantly impacted our results. Similarly, we are unaware of a more appropriate test, given our data format; Cochran's Q requires binary outcomes (i.e., participants would have had to answer only one question for each data recipient, preventing us from adequately controlling for data type) and a repeated measures ANOVA requires normality (our data was not normally distributed). Nonetheless, we repeated our analysis using only one randomly-selected data point per participant and found that our selected test was robust to this violation. Participants were significantly more concerned about having their data seen by humans (*vis-à-vis* app servers), though differences between specific human groups (between the public, friends, and work contacts) were not significant.

We do not claim that there is no distinction between the friends, public, and work contact recipients, because people have been shown to behave differently, especially in the social arena (cite studies here). People are more comfortable sharing certain data types with certain recipients (refer to the table of all questions with ranks by each recipient here). There are other data types which are universally concerning, and universally unconcerning, and the magnitude of the sentiment varies by recipient. As you can see in Table (J still needs to do this table), the VUR rate for each recipient is negatively correlated with the standard deviation of the answers. Additionally, we see that there is greater distinction between sharing people and the app server–people are comfortable sharing data they would feel uncomfortable sharing with at human recipient, and the magnitude of concerns is might more significant compared to the other recipients.

TODO: table of top and bottom 10 data types by recipient; refer to appendix; shoutout to when you were lying nervous or stressed work

4.1.3 Data Type and Data Recipient

We compared the 10 most concerning scenarios when sharing with an app servers versus with a humans. We observed that there was a substantial overlap between these groups, in that 6 of the most concerning scenarios were the same:

1. Bank account information
2. A video of you unclothed
3. Social security number
4. Video of you entering your PIN
5. An incriminating/embarrassing photo of you

Rank	Data	VUR	σ	Distribution
1	video of you unclothed	95.97%	0.31	
2	bank account information	95.91%	0.35	
3	social security number	94.84%	0.26	
4	video entering in a PIN at an ATM	92.67%	0.47	
5	photo of you unclothed	92.59%	0.46	
6	photo of you that is very embarrassing	91.39%	0.55	
7	username and password for websites	89.55%	0.62	
8	credit card information	88.98%	0.56	
9	video of you that is very embarrassing	88.41%	0.53	
10	photo of you at home	87.50%	0.60	
$\vdots$				
64	eye patterns (for eye tracking)	40.51%	1.27	
65	exercise patterns	38.66%	1.26	
66	when you are happy or having fun	34.75%	1.27	
67	television shows watched	30.20%	1.40	
68	when you are busy or interruptible	29.50%	1.26	
69	music on device	28.06%	1.43	
70	your heart rate	27.50%	1.40	
71	age	24.29%	1.43	
72	language spoken	15.86%	1.49	
73	gender	15.00%	1.45	

Table 1: The 10 most and least upsetting data types, across all recipients. For the complete list of all data types across all recipients, see Appendix [?].

6. A photo of you unclothed

While the concerning data types do not appreciably change based on the data recipient—even the non-overlapping scenarios all dealt with confidential data (e.g., passcode, credit card information, etc.)—only the level of concern changed. For instance, the 10th most concerning scenario for the non-human audience had a VUR of 66.67%, whereas the 10th most concerning scenario for a human audience has a VUR of 93.88%. This suggests that concern for different data types does not appear to vary relative to other data types based on recipient, but instead the recipient determines the overall magnitude of the concern.

4.1.4 Device
Participants had unique VURs for scenarios only differing in device. Our participants had a 58.79% VUR when asked about wearables and 46.64% VUR when asked about smartphones. The VURs for both devices for all 5 questions are in table 3. However, the effect the device has on the VUR is not considered to be statistically significant (see Table 4). Additionally, there is no statistically significant difference

between how people reacted in a given situation; although, participants were statistically significantly upset in Q2. The aforementioned results are only with respect to between subjects analysis, where answers are from participants who received either only the wearables or smartphone version of the 5 questions. Too few instances of participants answering both versions of questions occured (34 in total for all 5 questions) to perform a sound within-subjects analysis.

talk about the variance of the devices in general–was there more spread for the answers wrt smartphones, cubetastic, or were they kind of all the same?

4.1.5 Demographic Factors
Participants' responses were correlated with demographic factors. We observed that the biggest predictor of participants' decisions to rate a scenario as very upsetting was their self-reported level of general privacy concerns, as determined by the IUIPC scale [28]: a Spearman correlation yielded a statistically significant effect between average IUIPC scores with the VUR ($\rho = 0.446$, $p < 0.0005$). Similarly, we observed that age was a significant predictor of VUR ($\rho = $

Question	Wearable VUR	Smartphone VUR
All	58.79%	46.64%
Q1	14.81%	6.13%
Q2	44.11%	19.85%
Q3	87.09%	58.44%
Q4	52.77%	55.74%
Q5	86.49%	91.82%

Table 3: VURs for the questions described in Section 3.2.3, contrasting smartphones with the Cubetastic3000.

Question	χ^2	p-value	n	ϕ
All	2.202	<0.1378	3,588	0.001
Q1	2.500	<0.1139	714	0.004
Q2	17.333	<0.0001	708	0.024
Q3	0.020	<0.8886	699	0.000
Q4	1.413	<0.2345	730	0.002
Q5	1.604	<0.2054	709	0.002

Table 4: Chi-square test results comparing participants' VURs between the smartphone and Cubetastic3000 questions.

0.121, $p < 0.0005$). We suspect that the effect of age is due to the significant correlation between age and IUIPC scores ($\rho = 0.188$, $p < 0.0005$); others have observed that older individuals tend to be more protective of their privacy [46].

While we initially observed an effect on VURs based on whether or not participants claimed to already own wearable devices (57.0% vs. 60.8%, respectively; Mann-Whitney $U = 202,896$, $p < 0.032$), this difference did not remain significant upon correcting for multiple testing (Bonferroni corrected $\alpha = 0.01$), nor did the effect of gender. Finally, we observed no correlation between education level and VUR.

4.1.6 Regression Models
In order to examine the relative effect of each factor on participants' VURs, we constructed several statistical models to predict whether a participant would be "very upset" with a given scenario based on the data type, device, data recipient, and their demographic factors (i.e., age, education, gender, and privacy attitudes). We performed binary logistic regressions using generalized estimating equations, which account for our repeated measures experimental design (i.e., each participant contributed multiple data points).

We created several models using our three dependent variables as factors: device (smartphone vs. wearable), data recipient, and data type. We also used our collected demographic factors as covariates: age, gender, education, wearable device ownership (yes/no), and mean IUIPC score. For each model, we performed Wald's test to examine the model effects attributable to each of these eight parameters and observed that the only covariate that had an observable effect on our models was participants' IUIPC scores. Thus, we opted to remove the other covariates from our analysis. Similarly, we observed no statistically significant interaction effects between any of these four variables, which is why we

Parameters	χ^2	df	QIC
(Intercept)	254.5	1	18,699.82
(Intercept)	79.2	1	18,347.6
Device	391.0	1	
(Intercept)	232.1	1	17.897.0
IUIPC (covariate)	368.5	1	
(Intercept)	298.2	1	17,606.5
Data Recipient	900.0	4	
(Intercept)	370.9	1	14,970.7
Data Type	1,898.4	76	
(Intercept)	79.2	1	18,349.9
Device	391.0	1	
(Intercept)	29.2	1	14,114.8
Device	8.1	1	
Data Recipient	579.3	3	
Data Type	1,765.4	76	
(Intercept)	298.0	1	12,931.6
Device	10.4	1	
Data Recipient	626.2	3	
Data Type	1,997.5	76	
IUIPC (covariate)	374.8	1	

Table 5: Goodness-of-fit metrics for various binary logistic models of our data using general estimating equations to account for repeated measures. The columns represent the Wald test statistic for each parameter and the overall Quasi-Akaike Information Criterion (QIC) for each model. Each parameter listed was statistically significant at $p < 0.0005$.

did not include them in our models.

Table 5 shows the various models that we examined and the Quasi-Akaike Information Criterion (QIC), which is a goodness-of-fit metric for model selection (lower relative values indicate better fit). As can be seen, while the remaining four predictors all contributed to the predictive power of our model, the data type was the strongest predictor. Conversely, despite being significant, the device was the weakest predictor (i.e., whether participants were answering questions about a smartphone or a wearable device). This suggests that participants' biases towards specific wearable devices or wearable devices in general had a minimal effect on their results, and that they were primarily focused on the data type captured and how it will be shared.

4.2 Risk and Benefit Rankings
We asked participants to rate new capabilities related to wearable technologies (e.g., facial recognition) in terms of their risks and benefits. We also asked them to do this for technologies with which they were likely to be more familiar (e.g., smartphones and laptops) in addition to two examples of specific wearable devices, Google Glass and the fictitious Cubetastic3000. To calibrate our results, we also asked about four well-established technologies studied by Fischhoff *et al.* [18]. We found that participants generally rated familiar technologies and those related to wearables as being low-risk. Figure 2 depicts participants' median ratings. We found that the calibration technologies were all rated as the most risky. At the same time, with the exception of electricity, the calibration technologies were seen as lower benefit than the others.
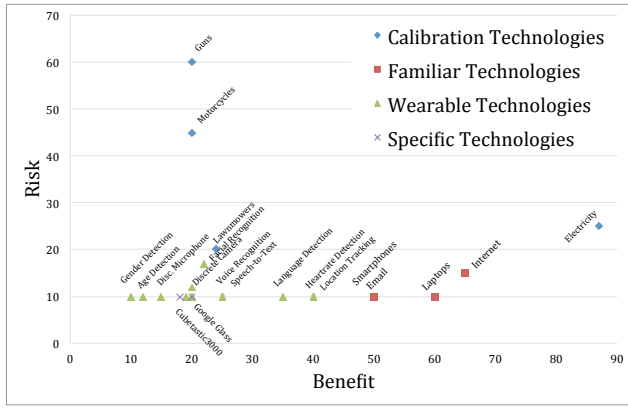
Figure 2: Participants' median risk-benefit ratings of technologies examined by Fischhoff *et al.* [18], which we used for calibration, alongside familiar technologies (e.g., laptops, the Internet, etc.), wearable technologies, as well as two specific wearable devices (Google Glass and the Cubetastic3000).

As a group, participants rated the familiar technologies as the most beneficial. We believe this is the result of exposure people have to these technologies—most people use these technologies daily. Of the wearable technologies, the most risky were ones perceived to be privacy-invasive; the most risky technologies were facial recognition, the Internet, and discrete cameras, whereas the remainder of the technologies were seen as having minimal—albeit equivalent—risk levels (i.e., a median of "10"). People are becoming increasingly aware of such privacy risks and are comparing these privacy invasion to real physical risks–for instance, the capacity for facial detection on a wearable device is perceived to be almost as risky as interacting with a lawnmower.

talk about the distributions and variance of the particular technologies, referring to the table in the appendix

people may have evaluated the risks only thinking of physical risk, not privacy risk. This might have happened because among the 5 presented options, the wearable-related one is the odd one out; all other options involve some physical risk scenario. These other options, by being the most prominent (4 versus 1), frame the risk perception in the user's mind as meaning "physical risk", and users may consequently ignore or downplay privacy risks. It would have been better to ask 4+4 questions (or 2+2 to keep the survey short) rather than the current 4+1. AND "with the exception of electricity, the calibration technologies were seen as lower benefit than the others". This is true for some, but not all of the others. Specifically, Google glass and Cubetastic3000 were about equally beneficial, and gender and age recognition were less beneficial. AND The differences in risk that *are* found between the different wearable-related are not tested for statistical significance, but given their minimal spread compared to the calibration options, the differences are negligible.

These perceptions of the most risky or beneficial technologies may not be reflective of actual risks or benefits. However, they do reflect the general public's exposure to these

Concern	Responses	Frequency
Privacy	452	25.32%
Being Unaware	275	15.40%
Health Risk	191	10.70%
Safety	185	10.42%
Social Impact	157	8.80%
Financial Cost	151	8.46%
Security	144	8.07%
Accidental Sharing	69	3.87%
Miscellaneous	57	3.19%
None	51	2.86%
Social Stigma	39	2.18%
False Information	33	1.85%
Don't know	31	1.74%
Aesthetics	19	1.06%
Don't care	11	0.62%

Table 6: A table listing the self-reported most common risks associated with owning a wearable device.

technologies and show that people perceive specific risks and benefits. We suspect that the similarity in assessments between the various wearable technologies are because most people are not consciously aware of the possibilities of these technologies or how they could be used. We suspect that performing this experiment longitudinally may yield more interesting results, as these technologies become more and more pervasive (and therefore more familiar to participants).

4.3 Self-Reported Concerns for Wearables

We also wanted to capture the participants' general reactions to wearable devices as a whole. To do this, we asked the participants the following open-ended question:

What do you think are the most likely risks associated with wearable devices?

This question was asked along with demographics questions (but before any IUIPC questions to avoid biasing). The participants were presented with a blank box to write in, with no character limit to their open-ended responses.

Without a doubt, the most common self-reported concern of wearables for the average user is the *possible loss of privacy* (see Table 6). Other significant concerns included being unaware of what the device is collecting, doing, or which information it is using (Being Unaware), long-term health effects caused from wearing the device such as cancer from emf waves (Health), safety hazards from wearing the device such as distractions which cause car accidents (Safety), resulting changes in social behaviors, such as dependencies on devices or spending less time with loved ones (Social Impact), the high financial cost of buying, replacing, or caring for the device (Financial), and information compromise (Security).

5. DISCUSSION

Here, we discuss complementary future research directions in fields of privacy, ubiquitous computing, and user studies,

along with specific limitations of this survey.

5.1 Interpreting the Survey Results
One of the main limitations of this work is that our participants might not have interest in or knowledge of wearables and their capabilities. 83% of our participants reported that they do not own a wearable device. These participants may have underestimated or overestimated the risk perceived in various scenarios. People may be overreacting to recent events for scenarios [2]. People may be underestimating the risk of sharing certain data due to unawareness of what can be inferred from the data, or not have an idea of how to rate a new technology with respect to familiar ones. Biometrics were generally not a concern for our participants, although there are many security and privacy implications [36]. Any economics or behavioral papers to support our claims and elaborate on this would be great here. Maybe ones on perception, estimation, etc.

We believed that getting a representative survey base was a useful endeavor. We could have easily recruited only wearables owners or people specifically interested in wearables. However, that will also have its own biases and limitations, since this does not reflect the general population. At the time of this writing, about 85% of the general population do not own wearable devices [31, 10], so our study reflective of the status quo. We expect user perceptions to change as rapidly as wearable technologies and the rate of adoption change.

Privacy concerns asked out of context differ from how users may react to these same concerns in real life [32, 24]. This is an unavoidable, yet important consideration of any study of this nature. This privacy paradox means that our findings may not be exactly representative of how upset users may be in real life, but do reflect their perceptions of wearable devices and various associated scenarios.

5.2 Future Research Directions
Further work can be done to expand various aspects of this study. Investigating more fine-grained data types (e.g. investigating if various types of location data, versus just location data in general) would be a useful endeavor to gain further insight into user perception. Adding more recipients, like "advertisers" or "acquaintances" may lead to more contrasting results.

While privacy and security concerns were expected, consider the following self-reported user concerns as inspiration for future research: addressing the high financial costs of wearables, communicating the reality of health concerns from constant use, creating distraction-free interfaces to prevent safety issues, minimizing negative social impacts of wearable device use, and improving device aesthetics.

Wearables are still in infancy. Perceptions of situations and capabilities will change rapidly with advancements and in-

---
[2]During our study, there were many stories covering injuries from exploding batteries (http://www.bloomberg.com/news/articles/2014-08-11/exploding-lithium-batteries-riskier-to-planes-research), which were explicitly and repeated mentioned when self-reporting concerns.

creased exposure. However, videos and textual information are considered to be significantly sensitive by our participants, along with past participants of smartphone user perception studies. Various systems which detect and take actions for sensitive objects in photos and videos will be critical as wearables and other devices become more ubiquitous.

6. CONCLUSION
We surveyed 2,250 internet users to determine what contributes to a violation of privacy or security, which technologies are risky, and what users think are the biggest risk for operating wearable devices. We examine how upset participants would in 304 scenarios, assessed the risk and benefit for 20 new technologies, and gave open-ended responses to express their concerns. We provide insight into how much and why data, recipient, and device contribute to users' perception of a situation, calibrate answers with existing smartphone literature, and provide a regression model. An assessment of a range of new technologies shows that users perceive new technologies to be low-risk and low-benefit, but we suspect this is due to limited exposure that an average person has with wearables techology. We also state what users perceived as the most significant concerns with respect to wearable devices. We conclude by discussing future research directions in the wearables and user study space.

7. REFERENCES
[1] G. D. Abowd and E. D. Mynatt. Charting past, present, and future research in ubiquitous computing. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 7(1):29–58, 2000.

[2] G. Abramovich. 15 mind-blowing stats about wearable technology. http://www.cmo.com/articles/2014/6/16/Mind_Blowing_Stats_Wearable_Tech.html. Accessed: 2014-12-19.

[3] A. Acquisti and J. Grossklags. Privacy and rationality in individual decision making. *IEEE Security & Privacy*, 2:24–30, 2005.

[4] V. Bellotti and A. Sellen. Design for privacy in ubiquitous computing environments. In *Proceedings of the Third European Conference on Computer-Supported Cooperative Work 13–17 September 1993, Milan, Italy ECSCW'93*, pages 77–92. Springer, 1993.

[5] S. Bogaty. Wearable tech device awareness surpasses 50 percent among us consumers, according to npd. https://www.npd.com/wps/portal/npd/us/news/press-releases/wearable-tech-device-awareness-surpasses-50-percent-among-us-consumers-according-to-npd/. Accessed: 2014-12-26.

[6] R. Böhme and J. Grossklags. The security cost of cheap user interaction. In *Proceedings of the 2011 workshop on New security paradigms workshop*, pages 67–82. ACM, 2011.

[7] J. Camp and Y. Chien. The internet as public space: concepts, issues, and implications in public policy. *ACM SIGCAS Computers and Society*, 30(3):13–19, 2000.

[8] L. J. Camp. Designing for trust. In *Trust, Reputation, and Security: Theories and Practice*, pages 15–29. Springer, 2003.

[9] E. Chin, A. P. Felt, V. Sekar, and D. Wagner.

Measuring user confidence in smartphone security and privacy. In *Proceedings of the Eighth Symposium on Usable Privacy and Security*, page 1. ACM, 2012.

[10] J. Comstock. Pwc: 1 in 5 americans owns a wearable, 1 in 10 wears them daily. `http://mobihealthnews.com/37543/pwc-1-in-5-americans-owns-a-wearable-1-in-10-wears-them-daily/`. Accessed: 2014-12-19.

[11] T. Denning, Z. Dehlawi, and T. Kohno. In situ with bystanders of augmented reality glasses: Perspectives on recording and privacy-mediating technologies. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages 2377–2386. ACM, 2014.

[12] C. Doug Gross. Google glass targeted as symbol by anti-tech crowd - cnn.com, 2014.

[13] J. Dvorak. Rest in peace, google glass: 2012-2014, 2014.

[14] W. Enck, D. Octeau, P. McDaniel, and S. Chaudhuri. A study of android application security. In *USENIX security symposium*, volume 2, page 2, 2011.

[15] A. P. Felt, S. Egelman, and D. Wagner. I've got 99 problems, but vibration ain't one: a survey of smartphone users' concerns. In *Proceedings of the second ACM workshop on Security and privacy in smartphones and mobile devices*, pages 33–44. ACM, 2012.

[16] A. P. Felt, M. Finifter, E. Chin, S. Hanna, and D. Wagner. A survey of mobile malware in the wild. In *Proceedings of the 1st ACM workshop on Security and privacy in smartphones and mobile devices*, pages 3–14. ACM, 2011.

[17] A. P. Felt, E. Ha, S. Egelman, A. Haney, E. Chin, and D. Wagner. Android permissions: User attention, comprehension, and behavior. In *Proceedings of the Eighth Symposium on Usable Privacy and Security*, page 3. ACM, 2012.

[18] B. Fischhoff, P. Slovic, S. Lichtenstein, S. Read, and B. Combs. How safe is safe enough? a psychometric study of attitudes towards technological risks and benefits. *Policy sciences*, 9(2):127–152, 1978.

[19] Fitbit.com. Fitbit official site for activity trackers & more, 2014.

[20] J. L. Fleiss, B. Levin, and M. C. Paik. *Statistical Methods for Rates and Proportions*. John Wiley & Sons, Inc., 3rd edition edition, 2003.

[21] Google.com. Google glass, 2015.

[22] K. Hill. Fitbit moves quickly after users' sex stats exposed. `http://www.forbes.com/sites/kashmirhill/2011/07/05/fitbit-moves-quickly-after-users-sex-stats-exposed/`. Accessed: 2014-12-26.

[23] J. I. Hong, J. D. Ng, S. Lederer, and J. A. Landay. Privacy risk models for designing privacy-sensitive ubiquitous computing systems. In *Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques*, pages 91–100. ACM, 2004.

[24] 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):203–227, 2005.

[25] X. Jiang, J. I. Hong, and J. A. Landay. Approximate information flows: Socially-based modeling of privacy in ubiquitous computing. In *Ubicomp 2002: ubiquitous computing*, pages 176–193. Springer, 2002.

[26] P. G. Kelley, L. F. Cranor, and N. Sadeh. Privacy as part of the app decision-making process. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3393–3402. ACM, 2013.

[27] M. Langheinrich. Privacy by Design—Principles of Privacy-Aware Ubiquitous Systems. In *Ubicomp 2001: Ubiquitous Computing*, pages 273–291. Springer, 2001.

[28] N. K. Malhotra, S. S. Kim, and J. Agarwal. Internet users' information privacy concerns (iuipc): the construct, the scale, and a causal model. *Information Systems Research*, 15(4):336–355, 2004.

[29] Mashable. Woman robbed, assaulted for wearing google glass in a bar, 2014.

[30] E. Morphy. Google glass drops facial recognition (for now). `http://www.forbes.com/sites/erikamorphy/2013/06/02/google-glass-drops-facial-recognition-for-now/`. Accessed: 2014-12-26.

[31] N. News. Are consumers really interested in wearing tech on their sleeves? `http://www.nielsen.com/us/en/insights/news/2014/tech-styles-are-consumers-really-interested-in-wearing-tech-on-their-sleeves.html`. Accessed: 2014-12-19.

[32] P. A. Norberg, D. R. Horne, and D. A. Horne. The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of Consumer Affairs*, 41(1):100–126, 2007.

[33] L. Palen and P. Dourish. Unpacking Privacy for a Networked World. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 129–136. ACM, 2003.

[34] L. Palen, M. Salzman, and E. Youngs. Going wireless: Behavior & practice of new mobile phone users. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, pages 201–210. ACM, 2000.

[35] Pebble and P. S. Smartwatch. Pebble smartwatch, 2014.

[36] S. Prabhakar, S. Pankanti, and A. K. Jain. Biometric recognition: Security and privacy concerns. *IEEE Security & Privacy*, 1(2):33–42, 2003.

[37] Readwrite.com, 2014.

[38] K. Russell. I was assaulted for wearing google glass in the wrong part of san francisco, 2014.

[39] 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 and Ubiquitous Computing*, 13(6):401–412, 2009.

[40] J. Scholtz and S. Consolvo. Toward a framework for evaluating ubiquitous computing applications. *Pervasive Computing, IEEE*, 3(2):82–88, 2004.

[41] I. Shklovski, S. D. Mainwaring, H. H. Skúladóttir, and H. Borgthorsson. Leakiness and creepiness in app space: Perceptions of privacy and mobile app use. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages

2347–2356. ACM, 2014.

[42] E. Spence. 2014 will be the year of wearable technology. http://www.forbes.com/sites/ewanspence/2013/11/02/2014-will-be-the-year-of-wearable-technology/. Accessed: 2014-12-19.

[43] S. Spiekermann, J. Grossklags, and B. Berendt. E-privacy in 2nd generation e-commerce: Privacy preferences versus actual behavior. In *Proceedings of the 3rd ACM conference on Electronic Commerce*, pages 38–47. ACM, 2001.

[44] TIME.com. 26 fitness trackers ranked from worst to first, 2014.

[45] J. Turi. The top 9 wearables you can buy right now, 2014.

[46] H. R. Varian, F. Wallenberg, and G. Woroch. The demographics of the do-not-call list. *IEEE Security & Privacy*, 3(1):34–39, 2005.

[47] T. Verge. The best wearables of ces 2014, 2014.

[48] Wikipedia. Google glass, 2015.

APPENDIX
A. FISCHHOFF PROMPTS
We would like to ask you to rate the <risks/benefits> associated with each of the following technologies.

Risks: *Consider all types of risks: the risk of physical harm or death, the risk to others or bystanders, the financial cost of the technology, any distress caused by the technology, what the consequences would be if the technology was misused, any impact on the public, work, or personal life, and other considerations. (e.g. for electricity, consider the risk of electrocution, the pollution caused by coal, the risk that miners need to take to mine the coal, the cost to build the infrastructure to deliver electricity, etc.) Give a global estimate over a long period of time (say, a year) of both intangible and tangible risks.*

Do not consider the costs or risks associated with these items. It is true, for example, that sometimes swimmers can drown. But evaluating such risks is not your present job. Your job is to assess the gross benefits, not the net benefits which remain after the costs and risks are subtracted out.

Please rate the following technologies below with a number. We know that this might be a bit hard to do, but please try to be as accurate as possible, adjusting the numbers until they feel they are right for you. Start with the least risky technology at 10 and assign higher numbers for the more risky technologies. (For instance, a technology rated 14 is half as risky as a technology rated 28.)

Benefits: *Consider all types of benefits: how many jobs are created, how much money is generated directly or indirectly, how much enjoyment is brought to people, how much a contribution is made to the people's health and welfare, what this technology promotes, and so on. (e.g. for swimming, consider the manufacture and sale of swimsuits, the time spent exercising, the social interactions during swimming, and the sport created around the activity.) Give a global estimate over a long period of time (say, a year) of both intangible and tangible benefits.*

Do not consider the costs or benefits associated with these items. It is true, for example, that electricity also creates a market for home appliances. But evaluating such benefits is not your present job. Your job is to assess the gross risks, not the net risks which remain after the costs and risks are subtracted out.

Please rate the following technologies below with a number. We know that this might be a bit hard to do, but please try to be as accurate as possible, adjusting the numbers until they feel they are right for you. Start with the least beneficial technology at 10 and assign higher numbers for the more beneficial technologies. (For instance, a technology rated 34 is twice as beneficial as a technology rated 17.)

B. CODING LABEL DEFINITIONS
Privacy: "privacy," revealing personal information, spying.
Security: "security," compromise, malware, hacking.
GPS tracking: "location," "GPS," being monitored.

Unaware use: using data without permission or in a different way than understood by user.
Unaware collection: collecting data without permission.
Unaware access: disclosure of data without permission.
False information: inaccurate or maliciously false data.

Health Risk: radiation, cancer, or long-term effects.
Safety: distractions causing car crashes or injuries, mugging or violence because of the device, injuries from device malfunctions (battery burns).
Discomfort: eye strain, headache, obscured vision, irritation.
Financial cost: getting ripped off by buying the device or device accessories, having to buy another device when broken or stolen, financial compromise caused by device.
Theft: the device getting stolen.

Social Impact: dependency, distance from friends and family, changes in decision making, social changes, etc.
Social Stigma: judgment, hate, or bystander discomfort.
Aesthetics: fashion, the device being ugly, mentions of not looking cool/dorky.

Miscellaneous: odd comments, uncommon concerns.
None: "None," no threat, perceiving no big concerns
Don't know: "do not know," hinting at confusion
Don't care: " do not care," nonchalant answers

Technology	Q1	Median	Q3	Distribution
Location Tracking	10.0	10.0	20.0	
Speech To Text	10.0	10.0	10.0	
Discreet Microphone	10.0	10.0	20.0	
Smartwatches	10.0	10.0	10.0	
Language Detection	10.0	10.0	10.0	
Laptops	10.0	10.0	15.0	
Smartphones	10.0	10.0	20.0	
Google Glass	10.0	10.0	20.0	
Cubetastic	10.0	10.0	30.0	
Gender Detection	10.0	10.0	13.5	
Voice Recognition	10.0	10.0	15.0	
Voice Based Emotion Detection	10.0	10.0	15.0	
Fitness Trackers	10.0	10.0	10.0	
Age Detection	10.0	10.0	15.0	
Facial Detection	10.0	10.0	25.0	
Email	10.0	10.0	18.0	
Heart Rate Detection	10.0	10.0	10.0	
Discreet Video Camera	12.0	10.0	30.0	
Internet	15.0	10.0	31.0	
Facial Recognition	17.0	10.0	30.0	
Lawnmower	20.0	12.0	30.0	
Electricity	25.0	15.0	40.0	
Motorcycle	45.0	27.0	70.0	
Handgun	60.0	40.0	100.0	

Table 7: The 10 most and least upsetting data types, across all recipients.

Technology	Q1	Median	Q3	Distribution
Gender Detection	10.0	10.0	15.0	
Age Detection	12.0	10.0	22.0	
Discreet Microphone	15.0	10.0	20.0	
Cubetastic	15.0	10.0	30.0	
Fitness Trackers	18.5	10.0	30.0	
Voice Based Emotion Detection	20.0	10.0	30.0	
Facial Detection	20.0	10.0	34.0	
Discreet Video Camera	20.0	15.0	30.0	
Google Glass	20.0	12.0	40.0	
Smartwatches	20.0	10.0	35.0	
Motorcycle	20.0	12.0	40.0	
Handgun	20.0	10.0	30.0	
Facial Recognition	22.0	12.5	42.5	
Lawnmower	24.0	15.0	40.0	
Speech To Text	25.0	15.0	40.0	
Voice Recognition	25.0	15.0	40.0	
Language Detection	35.0	15.0	60.0	
Heart Rate Detection	40.0	26.0	65.0	
Location Tracking	40.0	20.0	70.0	
Email	50.0	29.0	77.5	
Smartphones	50.0	30.0	75.0	
Laptops	60.0	40.0	80.0	
Internet	65.0	45.0	100.0	
Electricity	88.0	50.0	100.0	

Table 8: The 10 most and least upsetting data types, across all recipients.