UNIL | Université de Lausanne
HEC Lausanne

Qualitative Analysis of Interviews with Fitness Trackers Users: Understanding How They Perceive Privacy And Utility

Author:

MAZURYK Laura

Professor:

CHERUBINI Mauro

Expert:

HUGUENIN Kevin

Lausanne, 20th June 2021

Master in Information Systems (MScIS)

Faculty of Business and Economics

University of Lausanne

Abstract

Wearable technology like fitness trackers collect, store, and analyze fitness data such as heart rate, sleep pattern, step count, calories and so on. Thanks to their ability to do that seamlessly and continuously, they provide great benefits to users who want to achieve better lifestyle. Nonetheless, they also raise privacy concerns. They collect great amount of personal data from which other personal information could be inferred.

In order to understand fitness tracker users perception of utility and privacy in the context of fitness tracker, and to identify the opportunities to improve fitness tracker users' privacy without damaging their utility, we conducted interviews. They took place after a 4-month experiment in which participants were given a fitness tracker and filled a questionnaire. We ran 19 semi-structured interviews in order to get additional insights about participants perception. They were analyzed using the Thematic Analysis method.

In terms of privacy perception, we identified the reasons why participants were concerned about certain data types, their belief regarding the likelihood of inference and how it happens, their belief in terms of the consequences resulting from the inference and disclosure of their personal data, and what they do to protect their privacy. We also determined several opportunities for data minimization. They include differentiating 2 types of users, understanding why users did not use third party applications nor the Fitbit website, and why they would disable certain sensors.

Acknowledgment

I would like to thank the people without whom this work could not have existed.

First, I would like to thank my professors Mauro Cherubini and Kevin Huguenin for giving me the opportunity to take part to this very interesting study, and for their support and insightful feedback.

Then, I would like to express my gratitude to Kavous Salehzadeh Niksirat, for his great help and guidance.

I would also like to thank Lev Velykoivanenko for helping me during the interviews and for making them much more fun.

Finally, I would like to show my gratitude to Alexandre and my family, for being an incredible moral support during this challenging journey that writing this master thesis was.

Table of content

1 Introduction 8

2 Literature review 9

2.1 Benefits of using a fitness tracker 9

2.2 Privacy risks related to the features fitness trackers 9

2.3 Privacy risks of wearing a fitness tracker 10

2.4 Privacy knowledge and concerns 11

2.5 Research questions 13

3 Methodology 14

3.1 Survey 14

3.2 Interviews 15

3.2.1 Data collection 15

3.2.1.1 Participants selection 15

3.2.1.2 Recruiting participants 15

3.2.1.3 Participants' background 15

3.2.1.3.1 Participants' group distribution 15

3.2.1.3.2 General background 16

3.2.1.3.3 Technology background 16

3.2.1.4 Interview protocol 17

3.2.1.5 Conducting interviews 17

3.2.1.6 Interview transcription 18

3.2.2 Data analysis: Thematic Analysis (TA) 18

3.2.2.1 What is TA 18

3.2.2.2 Why chose TA 19

3.2.2.3 Quote selection 19

3.2.2.4 Coding process 19

3.2.2.5 Theme generation 19

4 Results 21

4.1 Privacy perception 21

4.1.1 Privacy concerns and belief about consequences 21

4.1.1.1 Not worried 21

4.1.1.1.1 The data is not sensitive 21

4.1.1.1.2 Negative consequences are unlikely 22

4.1.1.1.3 Trust in Fitbit 24

4.1.1.1.4 Safe country 24

4.1.1.2 Worried 25

4.1.1.2.1 Worried about fitness data and personal data 25

4.1.1.2.1.1 “It’s private, I don’t want to disclose it” 25

4.1.1.2.1.2 Family arguments 26

4.1.1.2.1.3 Profiling: manipulation and discrimination 26

4.1.1.2.2 Worried about location, psychological data, and pervasiveness 28

4.1.1.2.3 Cannot give concrete examples of negative consequences 29

4.1.1.2.4 Control over data 29

4.1.1.2.5 Cannot escape data collection 30

4.1.2 Inference 31

4.1.2.1 Inferable based on physiological data 31

4.1.2.2 Inferable if other data is used 31

4.1.2.3 Inferable because it was entered in the application 32

4.1.3 Privacy protection strategies (PPS) 34

4.1.3.1 Bracelet removal 34

4.1.3.1.1 Bracelet removal is useful to protect privacy 34

4.1.3.1.2 Bracelet removal is useful if it is removed for a long time 35

4.1.3.1.3 Removal is useless to protect privacy 35
4.1.3.2 Disabling sensors 36
4.1.3.2.1 Beliefs about disabling sensors 36
4.1.3.2.2 Design preference for disabling sensors 37
4.1.3.2.3 Perceived usefulness of disabling sensors 38
4.1.3.2.4 Reasons for disabling sensors 39
4.1.3.3 Other privacy protection strategies 40
4.1.3.3.1 Thinking before posting anything 40
4.1.3.3.2 Give as little personal information as possible 41
4.1.3.3.3 Permission management 41
4.1.4 Reaction to inference results 42
4.1.4.1 Surprise 42
4.1.4.2 Curiosity 42
4.1.4.3 Change in privacy perception 43
4.2 Perceived utility 44
4.2.1 Wearing attitude 44
4.2.1.1 Abandonment 44
4.2.1.2 Wear it only for sports 45
4.2.1.3 Wearing all day 45
4.2.1.4 Wore the bracelet more for experiment 46
4.2.2 Types of Fitbit usage 47
4.2.2.1 Website usage 47
4.2.2.2 Third-party applications 47
4.2.3 Data minimization 49
4.2.3.1 Steps intervals preference 49
4.2.3.1.1 100 and 500 steps interval 49

4.2.3.1.2 1000 and 2000 intervals 50
4.2.3.1.3 All intervals are useless compared to precise number 50
4.2.3.2 Precision need 51
4.2.3.2.1 A general idea is precise enough 51
4.2.3.2.2 As precise as possible 52
5 Discussion 54
5.1 Privacy perception 54
5.1.1 General privacy concerns 54
5.1.2 Inference knowledge and belief 54
5.1.3 Privacy protection strategies 55
5.2 Data minimization 57
5.3 Comparison with survey results 58
5.4 Limitations and future work 58
6 Conclusion 60
7 References 61
8 Appendix 65

List of figures

Figure 1 – Example of the different intervals as alternative to the precise number of steps 18

Figure 2 - Themes and sub-themes generated in the TA 20

List of tables

Table 1 - Participants' group distribution 16

Table 2 - Personal data asked about in the survey and interviews. Participants were asked if they believed the above-listed types of personal information could be inferred, and if they would be worried if it could 18

Table 3 – Is bracelet removal useful to protect privacy : distribution of participants' belief about usefulness of removing the Fitbit bracelet to protect their privacy 34

Table 4 – Are sensors really disabled : distribution of participants' belief about disabling sensors 36

Table 5 – Is disabling sensors useful to protect privacy : distribution of participants' belief about usefulness of disabling sensors to protect privacy 38

Table 6 – List of sensors users are willing to disable 39

Table 7 – Wearing habits : distribution of participants' wearing habit at the time of the interview ... 44

Table 8 – Intervals evaluation : distribution of how participants rated each interval 49

Table 9 – Level of detail used: distribution of participants' preferred level of detail to check their step count 51

1 Introduction

Wearable technology, and more particularly healthcare wearable like fitness trackers are increasingly popular. These devices sense, process, and store fitness-related data, such as steps, heart rate, calories, and even sleep pattern seamlessly and continuously. A survey with 1000 respondents conducted by PwC ("The Wearable Life 2.0: Connected living in a wearable world," 2016) established that health is first reason for purchase of wearable and that across all gender, age, and ethnicity, fitness trackers are the most popular wearable technology.

Fitness trackers are designed to motivate users to reach their fitness goals, be it to live a healthier lifestyle, to increase their physical activity level, to lose weight, etc. To do so, they use different types of behavior change techniques, such as leader boards, vibration alerts, rewards, feedback on performance, and so on, are implemented in activity trackers (Duncan et al., 2017; Mercer et al., 2016). But foremost, they collect seamlessly and continuously different types of data, such steps number, heart rate, or sleep pattern. This information is captured, processed, and stored in order to provide insights to their users. They are also sometimes shared with third party application based on the user need. However, it also raises privacy issues. Information seemingly inconsequential can be used to infer personal information that the users might not want to disclose. For instance, the data collected by a smartphone accelerometer can be used to infer alcohol consumption (Arnold et al., 2015). Unfortunately, users are generally unaware of the potential negative consequences stemming for the usage of fitness tracker.

In order to gather information on how fitness trackers could be designed to preserve privacy without damaging their perceived utility, we try to understand how users perceive privacy in the context of fitness trackers and what is their perception of utility when using wearable devices.

To do so, we conducted 19 interviews. They took place after a 4-month experiment carried between May 2020 and September 2020. Participants were given a fitness tracker, the Fitbit Inspire HR, in order to collect their fitness data (step count, sleep, and heart rate). At the end, they answered a survey which results were used to create the interview protocol. In this master thesis, we focus on the interviews; their analysis method, their results, and their implications in terms of privacy and utility perception.

The fitness tracker given to the participants, the Fitbit Inspire HR, will be named differently in this paper: "the Fitbit", "the bracelet", "the (fitness) tracker", or "the watch". All these terminologies refer to the same object, the Fitbit Inspire HR.

2 Literature review

2.1 Benefits of using a fitness tracker

Research has investigated the benefits of healthcare wearables, and more specifically of fitness tracker. (Fritz et al., 2014) found that users appreciate fitness tracker because it motivates and helps them make durable change in their activity, health, and well-being. Similarly, a PwC report ("The Wearable Life 2.0: Connected living in a wearable world," 2016) about wearable declared that health is the first reason why wearable are purchased. The benefits of fitness tracker to improve health has been demonstrated in many papers. For instance, (Randriambelonoro et al., 2017) found that activity trackers can help obese and diabetic users to pursue healthier lifestyle by increasing their physical activity awareness and therefore motivating them to increase they physical activity level. Similarly, (Saksono et al., 2018) found that fitness trackers can help even people with lower socio-economic status to increase their physical activity.

There exists fitness tracker users who are not primarily motivated by health. Called "self-quantifiers", this minority likes to track and quantify different aspects of their lives, in order to gain insights about their own behavior and habits. They use devices like fitness trackers to reach in their tracking goal. (Choe et al., 2014; Wolf, 2010)

Therefore, fitness tracker users' needs vary depending on the individual. (Burbach et al., 2019)'s results demonstrate that there are three types of users, each ranking features differently: "*facts enthusiasts"* prioritize accuracy, "*data protectors"* prioritize privacy, and "*benefit maximizers"* prioritize utility.

For users, important characteristics of a fitness tracker are usefulness, ease of use (Spil et al., 2017), playfulness, pervasiveness, persuasiveness, personalization, and privacy features (Gao et al., 2015; Randriambelonoro et al., 2017) . These characteristics are implemented through different features. The ones users utilize daily are accountability, getting credit, setting goals, winning rewards, self-reflecting, and engaging in social activities such as sharing results with peers, competing with them or receiving support (Fritz et al., 2014; Niess and Woźniak, 2018; Ridgers et al., 2018) ("The Wearable Life 2.0: Connected living in a wearable world," 2016).

2.2 Privacy risks related to the features fitness trackers

However, despite how useful these functions are, they are double-edged sword because can raise privacy risks. For instance, pervasiveness, persuasiveness, and personalization. While they increase users' experience and utility, they also increase their privacy risks by encouraging them to give as much personal data as possible. This is confirmed by (Aktypi et al., 2017) who found that users are encouraged and willing to enter precise personal information in the device in order to obtain the greatest accuracy and gain the best benefits from the device.

Social activities such as sharing are not exempt from privacy threats, even if users do not perceive it. Users' practice of sharing fitness data on social networks is a common habit that is motivated by its resulting benefits for the users. More precisely, the motivation to share fitness information are diverse and include accountability, advice, competition, providing help and motivation and emotional support to others, as well as track and improve health by sharing to physicians (Alqhatani and Lipford, 2019). Additionally, shared fitness data is usually met with support and approval of the sharing recipients, and no negative repercussions for sharing this information is experienced by users (Schneegass et al., 2019). In short, users receive great benefits by sharing their fitness data. Finally, (Hallam and Zanella, 2016) explain that because the privacy threats appear distant and vague to the users, their sharing behavior is driven more by the concrete short-term rewarding intentions (i.e. sharing on social media) than by the abstract long-term risk-avoiding intentions.

Regarding privacy features, they are considered important properties of wearable fitness devices (Gao et al., 2015), although being a top priority only for *"data protectors"* (Burbach et al., 2019). Similarly, privacy has been found to be the third barrier to adoption of fitness trackers, behind price and being unsure of using the device ("The Wearable Life 2.0: Connected living in a wearable world," 2016). Additionally, (Li et al., 2016) found that if the perceived benefits are higher than the perceived privacy risks, the adoption of the device is very likely. Likewise, according to (Yang et al., 2016; Zimmer et al., 2020), perceived benefits have a greater impact than perceived risks in the decision of fitness tracker adoption. Therefore, privacy influence adoption, but is not its main driver.

2.3 Privacy risks of wearing a fitness tracker

In spite of the usefulness of fitness trackers, they create privacy risks especially due to the continuous collection of personal data, the incentives to add more personal information, and users' sharing habits. The literatures have found that fitness trackers users are exposed to diverse privacy risks.

First, some of the risks found happen in the communication between the device and the cloud storage. For example, they arise from misuse of authorization protocol or from unauthorized data share (Mendoza et al., 2018), but also from traffic analysis that makes it possible to learn about the users' patterns and activities, such as their number and duration of workouts, despite the transferred data being fully encrypted (Kazlouski et al., 2021).

Second, privacy risks also emerge from users' personal data being disclosed to third parties or malicious individual, despite how innocuous some sensors and data may seem to the users. For instance, (Weiss et al., 2016) have been capable of inferring hand-based activities, such as eating, based on the accelerometers and gyroscopes contained in a smartwatch. Similarly, (Yang Liu, 2018) created a system that provides highly successful guesses of what users are typing on their phone, based on the accelerometer and gyroscope worn on their wrist. Another study investigated if the alcohol intoxication level of a smartphone users could be inferred from their gait, based on the accelerometer data of the phone. Their model had up to 70% of accuracy (Arnold et al., 2015). Similarly, a paper demonstrates that a user's transport mode, location, on-screen taps, smoking and heart disease can

be inferred with a high probability based on the sensor data and the profiling data provided by the user when registering to the service (Torre et al., 2016).
Moreover, the interest of the psychology field to use fitness trackers to gather data, and based on them to understand users' patterns of behavior, illustrates how powerful and efficient these devices are for inference (Ihsan and Furnham, 2018).

Despite companies regularly taking actions to increase the privacy security of their users, it is not always efficient. The Strava case illustrates this well: after their infamous incident where they released a heatmap of their users' activity, that revealed secret US military operations in sensitive locations (Alex Hern, 2018), they implemented a new privacy feature. Called "Endpoint Privacy Zone", it allowed users to hide sensitive locations on the map, such as home or workplace. However, a study found that 95,1% of moderately and highly active users are at risk of having their protected locations extracted by an attacker (Hassan et al., 2018).

2.4 Privacy knowledge and concerns

Overall, fitness tracker users have a basic understanding of privacy risks related to the collection of their fitness data (Lowens et al., 2017). They tend to know very little about how their data is being collected, stored, and anonymized (Aktypi et al., 2017). One factor contributing to this is the fact that only a minority of users read companies' privacy policies. However, despite reading them, users often do not remember their details, or they did not understand the complex writing and legalese used, which requires a lot of effort to comprehend (Aktypi et al., 2017; Vitak et al., 2018). For these reasons, privacy policies do not support decision making for concerned users (Antón et al., 2010; Jensen and Potts, 2004).

Furthermore, users to not understand entirely if, and how personal can be inferred based on their fitness tracker. (Rader and Slaker, 2017) found that users conceptualize the collection of *only* 3 types of data: the ones they manually entered (age, weight, etc.), the ones measured by the tracker (steps, heart rate, etc.), and the ones calculated by the device (calories burned, activity level, etc.). However, they do not apprehend that personal information could be inferred. (Schneegass et al., 2019)'s result show that this is particularly true for non-expert users who are not completely aware of the difference between sensor data (e.g. accelerometer raw data) and derived-information (e.g. steps count). Similarly, (Kröger, 2019) found that the potential privacy implications when sharing sensor data are not well-understood. Overall, users have no idea how raw sensor data could be used for purposes other than activity tracking, and how it could reveal personal information they did not want to disclose. Therefore, truly informed consent to data processing, as required by the GPDR (General Data Protection Regulation) can often not be obtained (Kröger, 2019).

Users tend to create distinct privacy rules for different data types as they consider some types of data as more acceptable to share than others (Zimmer et al., 2020). This is in line with (Alqhatani and Lipford, 2019)'s results, who found that sharing behavior hinges more on the acceptable norm and self-presentation, rather than on the sensitivity of the information. Finally, users' sharing behavior depends on the sharing recipient and the data type shared (Gabriele and Chiasson, 2020). In other

words, users share specific types of data with different groups, and on different platforms if they use several. For instance, (Prasad et al., 2012)'s participants shared their fitness data more with specific third parties than with family and friends, and less with the general public. Similarly, in (Raij et al., 2011)'s study, users tend to be more willing to share to researchers and other study participants, but their concerns increase when sharing it to the general public. However, other studies found that users preferred sharing with their close social circle, such as friends or colleagues (Gabriele and Chiasson, 2020; Zimmer et al., 2020).

Users of fitness wearables claim that privacy and security is their top overall concerns (Burbach et al., 2019). They worry the most about location-related data, as it is the most frequent concern cited in the literature (Kröger, 2019; Raij et al., 2011; Zimmer et al., 2020). But their apprehension also lies in media from cameras (Kröger, 2019), personal identifiers (Zimmer et al., 2020), private conversations (Raij et al., 2011), or having data uploaded to a server and health data being visible to others (Fritz et al., 2014). More specifically, they fear their data will be disclosed without their consent (Raij et al., 2011), especially to malicious parties, such as criminals or thieves (Motti and Caine, 2015). Similarly, (Lowens et al., 2017) found that they worry about the unintended usage of their personal data and the lack of control over it. In a study investigating users' perception towards Automatic Personality Assessment systems, participants reported that they would be less concerned if they could review, exclude, or pause the data collection of such system. In other words, their concerns would decrease if they had more control over their data (Kim et al., 2020).

However, they do not perceive the data collected by the fitness tracker as sensitive. Compared to other health information (such as diagnosis, pregnancies, ...), fitness data is considered less sensitive (Prasad et al., 2012). Trackers monitoring heart rate, steps, and pulse for instance, which are typically collected by fitness trackers, are usually seen as inoffensive to the users' privacy (Motti and Caine, 2015). Finally, (Lee et al., 2016) found that users are the least concerned about data related to measurements of their body. It is clear that the majority of fitness tracker users consider personal fitness data as low-sensitive or unsensitive. Therefore, they are unconcerned by it (Lowens et al., 2017; Zimmer et al., 2020) . These findings are in line with the fact that users are unaware of the potential negative consequences rising from the collection of fitness data, as they cannot worry about something they do not consider a threat.

When showed different privacy risks scenarios related to fitness tracker data, users generally consider them credible but unlikely (Gabriele and Chiasson, 2020). According to (Aktypi et al., 2017), one of the reasons why participants were unable to associate themselves with presented risks scenarios is because no one they know had been a victim of it. Similarly, the trust users have in fitness tracker companies is influenced by the media and what they cover regarding the businesses' practices (Zimmer et al., 2020). This contributes to the users' feelings that the threat is distant and vague, thus reducing their potential concerns (Hallam and Zanella, 2016). In short, if they have never heard of it happening to someone, they do not think it will happen to them. Even the most privacy concerned users have difficulties giving concrete examples and explanations of potential privacy risks and of the resulting negative consequences (Aktypi et al., 2017; Gabriele and Chiasson, 2020; Lowens et al., 2017; Zimmer et al., 2020).

Understandably, it has been found that users do little to protect their data (Gabriele and Chiasson, 2020). Furthermore, most users do not do the efforts required to actively protect their privacy, even if they claim caring about it (Vitak et al., 2018). This is called the Privacy Paradox (Norberg et al., 2007).

Fortunately, this state of unawareness in not static. Researchers have provided efforts to evaluate if and how users' concerns can evolve. They found that users' concerns increase when they are explained how their fitness information can be combined with other data in order to infer other personal information, or when told about the risk of reidentification (Raij et al., 2011; Zimmer et al., 2020). (Aktypi et al., 2017) developed a tool that help users visualize and understand how their sharing behavior on social networks and fitness tracker leads to identity exposure risks. Even skeptical users found the tool very helpful to understand the risks better and be more aware of them.

2.5 Research questions

Despite all the research on the topic, the understanding of fitness tracker users' utility and privacy perceptions is still limited. Therefore, the main objective of this master thesis is to better understand how users perceive the implications of privacy protection mechanisms on privacy and utility. This master thesis takes place after a 4-month experiment and a survey. In order to gain greater insights on the survey results, it has been decided to conduct interviews. Therefore, this master thesis focus on the qualitative analysis of the interviews, which were conducted in order to answer the following questions:

RQ1. How do users perceive **privacy** in the context of fitness trackers? In particular, what types of private information do users think can be inferred? What would users be willing to do to protect their privacy (such as removing their bracelets or disabling data collection)?

RQ2. What is the users' perception of **utility** when using wearable devices? In particular, which features, platforms, or types of data do users find useful and what benefits do people derive from fitness trackers?

3 Methodology

3.1 Survey

At the end of the 4 months experiment, the participants had to fill a questionnaire (n=227). We did not work on it, but since the interview protocol was created based on the survey, we will only briefly summarize its content and results here.

It contained questions related to the participants' perception of privacy and utility, in the context of their fitness tracker. More precisely, it included questions to evaluate participants' self-perceived privacy risk awareness, privacy-related actions, sensors-related knowledge and actions, inference-related concerns and perceived likelihood (see Table 2 for the types of data they were asked about), perceived usefulness of the Fitbit data, bracelet removal habits, and preference about the precision of the data.

Regarding the privacy-related questions, they found that more than half of the participants considered that they were at least slightly aware of the privacy risks associated with fitness trackers. About inference, they were the most concerned if Personality traits, Socioeconomic status, Political views, Alcohol and tobacco consumption, Sexual activity, and Illegal drug consumption could be inferred from fitness tracker data. A large majority of participants indicated that age, sexual activity, menstrual cycles, and drug consumption could be inferred with a moderate precision at least. They thought the opposite for religion, political views, and sexual orientation. Only a minority of participants performed privacy-related actions, which included changing the privacy settings, and changing the security settings. Concerning their wearing habits, only a minority reported removing it for privacy reasons. Regarding sensors, the participants expressed the most desire to disable the GPS and the heart rate sensor on their tracker. More than half of the participants would find an option to disable sensors 'Slightly' to 'Extremely' useful in terms of preserving their privacy.

In terms of data minimization potential, they discovered that a vast majority of participants never used other connected devices nor Fitbit's website to check their data. When checking their step data, they would rather see the total steps per day, and have little interest for more detailed results. Finally, when presented intervals as alternative to the total steps count, the majority found largest intervals of 1000 steps and 2000 steps to be useless compared to the precise number. On the other hand, the smallest interval of 100 was considered to be at most slightly less useful than having the precise number by more than a third of the participants.

Based on these results, survey questions worth going into more deeply were selected. They constituted the base of the interview questions.

3.2 Interviews

After analyzing the results of the survey, we wanted to get additional insights about the participants' perception of privacy and utility in the context of the Fitbit bracelet. Based on selected survey questions, we created questions for the interviews. Therefore, the structure and the questions of the interview protocol are based on the survey.

We conducted 19 individual semi-structured interviews between December 2020 and in January 2021 with survey respondents. Due to the COVID-19 situation, the interviews of January 2021 were conducted remotely via Zoom. A total of 14 interviews were done physically, and 5 were conducted remotely.

3.2.1 Data collection

3.2.1.1 Participants selection

To choose which survey respondents to interview, we classified them into 4 different groups, depending on when and how often they removed their bracelet (expect for charging it). By doing so, we hoped to get a better understanding of the relation between participants' privacy concerns related to the Fitbit bracelet, and their wearing habits.
The 4 groups are:

1. Removed the bracelet *frequently* during the day, and *rarely or never* wore it at night.
2. Removed the bracelet *frequently* during the day, and *often or always* wore it at night.
3. Removed the bracelet *rarely or never* during the day, and *rarely or never* wore it at night.
4. Removed the bracelet *rarely or never* during the day, and *often or always* wore it at night.

Our goal was to interview the same number of participants from each group, as well have an equal distribution of male and female. Based on these criteria, we established a list of participants to be contacted to take part in the interviews.

3.2.1.2 Recruiting participants

Participants matching the selection criteria were contacted through HEC-LABEX[1], the behavioral research laboratory of HEC Lausanne, using their own database of participants called "*ORSEE subject pool*". An email invitation was sent to selected participants to take part in an interview. Finally, 19 participants registered and all of them attended the scheduled interviews.

3.2.1.3 Participants' background

3.2.1.3.1 Participants' group distribution

Among the 19 participants we interviewed, 1 is from group 1 (5,3%), 3 are from group 2 (15,8%), 10 are from group 3 (52,6%) and 5 are from group 4 (26,3%). Therefore, we did not reach our distribution goal (cf. Table 1).

[1] https://www.unil.ch/hec-labex/home/menuinst/about-us.html

Group number (time when bracelet is removed)	**Number of participants in the group**	**Percentage of participants being in this group**
1 (day + night)	1	5,3%
2 (day)	3	15,8%
3 (night)	10	52,6%
4 (never)	5	26,3%

Table 1 - Participants' group distribution

3.2.1.3.2 General background

Regarding gender, 12 identify as female (63%) and the others as male. On average they are 21 years old, ranging from 19 to 25 years old. Regarding their education, they are all students: 11 are UNIL students (58%) and the others are EPFL students.

3.2.1.3.3 Technology background

All our participants reported using a smartphone and a laptop every day. They are at ease with technologies and have no difficulty using their basic and common features on a daily basis.

> *"I have a good relationship [with technology], I use it quite easily." (P8)*

> *"It's normal to work with technology, I don't know any other way to do it." (P6)*

They see technology as a normal and very useful daily tool, but due to the lack of alternative to it, they also consider it as a need. This necessity to own and use technology is exacerbated because of the current COVID-19 situation, where remote studying is the new norm students.

> *"We have to use technology nowadays, we don't really have a choice, we are bound to it for work." (P17)*

> *"[technologies are] very useful in my everyday life, especially in the current situation" (P32)*

Participants are not only reliant on work tools. Eight participants mentioned using social media regularly. In particular, one mentioned feeling dependent on it and on messaging applications, such as Instagram and WhatsApp. This emphasizes how using technology is part of their normal daily life activities.

> *"I feel quite close to technologies but also dependent on certain functions of technologies, like Instagram and WhatsApp" (P1)*

Despite how practical technologies can be, their use raises concerns related to privacy. Especially, they expressed how privacy cannot coexist with technology usage. We will develop this aspect later in the result section.

> *"Using technology means saying goodbye to part of your private life." (P6)*
>
> *"[Technology is] very useful, but it's scary sometimes how efficient it is, in relation to the amount of data collected, especially with social networks, they have even more info about us" (P22)*

Our participants were all comfortable with using technology, although they were not expert.

3.2.1.4 Interview protocol

We conducted individual semi-structured interviews, either physically or remotely using Zoom. As explained before, the structure of the protocol and its questions were based on the survey. We opted for this structure because we wanted participants to develop and explain their answers to the survey, as well as letting them speak freely about the topic.

We created personalized interview protocol for each participant, based on their answer to the survey. This way, we could ask very specific and personalized questions.

The interviews were recorded using an audio recorder, or using the Zoom recording feature for remote interviews, to facilitate transcription and analysis later.

Each participant filled a consent form before the interview started. It explained the course of the experiment and the participants' obligation in order to finish it and be eligible for the compensation payment. It also clarified that their data is kept confidentially and is used and processed for this experiment only. At the end of the interview, participants signed the financial form to receive the compensation payment. For remote interviews, the documents were sent by emails.

The interview protocol and the consent form can be found in appendix A and B, respectively.

3.2.1.5 Conducting interviews

Before starting the interviews, we tested the protocol with a relative of one of the researchers. The 19 interviews were conducted between December 2020 and in January 2021. Four interviews were conducted remotely in January 2021 via Zoom. Each interview was conducted with 1 participant and 2 interviewers; one researcher asked the questions, and the other took notes but was free to ask questions well. Participants were encouraged to explain and develop their answers.

To better remind the participants about the survey questions we were referring to during the interview, we showed them the following information. Figure 1 illustrates the question about interval. Table 1 lists the different types of data used in the inference questions.

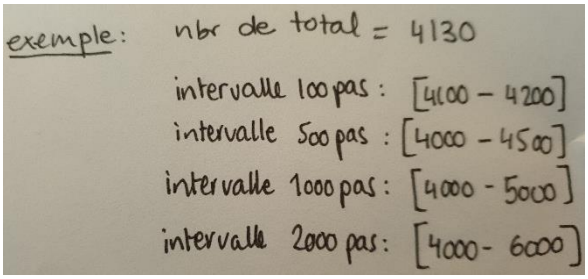


Figure 1 – Example of the different intervals as alternative to the precise number of steps

Sexual activity
Age
Alcohol and/or tobacco consumption
Drug consumption
Menstrual cycles
Gender
Political orientation
Sexual orientation
Personality
Religion
Socio-economic status

Table 2 - Personal data asked about in the survey and interviews. Participants were asked if they believed the above-listed types of personal information could be inferred, and if they would be worried if it could

On average interviews lasted 42min, with the shortest interview being 24min and the longest 1h03min.

After each interview, we debriefed and recorded it. During debriefing, we mentioned anything that stood out to us and tried to recall the participants' answers.

3.2.1.6 Interview transcription

Once all interviews had been conducted, they were transcribed on Word documents. To make this process faster, we used the notes taken during the interviews as a basis. We expanded them with the transcription of the debrief audios first. Finally, we completed it with the transcription of interview audios.

3.2.2 Data analysis: Thematic Analysis (TA)

3.2.2.1 What is TA

Thematic analysis (TA) is a qualitative analytic method. It is an iterative process that enables identifying, analyzing, and reporting key patterns, also called themes, within data. In practice, this is done by coding data and then combining codes into defined groups that represents the themes (Braun and Clarke, 2006).

One benefit of TA is its flexibility, as it can be used with a variety of theoretical frameworks, can be employed in different ways within them, and can be used on dataset of any size. This leads to a large spectrum of applications. In particular, it can be used to describe dataset amply or more precisely with a focus on specific aspects of the data. Themes can be generated from the data explicitly or at a more latent level, or with a mix of both. Finally, TA can be used for both inductive (data-driven) and deductive

(theory-driven) analyses. Its second advantage is accessibility, as the researchers do not need a high level of theoretical and technological knowledge to use it (Clarke and Braun, 2016).

3.2.2.2 Why chose TA

The flexibility and accessibility of TA are the reasons why this method was chosen over others, such as Grounded Theory. These methods require practice and time to be able to use them correctly. Due to the time constraint and this being a first analysis for its author, TA appeared as the best option.

3.2.2.3 Quote selection

The TA process started with the selection of the quotes to be coded. As not all that was said by participants was relevant in relation to the research questions, it was important to do it. This step was done in a deductive (theory-driven) way, as the quotes chosen were related to our research questions.

The quotes were selected by reading the transcriptions and highlighting relevant quotes. After that, the highlighted quotes were copied into a spreadsheet document that lists all the quotes per participants. In total, there are 382 quotes which equals to around 20.1 quotes per participants.

3.2.2.4 Coding process

The coding process is very important, as it constitutes the base for the analysis. It is also time consuming because each quote is coded individually, and 4 iterations were needed in order to create all the codes and to consistently code each the quotes.

This phase was planned to be done in a deductive (theory-driven) fashion, rather than in an inductive (data-driven) one. In other words, the coding was guided by the research question rather than by the quotes. However, in practice it happened to be more in a combination of both.

Finally, between 1 and 3 codes were associated for each quote. A total 137 different codes were used. The list of the codes can be found in appendix C.

3.2.2.5 Theme generation

Once the coding process was finished, we started to generate the themes. A theme covers an important aspect of the data, in relation to the research questions. There is no clear and fixed rules to determine what constitute a theme. It rests on the researchers' judgement (Braun and Clarke, 2006).

Generating themes and sub-themes is done by combining codes with similar meaning or ideas. To do so, we used an online tool called Miro to organize all the codes into a tree, which you can see in appendix E. We reviewed it until were satisfied with the code classification, the generated themes and the overall coherence between the themes and the codes.

Finally, 2 main themes (Privacy Perception, Perceived Utility) and 7 sub-themes (privacy concerns, inference, belief about negative consequences, privacy protection strategies, wearing attitude, types of Fitbit usage, data minimization) were generated. However, it is important to note that the 2 main

themes relate to what we asked participants; it is therefore normal that these topics are found in participants' responses.

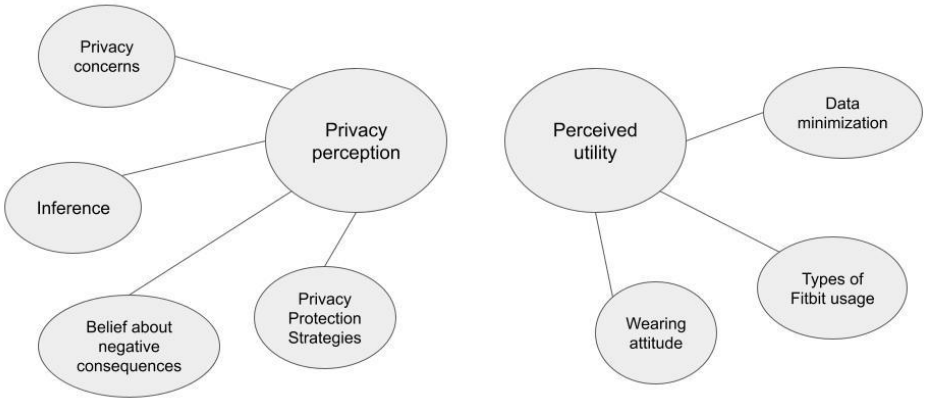


Figure 2 - Themes and sub-themes generated in the TA

The first themes related to Privacy Perception and to *RQ1: "How do users perceive privacy in the context of fitness trackers?"*. They include quotes in which participants talk about how they perceive privacy in the context of fitness tracker. More specifically, they comprise participants' concerns related to their data (Privacy concerns), their belief about inference likelihood and how it happens (Inference), their belief about the consequences resulting from inference (Belief about negative consequences), and what actions they take to protect their privacy (Privacy Protection Strategies (PPS)).

The second themes, relates to Perceived Utility and to *RQ2: "What is the users' perception of utility when using wearable devices?"*. They concern how the participants wear the bracelet (Wearing attitude), what features and platforms they use (Types of Fitbit usage), and what is level of detail they need to have their privacy preserved without damaging their utility (Data Minimization).

4 Results

In the following sections, we will first give a quick overview of the level of concerns. After that, we will summarize the reasons why many are unconcerned about their personal data being inferred. Then, we will focus on participants' concerns. Particularly, we try to understand why they are worried, what they believe can happen if their personal data is collected and what they do to protect their privacy. Finally, we will investigate some utility aspects. Specifically, what features they use the most and their preference in terms of precision.

It is important to note that the initial study stopped in September 2020. After that month, the participants were free to stop using the bracelet if they wanted. Therefore, during the interviews some participants had stopped wearing the bracelet, while other were still wearing it.

To preserve anonymity and for legibility concerns, participants will be named P1 to P32.

4.1 Privacy perception

4.1.1 Privacy concerns and belief about consequences

We asked participants if they would be concerned if their *personal data* (cf. Table 2) was inferred based on Fitbit data. We also asked them what data they perceived as sensitive.

4.1.1.1 Not worried

Among the 19 participants, 17 reported not being worried about at least one of the bracelet data, or by at least one of the *personal data* (cf. Table 2). We asked them to elaborate on why they were not concerned, and they gave us different reasons.

4.1.1.1.1 The data is not sensitive

The first reason is that they did not consider the data as sensitive personal information. This is particularly true for the data collected by the bracelet specifically (sleep, heart rate and steps count).

> *"There are not very personal things on my Fitbit account. And if the info collected by the bracelet were to become public, I wouldn't be worried." (P7)*

> *"No [I'm not worried], the data collected [by the Fitbit] is not part of my private life." (P19)*

> *"I'm not bothered by strangers having access to [the Fitbit] data, it's not a danger to my privacy. (P32)*

Regarding the *personal data* (cf. Table 2), they were not worried because they did not think it is taboo, and they believed they had "nothing to hide".

> *"[All personal data except sexual activity and socio-economic status] It's not taboo, it's normal, I don't mind talking about it so I don't worry about other people knowing about it" (P29)*

> *"I don't think it's worrying [if others know that] because I have no shame or anything to hide about this data [all personal data]." (P23)*

Additionally, for data such as age or gender, participants are so used to sharing it that they consider it normal to do so. Moreover, because it is data that defines them and that cannot be controlled, they did not mind if it was shared.

> *"[Not worried about Age, Gender because] It's info that I often give, I don't find it personal, it doesn't matter if other people know that. This is who I am, I have no control over it." (P6)*

> *"Age is who I am, I don't choose it so I'm not concerned [if others know it]." (P26)*

Moreover, they were less worried about sport-related data.

> *"[Not worried about age, gender, alcohol or tobacco consumption, drug consumption, sexual orientation, and menstrual cycle because there are] related to sports activities so it is useful information for my goals." (P13)*

> *"I don't mind being [localized] during sports activities." (P21)*

The main reason participants are not worried about their data is because they do not perceive them as sensitive. They are comfortable sharing data because they have "nothing to hide", are used to sharing it, or because it is related to their sport activities, therefore it is beneficial for them.

4.1.1.1.2 Negative consequences are unlikely

Another reason is that they did not think the collection of such data could lead to any negative consequences.

> *"This is not problematic data if exposed, I don't think it will have a negative impact for me." (P19)*

> *"I cannot determine to whom and for what it could be used, so I think it's ok [if others know all the personal data but alcohol and drug consumption]." (P32)*

Particularly, they believed such information could not be used against them, as it is not “dangerous” data.

> *“I don't think an ill-intentioned individual can use [age, gender, religion, sexual and political orientation, personality, menstrual cycle] against me. Because it's something "human" in society at large.” (P25)*

Even if some recognized that the data can be sensitive, they did not think they were a potential target. This means that they did not believe someone would choose to collect their data and harm them personally, either because they were not famous…

> *“I am not worried about my particular case, I do not think I am a potential target, I am not a public figure.” (P7)*

… or because they saw data collection as a large-scale process using anonymized information.

> *“As the data collection is on a very large scale, I don't think I'm really targeted personally, they don't focus on a specific person” (P1)*

> *“I don't pay too much attention to the info I give because it's anonymized and I don't think I'm a potential target.” (P18)*

Despite knowing their data could be used for commercial purposes, some participants did not mind it and perceived targeted ads as inoffensive.

> *“If it's to see what I've bought or where I've been, as long as it's used wisely I don't mind. Wisely, that is, targeted ads that I can ignore, but not repeated calls or people showing up at my door.” (P16)*

> *“Targeted advertisement is only numbers, it's not a personal application.” (P19)*

They believe nothing bad can happen to them if their *personal data* (cf. Table 2) or fitness data is known. They also think no one would be interested to attack them personally, and their data is anonymized for purposes like advertisement.

However, some participants are aware that privacy risks exist, but they do not mind them.

> *“No, [I have] no worries about my privacy, but I am aware that some things are not protected, but we cannot know what” (P8)*

> *“I am aware that [inferring sensitive information about me] can happen, but I don’t worry.” (P19)*

4.1.1.1.3 Trust in Fitbit

On top of that, some participants expressed feeling safe using Fitbit. They feel as such because it is a known brand.

> *"I have a feeling of trust and I love having the data so I wear [the Fitbit] all the time" (P18)*

Therefore, the participants' lack of concerns is maintained because of the trust they have in the company.

4.1.1.1.4 Safe country

Three participants (P23, P28, P32) acknowledge how much of a free country Switzerland is. Especially compared to others, where being yourself and sharing your opinions, religion, political or sexual orientation can lead to repercussions. They said they would be more worried if they did not live in Switzerland.

> *"I consider that [Alcohol or Tobacco consumption, Personality, Socioeconomic Status, and Political Orientation are] a way of keeping watch on people according to their opinion. It can be problematic if used in a more dictatorial country than Switzerland where there is strong repression against certain political opinions or sexual orientations." (P28)*

> *"We are in a safe country and so there is no risk for me if we guess my religion or whatever, but in other countries it wouldn't, so I understand that might bother some people. As for example with the Uyghurs in China." (P32)*

P23 talks about how discriminations are more likely in an authoritarian country.

> *"But if we were in an authoritarian regime that had access to [all personal data], and that on this basis we were discriminated against, that would be a problem for me." (P23)*

Similarly, P15 mentioned trusting how their data is used and managed in Switzerland thanks to data protection laws (LDP, GDPR). Therefore, they fear that their data would be handled outside of Europe.

> *"I'm especially scared of companies that I don't know of or companies with servers outside of Europe, because the data protection is worse." (P15)*

These participants are aware of how certain of their *person data* (cf. Table 2) could be used against them in some country. But since they live in Switzerland, they consider themselves safe.

4.1.1.2 Worried

4.1.1.2.1 Worried about fitness data and personal data

Reasons for participants' concerns are diverse. They include the fear of potential negative consequences, such as family arguments, targeted advertisements, or profiling. Another explanation is simply that they do not want to share certain information because it is private; they are "shy" about their data.

4.1.1.2.1.1 "It's private, I don't want to disclose it"

One of the reasons why they were worried about their data is that they were shy about it. They did not necessarily think their disclosure would lead to negative consequences. They simply do not want it to be known and would feel discomfort if it was.

Regarding the Fitbit data, only P22 mentioned being concerned about it. They felt uneasy at the idea of their sleep cycles and heart rate being known by others.

> *"I'm uncomfortable sharing my sleep cycles, it's my private life […]. Heart rate reflects emotional state, stress, anxiety, or health concerns or drug use. For me it's very personal, especially since it's being captured non-stop. This was what bothered me the most during the study." (P22)*

Regarding *personal data* (cf. Table 2), most our participants expressed concerns about at least some of it. They were worried when they considered it private. They believed it should not be shared or known by others.

> *"[This information: Personality, Socio-economic status, Political orientation, Sexual activity, Religion] is part of my private life, I don't want to share it. It must be kept confidential." (P13)*

> *"[all personal data but age and gender] is part of my privacy and I consider it a violation of my privacy to collect this data or to infer things about me […] but it's unlikely [that there will be any negative impact]." (P22)*

> *"[worried if Alcohol, Tobacco, and Drug Consumption are inferred because] it may be relevant for physical activity but I think they don't need to know that, it's none of their business […] No [I don't think it leads to] negative consequences, but we don't know where it's going, for whom, what, how, why, so we don't want to share." (P26)*

On top of that, they do not want it to be inferred because it is not related to their fitness activity.

> *"[I am worried about [Personality, Socio-economic status, Political orientation, Sexual activity, Religion] being inferred because] this information has nothing to do with the practice of my sport." (P13)*

> *"[I am worried about Political and Sexual Orientation, and Socio-Economic status being inferred because] these are data they don't need to know. These data are irrelevant to physical activity." (P26)*

They like to keep their personal information to themselves not because of potential consequences but rather because they are shy about it. They are particularly uncomfortable at the inference of not fitness related data, probably because they did not explicitly allow these *personal data* to be inferred.

4.1.1.2.1.2 Family arguments

Oftentimes, the source of concerns are the participants' parents. Four participants said that if personal data (cf. Table 2) were to be disclosed to their family, it could create arguments, as P16 and P32 said.

> *"If it's my parents [who have access to my location], yes there can be negative consequences." (P16)*

> *"It's about my family, I don't want [my drug and alcohol consumption] to be known, because they are against it." (P32)*

P25 explained the disagreement with their family would stem from the consumption of drug and alcohol.

> *"I am religious, and in my religion drugs and alcohol are forbidden, but I tested them out of curiosity and I don't want it to be known by my family. If they found out about it, yes it would have negative consequences for me." (P25)*

These users were afraid the disclosure of personal data could harm the relation they have with their family.

4.1.1.2.1.3 Profiling: manipulation and discrimination

Another worry of the participants is how their *personal data* (cf. Table 2) could be used to profile them.

> *"Profiling based on data scares me." (P15)*

For example, P26 mentioned how this data can be used to describe them precisely but can lead to a wrong definition of them.

> *"[worried if Religion, Political and Sexual Orientations, and Socio-Economic status are inferred because] it defines me but it's like a technical sheet at a certain point in time of an individual […] these are fluctuating data, and taken out of context they can lead to a distorted view of the individual. " (P26)*

Similarly, P1 said this data can be used to guess their personality traits.

> *"[worried if sexual activity, personality, alcohol/tobacco/drug use, socio-economic status, political orientation, and menstrual cycles are inferred because] I consider them as personality traits and quite personal things" (P1)*

Profiling is linked to 2 other worries, manipulation, and discrimination, that we will discuss in the 2 following sections.

4.1.1.2.1.3.1 Manipulation

The profiles can be used for targeted advertisement.

> *"[Worried about Socio-Economic status, and Political Orientation because] this kind of data can be used for targeted advertising, that's more worrying." (P19)*

Such marketing is often disliked as is it perceived as a form of manipulation, as expressed by P22 and P28.

> *"[extremely worried about the inferences of all but age and gender, because] if I'm being manipulated because they know me well, especially to sell me things I don't need" (P22)*

> *"My psychological or mental state should not be used to sell products." (P28)*

The more accurate the profile is, the more targeted the ads are. P1 expressed concerns regarding it.

> *"Increasingly targeted advertisement on Instagram [concerns me for my privacy]" (P1)*

4.1.1.2.1.3.2 Discrimination

The second worry linked to profiling is discrimination. P21 and P29 feared that their *personal data* (cf. Table 2) could be used to discriminate them.

> *"[Personality] if it can be guessed, it is concerning because of potential discriminations" (P21)*

> *"[Sensitive] data would be the kind that can affect my life, like getting a job" (P29)*

P11 and P22 stated that in the future, their data could be used against them. It would be too late to protect their privacy as a lot of personal data would have already been collected.

> *"So I tell myself, about the police, in the future if they ever want to know what we are consuming [alcohol, drugs]. But it is not a current fear." (P11)*

> *"I fear for the future if Switzerland becomes a state of mass surveillance, we are cornered because we have all already given too much information about ourselves." (P22)*

They do not want their *personal data* to be inferred because they could be used to generate a profile of them, which could be wrong. This profile could be used against them to manipulate them, such as in targeted advertisement, or to discriminate them in their daily life.

4.1.1.2.2 Worried about location, psychological data, and pervasiveness

Most participants expressed high concerns about their location, as they would be very worried if it was accessible to others.

> *"The location on the phone, or on the fitness tracker, it worries me." (P29)*

> *"I don't need these sensors and I don't want to give out this [location] information. Fitbit doesn't need to know where I've been and what route I took to calculate my physical activity." (P28)*

> *"My heart rate is not part of my private life, the geolocation a little more. I'm not worried about people having access to my heart rate, but I am worry if they have access to my location." (P15)*

> *"What worries me is the localization, especially Google that tracks us." (P11)*

They expressed a lot of concerns about location. They did not want it to be known by others.

P22 also talked about two others interesting worry.

> *"Heart rate reflects emotional state, stress, anxiety, or health concerns or drug use. For me it's very personal, especially since it's being captured non-stop. This was what bothered me the most during the study." (P22)*

First, psychological information ("*emotional state, stress, anxiety*"). This type of data tends to be perceived as sensitive by our participants. For instance, P28 said they are more sensitive than "physical" information.

> *"I have no problem if my physical appearance is guessed, but how I am psychologically or mentally, it worries me that it can be guessed." (P28)*

Second, the pervasiveness of data collection ("*especially since it's being captured non-stop*"). The fact that their data is being collected continually is an aspect that increases their concerns. P6 talked about it as well.

> *"The technology and things like the tracker are collecting [data] continuously, it can be used to learn about my habits and that worries me. That's why now I don't wear [the Fitbit] anymore." (P6)*

Location and psychological data are more often considered as sensitive information. On top of that, the pervasiveness of the data collection increases the privacy concerns of some participants.

4.1.1.2.3 Cannot give concrete examples of negative consequences

Apart from the consequences mentioned in the previous paragraphs (family disagreements, profiling, …), users have difficulties giving concrete examples of potential repercussions. Many, like P15 and P25, mentioned risks without explicating them.

> *"There is no point for me because I already know [my socio-economic status, and my drug, tobacco and alcohol consumption], and it can be used against me." (P15)*

> *"[I'm worried because] there are so many risks, people can use your information against you, you never know what can happen. Any little information can be traced back to you." (P25)*

Even when asked to clarify, they remained vague.

> *"I could not say what threats precisely, but the field of possible threats is very large" (P17)*

Similarly, P23 find hard to believe other consequences from targeted advertisements are possible.

> *"I can't imagine how it could go any further than [targeted ads]. And how it could be harmful to me." (P23)*

The participants who said negative consequences were possible were rarely able to give concrete examples. It shows they lack understanding of the privacy risks related to inference.

4.1.1.2.4 Control over data

Some participants expressed concerns about their lack of control on their own data.

> *"Once it's sent, it's left the watch, I lost control of my data. Even if I accept or refuse certain processing of my data, in the end I don't know what is actually being done." (P26)*

P22 said they would prefer that the data stay under their control. Despite enjoying having the Fitbit data, they were too concerned about their privacy to keep wearing the bracelet.

> *"I would prefer that my data does not leave my phone, that it remains under my control. […] What really bothers me is that this is sent to Google, that it leaves my phone. For my own use and analysis, I'm glad to see my heart rate." (P22)*

It was a source of worry for P15 and P16 who distrust how their data is being managed.

> *"I would feel more secure with my data if it didn't go through the [Fitbit] server." (P15)*

> *"I am not afraid of my data being collected, however I do worry not knowing to whom it is transmitted and what they do with it. It's not knowing what they're doing with [my data] behind [that worries me]." (P16)*

Users tend to not know how and where their personal data is processed and are concerned about it. It illustrates their lack of knowledge and control over it.

4.1.1.2.5 Cannot escape data collection

Users are resigned to having their data collected. They think that they cannot do anything to prevent it from happening. They think so because they consider other devices, such as their smartphone or computer, to be a greater privacy risks than the bracelet. Therefore, they do not see the point of stopping using the fitness tracker if they still continue to use other devices that collect similar or more personal information.

> *"Yes [I'd be worried if it was possible to deduce sensitive information about me], but I think you can't really protect yourself from it, because you would have to give up your phone, computer etc. […] Using technology means saying goodbye to part of your private life." (P6)*

> *"Bracelet, Smartphone, Computer: to have one is to give up your privacy." (P25)*

> *"For example for location, there is always another way to locate us even if we deactivate something. For example automatic connections to Wi-Fi because of which we can be located." (P19)*

> *"On a global scale it is possible to do things that I cannot even imagine, particularly because to use a service you often have to agree to give a lot of data, especially since the data is probably stored for a long time." (P22)*

This idea that they cannot escape data collection will be found again across the different results sections.

4.1.2 Inference

We were wondering how users believe inference happens. In the interviews, we had the opportunity to asked participants why they believed some data was inferable and some was not. Some participants, for instance P7 and P15, did not believe inference was possible.

> *"If it were possible [to infer my religion, socio-economic status, and political orientation from the Fitbit], I wouldn't necessarily have agreed to take part in the study." (P7)*

> *"I don't think they would be able to [infer socio-economic status, drug, tobacco or alcohol consumption]." (P15)*

Regarding the participants that believed inference was feasible, they mentioned 3 different ways inference happen: 1) based on physiological data, 2) using other data, and 3) based on manually entered data.

4.1.2.1 Inferable based on physiological data

First, they believed inference of body-related data was possible because the bracelet collects physiological data. Therefore, data such as sexual activity, alcohol and drug consumption, can be inferred from it.

> *"The observations of the watch are physical, they are measurements of the body, so it allows precise inference of physical data." (P19)*

> *"The heart rate makes it possible to guess [drugs, alcohol and tobacco consumption precisely]." (P29)*

According to P15 and P32, these corporal data also informed about their habits, especially through sleep and heart rate data. These habits can be used to guess their age or their personality.

> *"Personality [is inferable] from the regularity, the time I get up, my heart rate." (P15)*

> *"An age interval can be determined from sleep especially, because it gives information about the lifestyle." (P32)*

4.1.2.2 Inferable if other data is used

Many participants did not believe the inference was possible based solely on Fitbit data. They mainly thought so for political orientation, sexual orientation, religion, and socio-economic status.

> *"[political and sexual orientation, religion, socio-economic status] are more thoughts, it's not physical, so the watch can't measure it." (P19)*

> *"[Sexual activity, religion, sexual orientation, personality, socio-economic status, political orientation are not precise because] it is not possible to know that from the data collected by the Fitbit bracelet." (P29)*

Therefore, a few participants believed inference was possible by using other data in addition to the Fitbit data.

> *"I don't really see how [sexual orientation, personality, religion, political orientation] can be guessed, but I think that with a lot of data we can maybe deduce a bit, get an idea." (P22)*

> *"I don't see how [religion, sexual orientation] can be guessed just with this kind of bracelet. Maybe by cross-checking with other data, yes." (P23)*

> *"I tell myself that by observing someone a little bit, we can deduce a lot of things. So, with a little digital information we can surely come to some conclusions." (P25)*

P11 and P22 suggested that the "other data" could be location:

> *"I can't quite explain why [political and sexual orientation, religion, personality] can be precisely inferred, but I think from physical data and location it can be more or less inferred." (P11)*

> *"[Socio-economic status can be inferred] with location by neighborhood, trips, workplace. [Religion can be inferred] whether one is religious, whether one goes to places of worship." (P22)*

P1 implied that internet activity and purchase list can be used to infer different data type:

> *"[Age is accurately inferred] based on the content we watch, our activity on the internet in general, [Socio-economic is slightly accurately inferred] if they know what we're buying." (P1)*

4.1.2.3 Inferable because it was entered in the application

The third way participants said inference happen is by entering the data in the app. Eight participant said so for age, gender, and menstrual cycle. Since this is not what inference is, these answers show that users lack understanding of it.

> *"[Age, gender and menstrual cycle can be inferred precisely because] I entered it in the application." (P1)*

> *"[Age, gender and menstrual cycle can be precisely inferred when it's] based on physical activity, heart rate, or if you enter it in the application" (P16)*

> *"[Age, gender and menstrual cycle can be inferred precisely because] because I can give this information" (P29)*

Some participants had good intuition about how some data could be used alone or combined with others to infer other information. However, they seem to lack knowledge of how inference happens and its extend, especially because they did not mention the usage of raw data (such as accelerometer raw data), and because they considered data entered in the application as a form of inference.

4.1.3 Privacy protection strategies (PPS)

We asked our participants different questions to understand how they protect their privacy. More specifically, we asked if removing their bracelet, or disabling sensors on their bracelet, could be useful to protect their privacy.

4.1.3.1 Bracelet removal

First, we asked our participants if they believed removing their bracelet could be useful to protect their privacy. The following table recapitulate their answers:

Is removing the bracelet useful to protect privacy?	Number of participants	Participants' Number
Useful	9	6, 8, 11, 13, 16, 21, 23, 26, 29
Useless	5	7, 15, 19, 25, 32
Useful if removed for a long time	2	1, 22
No answer	*3*	*17, 18, 28*

Table 3 – Is bracelet removal useful to protect privacy: distribution of participants' belief about usefulness of removing the Fitbit bracelet to protect their privacy

4.1.3.1.1 Bracelet removal is useful to protect privacy

Nine participants believed removing their bracelet can be useful to protect their privacy. Among the participants who explained why they thought so, P26 and P13 believed that if the bracelet is removed, it cannot collect any personal data, thus privacy is preserved.

> *"[removing the bracelet in useful to protect my privacy] because there is no more traces of me" (P13)*

> *"Yes, if I no longer wear it, then only inactivity is collected, so my privacy is protected" (P26)*

P11 gave a concrete example of a situation where they removed their bracelet to protect their privacy. They did not want any data to be collected:

> *"[I took off my bracelet because] I went to a rave and I didn't want any data to be collected at that time, any type of data, about me but also about others" (P11)*

Other participants talked about specific data. For example, P29 and P21 thought removing the bracelet was useful to "*hide their location*". Similarly, P8 said removal during specific activities, such as the consumption of alcohol, could protect their related privacy:

> *"Yes [removal can help to protect privacy], at certain times: for example, when consuming alcohol, tobacco, drugs or during sexual activity." (P8)*

They believed if their data is not being collected, their privacy is preserved. Therefore, they believe removing their bracelet protects their privacy.

4.1.3.1.2 Bracelet removal is useful if it is removed for a long time

P22 and P1 gave similar interesting opinions. P1 remarked that removing the bracelet only for a short period of time is not enough to prevent inference from happening, especially if it is worn the rest of the time.

> *"[Taking off the bracelet] all day or for a long time, yes [it can be useful to protect my privacy], otherwise removing it for 15 min for example is useless, especially on 8 hours of use. For the rest of the time it is worn and it is more than enough to compromise our privacy" (P1)*

P22 claimed you would have to be very meticulous to avert inference:

> *"No [removing the bracelet does not protect privacy], unless you are very rigorous in interfering with the data because removing it once does not prevent inferences, especially in the long term because our habits will still be captured" (P22)*

Contrary to the previous section, here participants thought that interruption in data collection is not enough to protect their privacy.

4.1.3.1.3 Removal is useless to protect privacy

Regarding the participants who did not think removing the bracelet can help protect their privacy, the majority thought so because they did not consider the data to be sensitive. However, one of the participants, P25, said something interesting:

> *"There are other ways to get information about me, and removing the bracelet won't change that; it won't protect my privacy" (P25)*

This idea is in line with the feeling of not being able to escape data collection, that we mentioned earlier in the *worried* section. P25 implies that, among all the ways their data can be collected, the bracelet is a minor one. They believe that similar information can be inferred even if the bracelet is removed, using different means. Therefore, they do not think removing the bracelet can help them protect their privacy.

The majority thought removing the bracelet was useful to protect their privacy because it stops to collect data. When data is not collected, privacy is preserved. However, two participants argued that it is not true. You would have to remove the bracelet very consciously to prevent data collection. Finally, those who perceived bracelet removal as a useless strategy thought so either because their did not

perceived the Fitbit data as sensitive, or because they did not consider the Fitbit as a major source of data collection. Therefore, they believed that their data would still be collected via other means, despite removing their bracelet.

4.1.3.2 Disabling sensors

We asked participants about sensors perception. We wondered if they believed disabling sensors actually disabled them, and if this would change if the design of the button was different. We then wanted to know if they thought disabling sensors could be useful to protect their privacy. Finally, we asked what bracelet sensor they would disable if they could and why.

4.1.3.2.1 Beliefs about disabling sensors

Are sensors really disabled?	Number of participants	Participants' Number
Don't know/Have doubts	7	7, 8, 13, 16, 17, 21, 28
Yes	5	1, 11, 22, 29, 32
No	3	18, 23, 25
Yes, but the information can be obtained differently	3	6, 15, 19
It depends	1	26

Table 4 – Are sensors really disabled: distribution of participants' belief about disabling sensors

We asked participants if they believed disabling the sensors on their phone, or on another device, was *really* disabling them, or if they were still running somehow. Only five participants believed it was really disabled. Most were unsure whether it was really disabled.

> *"I don't have any knowledge of it, maybe it still works." (P17)*

> *"I hope it is, but I don't think it's completely disabled." (P28)*

Some participants had different beliefs about the impacts of disabling sensors. For instance, P25 implied that it is only effective for some stakeholders.

> *"I think [that by disabling the sensors] you make yourself invisible to some stakeholders, but it's still running in the background" (P25)*

P26 thought that it depends on how you would disable the sensors.

> *"In short-cut settings it's not really disabled. But directly in the settings it is really disabled" (P26)*

Three participants (P6, P15, P19) stated that, although the sensors were really disabled, it would not prevent data collection.

> *"Yes [it is useful to disable sensors to protect privacy], but the information can be obtained differently" (P6)*

P15 gave the example of location data, and said that despite the GPS being disabled, location could still be obtained.

> *"Geolocation is never disabled on the phone by triangulation, so no need for the GPS to be enabled [to locate me]." (P15)*

P19 explained that it is difficult to know what sensors are really disabled because there exist other means to collect data. It makes it harder to know where the data have been collected. It comes back to this idea that they cannot escape data collection.

> *"It's difficult to say [if sensors are really disabled], because even if we disable something, there is another way to get the same information" (P19)*

These beliefs and doubts show that they lack knowledge of how sensors work on phone or other devices.

4.1.3.2.2 Design preference for disabling sensors

We asked participants if their perception of the effectiveness of the button to enable and disable sensors on the bracelet would change if it was *physical* (physical button or switch) rather than *software* (on a touchscreen). The vast majority (14 participants) would not perceive the button differently if it was *physical* instead of *software*. Only two participants, P18 and P19, believe they would trust a physical control more if they had the guarantee that it cuts the electrical circuit of the sensor.

> *"No [my perception would not change if the button was different], unless it's a kill switch, circuit breaker." (P18)*

> *"Yes [my perception would change if the button was different], if it's like a plug that you unplug you have more confidence." (P19)*

Moreover, participants mentioned advantages and disadvantages of a physical switch. The first downside is that a it can be activated or deactivated accidentally and is less convenient because users are not used to it.

> *"[I'd prefer a button] on the app or website it's more convenient, a physical button can be activated by chance or by mistake." (P6)*

> *"[I'd prefer a button] on the application instead. It's more convenient. Physically it can be reactivated by accident." (P13)*

Another drawback given by P32 is the aesthetic of such switch.

> *"If I'm convinced that it's the same result [between the 2 types of buttons], then I don't want an extra button, it's not aesthetic." (P32)*

Regarding the advantages of this kind of button, P21 stated it can act as a visual reminder that sensors are enabled.

> *"I think seeing the physical button can raise awareness among people, it's a visual reminder that we have sensors on." (P21)*

P16 argued that it is useful if you need to use the switch often.

> *"If I deactivate / reactivate it often yes it can be convenient." (P16)*

Overall, participants would not perceive the effectiveness of the button differently if it was physical. They tend to prefer to stay on the status quo, the software button, because this is what they use every day and have no problem with it.

4.1.3.2.3 Perceived usefulness of disabling sensors

We asked the participants if they believed disabling the sensors on their bracelet could be useful to protect their privacy.

Is disabling sensors useful to protect privacy?	Number of participants	Participants' Number
Useful	4	1, 13, 16, 29
Useless	9	7, 15, 17, 18, 19, 21, 23, 25, 28
Useless on the long term	2	6, 22
Other	1	11
No answer	*3*	*8, 26, 32*

Table 5 – Is disabling sensors useful to protect privacy: distribution of participants' belief about usefulness of disabling sensors to protect privacy

Since the majority were not sure whether sensors were really disabled, it made sense to expect participants to say they did not perceive disabling sensors as a useful privacy protection mechanism because of that. However, none of the participants gave this answer. They had different reasons to believe it is useless to protect their privacy. The majority thought so because they do not consider Fitbit data as sensitive.

> *"It's useful for sensitive data, but not for Fitbit data" (P28)*

P22 believed disabling sensors was useless to protect their privacy if it was disabled only temporarily. They thought the same for bracelet removal. P6's idea was similar. Both said that on the long term, the same data will be inferred.

> *"Even if [the sensors] are disabled for a while, there is still data from the rest of the time with the bracelet. And we can deduce a lot even with a little bit of data." (P6)*

> *"Not very useful [to disable sensors to protect my privacy] because even if we disable them temporarily, in the long term they can still manage to infer a lot of information based on when it is activated." (P22)*

Regarding P21 and P23, they believe it will not protect their privacy because the same data can be inferred differently. Again, this is the idea of being able to escape data collection. In other words, they believe that even if they disable Fitbit's sensors, other means (such as smartphones) can be used to collect their data.

> *"If someone wants to control you, they can do it even if you turn off some sensors." (P21)*

> *"On a daily basis anyway we have the smartphone, so if we disable the watch without disabling the phone, it boils down to the same thing" (P23)*

Overall, they mostly do not believe disabling sensors can be useful to protect their privacy. Either because they do not consider Fitbit data to be part of their privacy, or because they do not think it will prevent their data from being collected. They believe that even if the data is not collected through the Fitbit's sensors, it is gathered through other means, such as smartphones.

4.1.3.2.4 Reasons for disabling sensors

Sensors participants would disable	Number of participants
None	5
GPS	9
Heart Rate	1
Blood Oxygenation Sensor	2

Table 6 – List of sensors users are willing to disable

During the interviews, we asked them to explain why they would disable -or not- certain sensors on their tracker. The vast majority said they would disable the *GPS* for two reasons. First, because of privacy concerns.

> *"[I will deactivate the GPS because] it is too intrusive to know where I am constantly." (P1)*

> *"[I would disable the GPS] so I wouldn't be located all the time. I don't want to be located all the time." (P13)*

Second, because they do not need it.

> *"[I would turn off the GPS] all the time because I don't need it, and nobody else needs to know my location at all times" (P7)*

> *"[I would disable location because it is] not useful for a bracelet like this. I know my routes; I don't need it." (P16)*

Regarding *heart rate*, only P22 considered it as very sensitive. Therefore, they wanted to disable it.

> *"Heart rate reflects emotional state, stress, anxiety, or health concerns or drug use. For me it's very personal." (P22)*

Two participants said they would disable the blood oxygenation sensor because they do not need it. However, there is no such sensor in their Fitbit bracelet. It shows they lack understanding of what data the bracelet actually collects.

Users would disable sensors that they do not need, and that collect data perceived as sensitive. Since fitness data are not perceived as sensitive, and are considered useful, users would not disable them. The following quote from P28 illustrates this well.

> *"[Heart rate and number of steps] are useful for me [so I won't disable them]. But the rest, yes I will disable if it could help [to protect my privacy]" (P28)*

4.1.3.3 Other privacy protection strategies

In this section we present other privacy protection strategies users mentioned using in the interviews.

4.1.3.3.1 Thinking before posting anything

One strategy the participants mentioned was to think before posting anything. It means that they evaluate if the content they are about to publish will harm their privacy. If it does, they refrain from posting it.

> *"I control what I put on social media before, so I protect my privacy beforehand." (P8)*

> *"I think before I post anything, so I don't have anything compromising on the internet." (P16)*

They use this strategy for data they perceive as sensitive ("*compromising*"). However, the majority of our participants consider fitness data as unsensitive. Therefore, it is very unlikely that they do this with when sharing their Fitbit data.

4.1.3.3.2 Give as little personal information as possible

Another strategy is to give as little personal information as possible. They hope that retaining as much personal information as possible will protect their privacy.

> *"I don't want other people to know about me. On social media, I give minimal information about myself. I try to give the minimum to companies too to protect myself." (P13)*

> *"I tried to put as little information as possible [on the Fitbit app]." (P26)*

P15 mentioned lying about the information they give:

> *"Sometimes I voluntarily give false information to "cover my tracks": false address, false last name" (P15)*

However, this strategy is only effective for data manually entered in applications and cannot be used for sensors' data. The extend of this strategy is very limited for fitness trackers.

4.1.3.3.3 Permission management

Three participants mentioned managing the app permissions on their phone.

> *"I manage the application permissions on my phone." (P15)*

P13 is particularly careful about their location:

> *"If it's not necessary, in general I don't give access to my location for apps." (P13)*

P26, on the other hand, cares about all type of permissions and prefer to authorize them only when it's needed:

> *"By default, I refuse all accesses, and eventually I re-authorize later if needed." (P26)*

Permission management is way to have some control over personal data. Participants can choose, to a certain degree, what app can access what data type. However, with the Fitbit model they have, they must enable internet connection, and for Android users, GPS too (Fitbit, 2021a, 2021b). Again, this strategy is very limited for this fitness tracker.

4.1.4 Reaction to inference results

During the interview, we told to our participants that we have been able to infer their religion and personality. This is not completely true, as it was done by another team and only Islam and neuroticism had been precisely inferred so far. The main goal of this statement was to create a reaction and capture it.

Ten participants reacted to the statement, either right after it, or during a discussion at the end of the interview. The three main reactions were curiosity, surprise, and change in their privacy perception.

A summary of their answers to the survey related to religion and personality inference can be found in appendix E.

4.1.4.1 Surprise

A few participants expressed a clear surprised about this statement.

> *"Is that really true?!" (P15)*
>
> *"I'm surprised at the results you managed to get." (P21)*

P11 was surprised but doubtful about the accuracy of our results:

> *"Actually what I find interesting, you said that you were able to deduce our religion […] but is that really... In the sense that I have Muslim friends and they drink alcohol, you know what I mean?" (P11)*

Their reaction was expected as they all believed the inference of religion and personality was moderately accurate at best.

4.1.4.2 Curiosity

The other reaction was curiosity. Four participants wanted to know if we really succeeded in inferring their personal data.

> *"I would like to know if you were able to guess the religion and everything" (P8)*
>
> *"Oh yes, I have a question, so it's true that you were able to see the consumption of alcohol and drugs?" (P15)*
>
> *"And what were the results for me [about personality]? […] And for religion, what kind of religion am I?" (P21)*
>
> *"Were you able to deduce things from the data?" (P23)*

Apart from P21, in the survey they all said to be slightly worried or not worried if their religion and personality were inferred. Which means that P21 may have asked these questions out of concerns, while the other asked out of simple curiosity.

4.1.4.3 Change in privacy perception

Apart from how we inferred the data, participants P11 and P17 were also curious about how their Fitbit data was managed. As they both stated not being worried in the survey, their questions can be interpretated as a raise in their privacy concerns.

> *"I have another question... What do they do with our data? Honestly I haven't read the terms and conditions and such..." (P11)*

> *" And what about data protection? Since I don't know much about it..." (P17)*

Two participants reflected on their behavior. P1, who is worried about the inference of their personality, claimed that it is good they did not wear the Fitbit tracker all the time.

> *"So I did the right thing not to sleep with it in this case" (P1)*

P28, who in the survey said to be worried and thought inference was moderately accurate for personality and not accurate for religion, is thinking of changing his behavior so to not send his data to the Fitbit servers.

> *"Maybe I'll synchronize less and less... Maybe look at the end of the day to see how much I've done, maybe write it down if I need to. [...] because now that I know that it's going to the servers, well... I might as well look at it myself and write it down, rather than my information be sent somewhere else" (P28)*

Finally, P32 perceived risks increased after the statement. In the survey, P32 thought inference was slightly precise at best. However, at the end of the interview he mentioned "*inference of personal data*" as a privacy risks emanating from the Fitbit servers.

> *"The server [presents privacy risks]. In relation to what you just said to me... Inference of personal data." (P32)*

Overall, their reaction seemed to be in line and not contradictory with their answers to the survey. Some participants seem to have reconsider the likelihood of inference. Although it is difficult to claim it based on quotation only.

4.2 Perceived utility

4.2.1 Wearing attitude

We asked our participants if they were still wearing the bracelet at the time of the interview. Seven stopped wearing it, seven continued to wear it but only for sports, and five are still wearing it all day.

Wearing habit	Nb of participants	Participants' numbers
Stopped wearing	7	1, 6, 7, 16, 21, 22, 23
Only for sport	7	8, 11, 13, 19, 26, 28, 29
Wearing all day	5	15, 17, 18, 25, 32

Table 7 – Wearing habits : distribution of participants' wearing habit at the time of the interview

4.2.1.1 Abandonment

Different reasons for the abandonment of the device were discovered. The main one is their lack of interest in the device. For P1 and P6 the data were interesting at first, but they lost interest over time.

> *"Yes [the way I wear the bracelet has changed], I used to wear it less and less, [...] now I don't wear it anymore." (P1)*

> *"It was interesting to see this type of data over a short period of time, but now it doesn't bring me anything anymore, I don't really find it interesting anymore." (P6)*

P21 lost interest because they found the bracelet not precise enough, therefore useless.

> *"I don't wear it anymore because I don't really need it. […] I didn't like it too much because it's not precise enough." (P21)*

For P7 and P23, they stopped wearing it because it was not practical, nor useful for sports:

> *"No [I don't wear it anymore] because it's not practical to wear every day and I don't use it for sports either because I don't see the use" (P7)*

> *"I don't wear it anymore because it's uncomfortable and not practical for sports […]. And apart from sport I don't see the point of the bracelet." (P23)*

Only two participants, P6 and P22, stopped wearing the bracelet for privacy reasons. P6 explicitly expressed worries with regards to data collection.

> *"The technology and things like the tracker are collecting [data] continuously, it can be used to learn about my habits and that worries me. That's why now I don't wear [the Fitbit] anymore." (P6)*

The main reason why P22 stopped wearing it is to protect their privacy.

> *"Even if the app had been useful for sports, I would not have kept it because of the daily data collected: heart rate, walking speed, location." (P22)*

The first reason for abandonment is the lack of usefulness of the device. Only a minority stopped wearing it due to of privacy concerns.

4.2.1.2 Wear it only for sports

Regarding those who still wear the bracelet, the majority wears it only during sport activities. One of the explanations is finding the bracelet "*not practical*". This feeling, that is one of the reasons for abandonment too, is shared by participants P19 and P29.

> *"I used to wear it less because of the allergies [it gave me] and it bothered me for manual work [in architecture at EPFL: using plaster, etc.]. Now I only wear it when I exercise." (P29)*

> *"[I wear the Fitbit] not every day because it's not comfortable when working on the computer. […] Now I only wear it when I work out to get data about it." (P19)*

Other reasons stated by P13 are the design of the bracelet and privacy. P13 found balance between benefits and privacy concerns related to location, by deciding to wear the Fitbit only when exercising.

> *"I prefer my current watch aesthetically, and I don't like the fact that the app knows where I am all the time […]. I use [the Fitbit] just for working out." (P13)*

They wore the bracelet only for sports because it was not practical to wear it all the time, or for privacy reasons.

4.2.1.3 Wearing all day

Five participants were still wearing the bracelet all day when we interviewed them. However, three of them had a lower interest in the Fitbit data. P15 said they looked at the data less frequently.

> *"Yes [I still wear the bracelet] but I look at the data less and less." (P15)*

Similarly, P32 and P25 only used the Fitbit as a watch.

> *"I rarely use the app, I only go there to set the time. I use the Fitbit as a watch." (P32)*

> *"I used to wear it as a watch [before I lost the charger and couldn't wear it anymore]" (P25)*

Among the participants who are still interested in the tracker, they really enjoy having the its data.

> *"I rarely took [the Fitbit] off [during the experiment] because I wanted the complete data for the day, km, calories, all that, and for the night" (P17)*

> *"The more details the better. I like stats, I'm interested in seeing that, comparing my different activity levels." (P32)*

P17 said the bracelet motivated them to practice more sports.

> *"Since wearing the bracelet, it has motivated me [to exercise]" (P17)*

The majority of participants who were still wearing the bracelet all the time used it mostly as a watch. They accepted having their data collected despite using the tracker only as a watch. A minority enjoyed having access to their fitness data. They appreciated doing statistics with it, and found it motivating.

4.2.1.4 Wore the bracelet more for experiment

Five participants (P6, P13, P18, P19, P29) admitted wearing the bracelet as much as possible during the experiment.

> *"For the experiment I wanted to wear it all the time, but since it is over I no longer wear it, precisely to protect my privacy." (P6)*

> *"For the study I wore [the Fitbit watch] all the time, [...] I made the effort to wear it as much as possible." (P18)*

> *"I used to force myself to wear it all the time during the experiment, now I only wear it when I work out to get data about it." (P19)*

> *"During the day I wore it as much as possible to play the game and bring you the maximum of information" (P29)*

They wanted to please the researchers by having a maximum of personal data being collected. It implies that they may have worn the bracelet differently for the experiment, compared to how they would have normally. This is a potential bias in our experiment.

4.2.2 Types of Fitbit usage

4.2.2.1 Website usage

We asked participants if they used the website, and why. Only P18 used the website because they preferred the level of detail it offered. P18 really enjoyed having advanced statistics.

> *"I like the level of detail on the website. The graphs are nice, the trends too" (P18)*

The other participants did not use the website because they find it more convenient to check their data on their phone or fitness tracker.

> *"The app [...] is more convenient because you are already logged in." (P13)*

> *"I didn't use [the website] because it's less convenient and easy to access than the application" (P19)*

They also were satisfied with the level of detail offered by the smartphone app.

> *"I know the site has more stuff, but I didn't see the use." (P11)*

> *"The smartphone provided all the info I needed, I didn't feel the need to look elsewhere." (P32)*

P18 explained they were not very interested in the data, so they never needed to check them differently.

> *"I'm not very interested in the bracelet data, so I have no interest in seeing it in more detail on the site." (P18)*

Except for P18, the participants were not interested enough in the data to check them in more details on the website. They were satisfied with the level of precision available on the Fitbit app.

4.2.2.2 Third-party applications

We wanted to know if participants used third party apps with their Fitbit. Whether they did connect other applications with their Fitbit account or not, we were interested to learn if they had privacy concerns related to it.

We found that only 3 participants, P15, P17 and P18, linked a third-party app to their Fitbit account. P15 authorized their insurance to access their steps count. They did so to have a discount on their insurance policy.

> *"Helsana insurance is connected to my tracker and collects my steps. […] At the end I can have a reduction [on my insurance policy]." (P15)*

P15 had no privacy concerns doing so because they trust the Swiss law LPD (Loi fédérale sur la Protection des Données).

> *"[My authorized my insurance company to access my step count because] I trust insurance companies more because they are more protective of data, the Swiss law "LPD" protects data." (P15)*

P17 connected with Strava because their friends use it too.

> *"A lot of people around me use Strava, and it's convenient to use Strava with Fitbit" (P17)*

They are not worried about linking Strava to their Fitbit account because they know other people who use it too.

> *"No [I'm not worried], lots of people use [Strava], so I think it's low risk" (P17)*

Finally, P18 used Fatsecret with Fitbit in order to count calories better. They used this application because it was more convenient to record the Swiss food they ate. They expressed no privacy concerns related to it.

> *"Because the application corresponds better to my needs, it is more useful for registering Swiss food." (P18)*

These 3 participants benefit from using the third-party app they authorized. It is convenient, they can save money or increase their social experience with friends. Regarding the other participants, only P25 said to have privacy concerns using third-party app. They agreed to link their Fitbit account to trusted entities only.

> *"I agree to share for Unil and science but not the others." (P25)*

The rest of the participants either were not interested in linking their Fitbit account to third-parties, or wanted to do it but did not have any relevant application to do so.

4.2.3 Data minimization

4.2.3.1 Steps intervals preference

In the survey, we presented participants different intervals (100, 500, 1000, 2000) as alternative to the total step count. We asked them to compare having an interval with having the precise step number. The interval could be evaluated from "as useful as the precise number" to "useless compared to the precise number". The results are summarized in the table below. We asked them to develop their answers in the interviews.

Evaluated as / Intervals	100	500	1000	2000
As useful	5	0	0	0
Slightly less useful	7	6	1	0
Less useful	1	4	5	1
Not useful	5	8	12	17

Table 8 – Intervals evaluation: distribution of how participants rated each interval

4.2.3.1.1 100 and 500 steps interval

The 100 and 500 steps intervals were rated "as useful" or "slightly less useful" for different reasons. For P7 and P8, it was because they considered it precise enough.

> *"[intervals of] 500 or 100, that's enough information, that's precise enough" (P7)*

> *"[With the interval of] 100 we know quite precisely how many steps we have walked during the day." (P8)*

P22 even argue that 100 steps is such a small number that they would not feel like the physical difference of walking 100 steps more or less.

> *"100 steps is very close to the real number, I don't feel the physical difference." (P22)*

P19 state that because the bracelet is not accurate, an interval of 100 is precise enough.

> *"The 100 steps interval is the best [among the 4 intervals] because it gives a good idea and the bracelet is not very precise anyway." (P19)*

Another argument in favor of the smallest intervals is that it motivates the user to do more steps by wanting to reach the next interval.

> *"An interval may motivate to exercise more to reach the next interval, but this is not the case if the interval is too large." (P21)*

The smallest intervals are considered useful when perceived as precise enough for this kind of bracelet. Moreover, they can act as motivation to increase physical activity by being an incentive to reach the next interval.

4.2.3.1.2 1000 and 2000 intervals

Larger intervals were unanimously considered as not useful compared to the precise number. Participants regarded them as useless because there are not precise enough.

> *"The large intervals (1000 and 2000) are not precise enough, it does not give any information in the end." (P7)*

> *"The reason I buy a pedometer watch is so that it accurately calculates my steps and not other movements. So if the interval is too large, it is not useful." (P28)*

P11 would have the impression that the tracker is less effective if it had larger intervals.

> *"[A large interval,] it's irrelevant because it's really too imprecise, I would have the impression that it counts the steps less accurately" (P11)*

P22 explain that they are particularly useless if you try to reach a goal because it is difficult to know if your reached it or not.

> *"A long interval is impractical, especially if you set a goal, [...] you don't really know what you've done." (P22)*

Larger intervals are disliked because they do not give enough information. Therefore, it makes it difficult to set fitness goals, and the tracker can appear as less effective.

4.2.3.1.3 All intervals are useless compared to precise number

A few participants disliked all intervals. For them, only a precise number is useful.

> *"I want the precise result at the end of the day, I would feel like I was missing something if it was an interval, because I would have a less good overall idea of my steps." (P17)*

> *"An interval is less precise so there is less information and less interest." (P23)*

They also want a precise number to be able to compare and precisely measure their results.

> *"You can't compare intervals, that doesn't allow you to draw trends. I prefer a margin of error and a specific number." (P15)*

> *"I would say I want the precise daily number. With an interval we cannot calculate an average. I need to calculate it because I find it relevant and satisfying." (P25)*

> *"I would feel like I was losing information, because every day we would have the same interval and that's useless, especially for the larger intervals, we can't compare our results." (P32)*

P13 implies they would not be interested in a device displaying intervals instead of a precise number.

> *"I find it a shame to pay for a device that is not precise." (P13)*

They dislike all intervals because they cannot be compared and used to draw trends; they do not allow precise analysis of their performance.

4.2.3.2 Precision need

We asked participants what level of precision they used and if they needed the bracelet to be precise. We had 2 clearly opposite types of answer. Participants either stated to only need a general idea of their results, or they wanted the device to be as precise as possible. The level of detail they used to check they step count is summarized in the table 9 below.

Level of detail used for steps	Number of participants	Participants' Number
Total steps per day	14	6, 7, 8, 11, 13, 15, 16, 18, 21, 22, 23, 25, 26, 29
More precise than Total Steps Per Day (per hour, …)	4	1, 17, 19, 32
Other: Activity duration and number of kilometers travelled	1	26

Table 9 – Level of detail used: distribution of participants' preferred level of detail to check their step count

4.2.3.2.1 A general idea is precise enough

Many participants reported not needing the bracelet to be highly accurate, and were satisfied with just having a general idea of their results. Most claimed not needing detailed information about their steps count and were satisfied with using just the total steps count. Either because they had no need to see it in more detail, or because they were interested in comparing their performance, thus they only needed their total steps number.

> *"The total number is enough for me, no need for the detail throughout the day." (P8)*

> *"In general, I look at the total number, not precisely but to have a point of reference and compare the days to see if I walked more or less" (P22)*

> *"I mostly look at the total number of steps per day, per weekend and per week. I like to compare the days with each other. I have little interest in more detail." (P26)*

P11 and P16 expressed being satisfied with the current precision of the fitness tracker.

> *"No, it's not important [that the bracelet is accurate] and I don't need it, a margin of error of +-150 steps is enough." (P11)*

> *"The current precision is fine with me, it was out of curiosity more than need. Having a general idea is enough." (P16)*

P15 and P25 explained they were also satisfied with a general idea because they believed the tracker was not precise.

> *"I'm just interested in the trend, because I don't know what the margin of error of the bracelet is." (P15)*

> *"Although it is not perfect because it counts arm movements as steps, [the accuracy is] sufficient." (P25)*

Finally, they are happy with having just a general idea because they do not need the tracker for serious reasons, like medical problems or measurements of "athlete" performance.

> *"For me, not really [important that the bracelet is precise]. Getting a general idea is enough, it's interesting to see once and I don't need it for medical reasons." (P6)*

> *"Accurate to 1-2 beats/min, what do I care, I don't have any heart problems." (P15)*

> *"No [the bracelet does not need to be precise], if it is for leisure and if it is mainly used as a motivation and to give an overall idea, it is not important that it is very precise" (P21)*

> *"No for others like me who just work out at home, it's not that important [that the bracelet is accurate]." (P29)*

The current precision of the device is good enough for most of the participants. They mostly check their results to have a general idea of their daily performance, and have no desire for more detail.

4.2.3.2.2 As precise as possible

One reason some participants wanted the tracker to be as precise as possible, is because they like stats and are interested in their detailed results.

> *"I look precisely at the detail throughout the day and the track record at the end of the day, heart rate, water, steps, sleep. […] I don't need it to be accurate to the step, but I would like to have as much precision as possible." (P17)*

> *"I want as much detail as possible to see the different levels of activity." (P19)*

> *"The more details the better. I like stats, I'm interested in seeing that, comparing my different activity levels. […] I want optimal data quality and accuracy." (P32)*

However, they often use only certain types of data. Therefore, they need the highest accuracy for these and not the others.

> *"Sleep and heart rate, yes [that must be precise] because that's data that interests me. For the number of steps, I don't mind if there is a small gap." (P8)*

> *"I really want specific information. Because I need to know, for my heart rate, if I'm giving my all. For the number of steps, it is fine if there are a few steps errors." (P13)*

> *"Steps and calories, I need it to be accurate because that's what I use the most." (P18)*

> *"Yes [it is important that the bracelet is accurate], to have statistics on the data that interests me, number of steps and heart rate" (P26)*

Participants want the data that they are interested in to be precise. This is particularly true for those who like meticulously tracking their performance.

5 Discussion

5.1 Privacy perception

5.1.1 General privacy concerns

Similarly to what we found in the literature review, fitness tracker users were not worried about their fitness data (sleep, heart rate and step count). They did not it perceive as sensitive and they did not believe negative consequences were likely. However, despite the Fitbit tracker not having a GPS, they thought it contained one and were clearly worried about their location being disclosed.

Regarding the *personal data* (cf. Table 2), our results are similar to (Gabriele and Chiasson, 2020; Zimmer et al., 2020)'s in that our participants perceived some types of data as more acceptable to share than others. They were worried about the data they were shy about, and the data that could lead to consequences like family disagreements, profiling, manipulation, and discrimination. They seemed to be more worried about psychological data than physiological data, and information perceived as not related to sport. Another source of worry is the pervasiveness of the tracker. Despite it being an expected feature (Randriambelonoro et al., 2017), it increased the privacy concerns of a few users. Finally, only a minority of our participants were "data protectors" (Burbach et al., 2019), as only a minority stopped wearing the device for privacy reasons.

At the time of the interview, the participants were not obligated to wear the fitness tracker anymore. The majority either stopped wearing the tracker or they wear it only wear doing physical activities. The first reason for abandonment of the device is the lack of usefulness of the device. Only a minority stopped wearing it due to of privacy concerns. For those who wear it only for sports, their reason is that it was not practical to wear it all the time, or for privacy reasons. Finally, for those who were still wearing the bracelet all the time, they used it mostly as a watch. A minority enjoyed having access to their fitness data. Therefore, we can see a clear decrease in interest in the device after the experiment ended. This decline in use is well observed among buyers of fitness tracker (Finkelstein et al., 2016; "PWC: The Wearable Life 2.0: Connected living in a wearable world," 2016), so this is something we expected.

5.1.2 Inference knowledge and belief

Several findings show that fitness tracker users lack knowledge of privacy risks related to inference. First, they had a hard time giving concrete example of negative consequence and remained quite vague. This result is similar to our literature review in which we found that even the most privacy concerned users have difficulties giving concrete examples and explanations of potential privacy risks and of the resulting negative consequences (Aktypi et al., 2017; Gabriele and Chiasson, 2020; Lowens et al., 2017; Zimmer et al., 2020). Second, they do not know how and where their data is processed, and they are concerned about it, which related to (Lowens et al., 2017). Third, the majority believed the inference was possible for physiological data only, or when other data was used. No participants mentioned the

inference possibility using raw sensor data. Rather, some participants explained the data they manually entered in the app could be inferred. This relates to (Rader and Slaker, 2017) who found that users conceptualize the collection of only 3 types of data: the ones they manually entered (age, weight, etc.), the ones measured by the tracker (steps, heart rate, etc.), and the ones calculated by the device (calories burned, activity level, etc.).

We identified 1 reason that contributes to participants' lack of concerns about inference and their perceived unlikelihood of negative consequences, in the context of fitness tracker. The reason is that they live in a "free country". The potential consequences of disclosing personal information such religion, socio-economic status, sexual activity, or sexual orientation, are for most people "mild" and not life-threatening. However, this belief can be perceived as naive. Even if the users do not face consequences from the government, there are still companies that could use their information against them. The consequences could be job discrimination, insurance discrimination, or forms of manipulation to sell product to the users.

Their reactions to the statement about successful inferences were in line with their answers to the survey. As they believed inference of religion and personality was not possible, they were surprised that we said we succeeded. Some were curious to know how we did it and what we inferred about them. Finally, some participants' perception of inference risks seem to have changed. Their privacy concerns seem to have increased and one participant is thinking of changing their behavior in order to protect their privacy. It shows that when users are faced with proof that inference successfully happen, their concerns and behavior may change. However, it is difficult to measure the change based on quotes only. It would require quantitative data measuring their perception of the privacy risk and concerns before and after the interviews.

5.1.3 Privacy protection strategies

Regarding the actions users were ready to take to protect their privacy, the majority considered bracelet removal useful to protect their privacy. They believed if their data was not being collected, their privacy would be preserved. However, a minority of users believed the interruption of data collection was not enough to protect privacy. They claimed that on the long term, the same among of data can be inferred despite removing the bracelet from time to time.

On the other hand, they did not perceive disabling sensors useful to protect their privacy. Despite not knowing whether sensors were ever really disabled, they did not give it as the reason why they believed disabling sensors was useless to protect privacy. The reasons were either the data can be collected through other means, or the Fitbit data are not sensitive.

Users did not perceive disabling sensors as useful to protect their privacy, however they were willing to disable sensors that collect data they perceive as sensitive and sensors that collect data they do not need. For the first one, since they do not perceive fitness data as sensitive, they would not disable them for this reason. For the second one, many users were not interested in the 3 Fitbit data, therefore

they would be willing to disable the ones they do not use. It is important to note that the Fitbit they own, the Fitbit Inspire HR, only allows disabling heartrate recording (Fitbit, 2019). Therefore, they should be able to choose what sensors to activate to increase their privacy protection. This way, their privacy protection could be increased.

The few participants that believed sensors were never really disabled, also thought disabling sensors was *useless* to protect privacy. Apart from that, there is no correlation between the perceived usefulness of disabling sensors and the belief of them being really disabled.

When investigating an alternative button for disabling sensors, participants said having a *physical* button instead of a *software* one would not change their perception of it. The only advantage of having a physical button is that it can act as a visual reminder that sensors are on. But that be done in a more aesthetic way, with a small light on the device for instance. Users are satisfied with a software button, and we recommend keeping it that way. However, it is important to remember that they may have a status quo bias (SAMUELSON and ZECKHAUSER, 1988). They may have said to prefer the software button precisely because they do not want to change what they currently have and stick with the status quo.

They use other PPS on the internet: think before posting, give as little information as possible, permission management. But they have really limited impacts in the context of fitness tracker. For the first one, since they do not perceive fitness data as sensitive, they will continue to share it. For the second one, they can give as little information as possible while registering to the service, by filling only the required field and even lying in them, but they cannot change how sensors collect data. Finally, permission management is limited because the users have to enable an internet connection, and for Android users the GPS too, in order to synchronize their data (Fitbit, 2021b, 2021a). Therefore, if they want to benefit from the fitness tracker, they have to allow these accesses.

We identify one reason that prevent fitness tracker users from adopting privacy preserving strategies. The reason is their feeling and belief of being unable to “escape data collection”. They think other devices, such as their smartphone or personal computer, collect similar or more personal information than their fitness tracker. Therefore, they see the fitness tracker as a minor source for data collection. In this context, they claim removing the bracelet or disabling sensors will not protect their privacy, as it will not prevent their data from being collected.

5.2 Data minimization

The utility preference is clearly influenced by the participants' needs and goals. The majority of participants practiced little physical activity or had little interest in the Fitbit data. These users reported being satisfied with a lower precision, stating that "*just having a general idea is enough*". Based on (Burbach et al., 2019)'s types, they fall into the "*benefit maximizers*" category. The majority would not mind if the precision of the device was lower.

On the other hand, participants who practices sports regularly and more "seriously", or who were really interested in the Fitbit data, wanted as much precision as possible. We can categorize them as *"facts enthusiasts"* (Burbach et al., 2019) or *"self-quantifiers"* (Choe et al., 2014; Wolf, 2010). These users typically found the Fitbit not precise enough and would not accept a reduction in the precision of the device.

Regarding using intervals instead of the precise step number, this is overall not appreciated at all.
The main reasons are that they are too imprecise and that they cannot be used to do comparison or draw trends. Another reason for the dislike of intervals might be the status quo bias (SAMUELSON and ZECKHAUSER, 1988), because of which participants prefer to stick to what they already have and are averse to change. An alternative suggested by participant P15 could be to display a precise number, even if it is wrong, with its margin of error. It would be a good alternative for participants who are only interested in roughly following their activity level.
However, the 100-steps interval was mostly reported satisfactory by participants, stating that "it is precise enough" or arguing that the Fitbit was "not precise anyway". Also, it can motivate users to increase their physical activity by wanting to reach the next interval. Therefore, only a small interval could be suggested to users who are satisfied with just "*a general idea*" as an alternative to the precise step count. Apart from that, intervals appear to be a bad alternative to having the precise number.

Participants used only the features on the bracelet and on the app. They did not use the website at all. Therefore, they would be fine without it. About third party app, only a minority of participants linked one to their Fitbit account. Apart from the participants who cared a lot about precision and data quality, they would prefer that their data do not go to the Fitbit servers. Therefore, connectivity to Fitbit servers could be removed for these two features without harming the utility of the majority of users.

On top of that, choosing not the send data to servers and choosing which sensors to disable would help users feel in control of their data. This was one of the concerns identified in the results, which corresponds to (Lowens et al., 2017)'s results that users worry about the lack of control over their personal data.

5.3 Comparison with survey results

The interviews contributed greatly to put the survey's answers into context and to understand the reasoning of participants. Overall, the results of the survey and of the interviews are very similar.

However, we identified some differences. First, contrary to the survey, we found that bracelet removal is perceived as useful to protect privacy. Second, contrary to the survey too, we discovered that disabling sensors is perceived as useless to protect privacy. The difference is due to the fact that the interviews results are based on what the participants said, and not what they did, which was what the survey measures. Therefore, even if participants do not *think* disabling *fitness* sensors will protect their privacy, they were *doing* it for other sensors they perceived as sensitive (such as GPS) or for sensors they do not need.

5.4 Limitations and future work

We have identified some limits to this work. First, our participants may not represent "real" fitness trackers users as they were given the Fitbit for this experiment. Also, all our participants are young university students. It would have been better to conduct interviews with real fitness tracker users, with different age and background to represent the real fitness tracker users better.

Second, we distributed survey participants' into four groups according to their wearing behavior, and we aimed at interviewing an equal number of participants from each group. However, we did not succeed in achieving this goal. On top of that, by choosing to interview participants from the four groups, the distribution for the interview participants was different than the survey's. Therefore, the interviews are not representative of the survey participants, but rather of the different wearing behavior. This is not necessarily a bad thing, but it is important to that note of it when comparing the survey results with the interview results.

Third, there might be some bias in our experiment. The first one was mentioned by participants themselves. They claimed to have worn the Fitbit as much as possible during the experiment. They did so to give us "as much data as possible". Therefore, the way they wore the bracelet was to please the researchers and may have been different from their "normal" behavior. The second bias is related to one of the survey questions. We asked participants if they "thought sensors were really disabled or not". The question itself may imply that sensors were never really disabled. Because of the way we phrased the question, participants may have tried to anticipate a correct answer, and say something different from their belief. We should have asked the question in a more neutral manner, for instance "what do you think would happen if you pressed this button? [show button to disable sensor]".

Finally, some of the interviews were conducted remotely via Zoom. This is not an optimal way to interview participants. Many were at home, and they may have not felt comfortable to talk about certain data types (such as alcohol and drug consumption) because the people living with them might

have heard them. This may have limited the quality of their responses. On top of that, the audio recording of the Zoom interviews were not always of good quality, which made the transcription difficult. We would not recommend conducting remote interviews, except if there is no other options.

In the future, it would be interesting to do other face-to-face interviews with participants more representative of actual fitness tracker users. This way, we would have better insights of their actual belief, habits and needs in terms fitness tracker utility and privacy.

6 Conclusion

The qualitative analysis of the interviews enabled us to answer our research questions RQ1 and RQ2.

First, we have been able to understand better how users perceived privacy in the context of fitness trackers. We found that they did not consider fitness data (sleep pattern, heart rate, and step count) as sensitive. Regarding other *personal data* (cf. Table 2), they did not think they can be inferred solely based on the fitness data, except if they are body-related data. They believed that other information is needed in order to infer *personal data*. We also found that, when users are faced with proof that inference happened based on fitness data, their privacy concerns increase, and they are thinking of changing their behavior in order to protect heir privacy.

In terms of the privacy protection strategies (PPS) they could use, they were willing to remove their bracelet and to disable sensors. However, they were not convinced that it will help protect their privacy. Rather, they are willing to remove the bracelet and disable sensors because they do not need it. Other PPS they use on a daily basis, such as thinking before posting, giving as little information as possible, and permission management, have very limited impact on privacy protection in the context of fitness tracker.

Then, we were able to get a better understanding of users' perception of utility when using their fitness tracker. They only used the features available on the wrist band and on the smartphone app. Only one participant used the website, and very few participants linked third party applications to their Fitbit account.

We identified 2 types of fitness tracker users. The first one constituted the majority of the users. They are not very interested in the fitness data. Rather, they are satisfied with the current level of precision of their bracelet. The second one is composed of users who love having their detailed information, in order to precisely track their activity.

The first group is the easiest to protect in terms of privacy because they use only their bracelet and their phone and do not need high precision. Therefore, they are more likely to have their utility not diminished after data minimization actions would be taken. The second group, on the other hand, prioritize data quality and accuracy, therefore is it more difficult to reduce their amount of data collected without hurting their utility. However, both types never used the website, and rarely used third party application. Therefore, these are opportunities for data minimization that would be appropriate for both types of user.

7 References

Aktypi, A., Nurse, J.R.C., Goldsmith, M., 2017. Unwinding Ariadne's Identity Thread: Privacy Risks with Fitness Trackers and Online Social Networks, in: Proceedings of the 2017 on Multimedia Privacy and Security - MPS '17, MPS '17. ACM, Dallas, Texas, USA, pp. 1–11. https://doi.org/10.1145/3137616.3137617

Alex Hern, 2018. Fitness tracking app Strava gives away location of secret US army bases.

Alqhatani, A., Lipford, H.R., 2019. "There is nothing that I need to keep secret": Sharing Practices and Concerns of Wearable Fitness Data. Presented at the Fifteenth Symposium on Usable Privacy and Security ({SOUPS} 2019).

Antón, A.I., Earp, J.B., Young, J.D., 2010. How Internet Users' Privacy Concerns Have Evolved since 2002. IEEE Security & Privacy 8, 21–27. https://doi.org/10.1109/MSP.2010.38

Arnold, Z., Larose, D., Agu, E., 2015. Smartphone Inference of Alcohol Consumption Levels from Gait, in: 2015 International Conference on Healthcare Informatics. Presented at the 2015 International Conference on Healthcare Informatics, pp. 417–426. https://doi.org/10.1109/ICHI.2015.59

Braun, V., Clarke, V., 2006. Using thematic analysis in psychology. Qualitative Research in Psychology 3, 77–101. https://doi.org/10.1191/1478088706qp063oa

Burbach, L., Lidynia, C., Brauner, P., Ziefle, M., 2019. Data protectors, benefit maximizers, or facts enthusiasts: Identifying user profiles for life-logging technologies. Computers in Human Behavior 99, 9–21. https://doi.org/10.1016/j.chb.2019.05.004

Choe, E.K., Lee, N.B., Lee, B., Pratt, W., Kientz, J.A., 2014. Understanding quantified-selfers' practices in collecting and exploring personal data, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Presented at the CHI '14: CHI Conference on Human Factors in Computing Systems, ACM, Toronto Ontario Canada, pp. 1143–1152. https://doi.org/10.1145/2556288.2557372

Clarke, V., Braun, V., 2016. Thematic analysis. The Journal of Positive Psychology 12, 1–2. https://doi.org/10.1080/17439760.2016.1262613

Duncan, M., Murawski, B., Short, C.E., Rebar, A.L., Schoeppe, S., Alley, S., Vandelanotte, C., Kirwan, M., 2017. Activity Trackers Implement Different Behavior Change Techniques for Activity, Sleep, and Sedentary Behaviors. Interactive Journal of Medical Research 6, e13. https://doi.org/10.2196/ijmr.6685

Finkelstein, E.A., Haaland, B.A., Bilger, M., Sahasranaman, A., Sloan, R.A., Nang, E.E.K., Evenson, K.R., 2016. Effectiveness of activity trackers with and without incentives to increase physical activity (TRIPPA): a randomised controlled trial. Lancet Diabetes Endocrinol 4, 983–995. https://doi.org/10.1016/S2213-8587(16)30284-4

Fitbit, 2021a. Why is the Fitbit app prompting me to turn on location services? [WWW Document]. Fitbit | HELP. URL https://help.fitbit.com/articles/en_US/Help_article/2134.htm?Highlight=syncing (accessed 5.13.21).

Fitbit, 2021b. Why won't my Fitbit device sync? [WWW Document]. Fitbit | HELP. URL https://help.fitbit.com/articles/en_US/Help_article/1866.htm (accessed 5.14.21).

Fitbit, 2019. Fitbit Inspire HR User Manual.

Fritz, T., Huang, E.M., Murphy, G.C., Zimmermann, T., 2014. Persuasive technology in the real world: a study of long-term use of activity sensing devices for fitness, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '14. ACM, Toronto, Ontario, Canada, pp. 487–496. https://doi.org/10.1145/2556288.2557383

Gabriele, S., Chiasson, S., 2020. Understanding Fitness Tracker Users' Security and Privacy Knowledge, Attitudes and Behaviours, in: Proceedings of the 2020 CHI Conference on Human Factors in

Computing Systems. Presented at the CHI '20: CHI Conference on Human Factors in Computing Systems, ACM, Honolulu HI USA, pp. 1–12. https://doi.org/10.1145/3313831.3376651

Gao, Y., Li, H., Luo, Y., 2015. An empirical study of wearable technology acceptance in healthcare. Industr Mngmnt & Data Systems 115, 1704–1723. https://doi.org/10.1108/IMDS-03-2015-0087

Hallam, C., Zanella, G., 2016. Wearable device data and privacy: A study of perception and behavior. World Journal of Management 7.

Hassan, W.U., Hussain, S., Bates, A., 2018. Analysis of Privacy Protections in Fitness Tracking Social Networks -or- You can run, but can you hide?, in: Proceedings of the USENIX Symposium on Security. USENIX Association, Baltimore, MD, pp. 497–512.

Ihsan, Z., Furnham, A., 2018. The new technologies in personality assessment: A review. Consulting Psychology Journal: Practice and Research 70, 147. https://doi.org/10.1037/cpb0000106

Jensen, C., Potts, C., 2004. Privacy policies as decision-making tools: an evaluation of online privacy notices, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '04. Association for Computing Machinery, New York, NY, USA, pp. 471–478. https://doi.org/10.1145/985692.985752

Kazlouski, A., Marchioro, T., Manifavas, H., Markatos, E., 2021. I still See You! Inferring Fitness Data from Encrypted Traffic of Wearables:, in: Proceedings of the 14th International Joint Conference on Biomedical Engineering Systems and Technologies. Presented at the 14th International Conference on Health Informatics, SCITEPRESS - Science and Technology Publications, Online Streaming, --- Select a Country ---, pp. 369–376. https://doi.org/10.5220/0010233103690376

Kim, S., Thakur, A., Kim, J., 2020. Understanding Users' Perception Towards Automated Personality Detection with Group-specific Behavioral Data, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. Presented at the CHI '20: CHI Conference on Human Factors in Computing Systems, ACM, Honolulu HI USA, pp. 1–12. https://doi.org/10.1145/3313831.3376250

Kröger, J., 2019. Unexpected Inferences from Sensor Data: A Hidden Privacy Threat in the Internet of Things. pp. 147–159. https://doi.org/10.1007/978-3-030-15651-0_13

Lee, L., Lee, J., Egelman, S., Wagner, D., 2016. Information Disclosure Concerns in The Age of Wearable Computing, in: Proceedings 2016 Workshop on Usable Security. Presented at the Workshop on Usable Security, Internet Society, San Diego, CA. https://doi.org/10.14722/usec.2016.23006

Li, H., Wu, J., Gao, Y., Shi, Y., 2016. Examining individuals' adoption of healthcare wearable devices: An empirical study from privacy calculus perspective. International Journal of Medical Informatics 10.

Lowens, B., Motti, V.G., Caine, K., 2017. Wearable privacy: Skeletons in the data closet, in: IEEE Int. Conf. on Healthcare Informatics (ICHI). pp. 295–304.

Mendoza, F., Alonso, L., López, A., and Patricia Arias Cabarcos, D., 2018. Assessment of Fitness Tracker Security: A Case of Study. Multidisciplinary Digital Publishing Institute Proceedings 2, 1235. https://doi.org/10.3390/proceedings2191235

Mercer, K., Li, M., Giangregorio, L., Burns, C., Grindrod, K., 2016. Behavior Change Techniques Present in Wearable Activity Trackers: A Critical Analysis. JMIR mHealth uHealth 4, e40. https://doi.org/10.2196/mhealth.4461

Motti, V.G., Caine, K., 2015. Users' Privacy Concerns About Wearables, in: Brenner, M., Christin, N., Johnson, B., Rohloff, K. (Eds.), Financial Cryptography and Data Security, Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp. 231–244. https://doi.org/10.1007/978-3-662-48051-9_17

Niess, J., Woźniak, P.W., 2018. Supporting Meaningful Personal Fitness: the Tracker Goal Evolution Model, in: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, CHI '18. Association for Computing Machinery, New York, NY, USA, pp. 1–12. https://doi.org/10.1145/3173574.3173745

Norberg, P.A., Horne, D.R., Horne, D.A., 2007. The Privacy Paradox: Personal Information Disclosure Intentions versus Behaviors. Journal of Consumer Affairs 41, 100–126. https://doi.org/10.1111/j.1745-6606.2006.00070.x

Prasad, A., Sorber, J., Stablein, T., Anthony, D., Kotz, D., 2012. Understanding Sharing Preferences and Behavior for MHealth Devices, in: Proceedings of the 2012 ACM Workshop on Privacy in the Electronic Society, WPES '12. ACM, Raleigh, North Carolina, USA, pp. 117–128. https://doi.org/10.1145/2381966.2381983

PWC: The Wearable Life 2.0: Connected living in a wearable world, n.d.

Rader, E., Slaker, J., 2017. The importance of visibility for folk theories of sensor data. Presented at the Thirteenth Symposium on Usable Privacy and Security ({SOUPS} 2017), pp. 257–270.

Raij, A., Ghosh, A., Kumar, S., Srivastava, M., 2011. Privacy risks emerging from the adoption of innocuous wearable sensors in the mobile environment, in: Proc. of the Conf. on Human Factors in Computing Systems (CHI), CHI '11. Association for Computing Machinery, Vancouver, BC, Canada, pp. 11–20. https://doi.org/10.1145/1978942.1978945

Randriambelonoro, M., Chen, Y., Pu, P., 2017. Can Fitness Trackers Help Diabetic and Obese Users Make and Sustain Lifestyle Changes? Computer 50, 20–29. https://doi.org/10.1109/MC.2017.92

Ridgers, N.D., Timperio, A., Brown, H., Ball, K., Macfarlane, S., Lai, S.K., Richards, K., Mackintosh, K.A., McNarry, M.A., Foster, M., Salmon, J., 2018. Wearable Activity Tracker Use Among Australian Adolescents: Usability and Acceptability Study. JMIR Mhealth Uhealth 6, e86. https://doi.org/10.2196/mhealth.9199

Saksono, H., Castaneda-Sceppa, C., Hoffman, J., Seif El-Nasr, M., Morris, V., Parker, A.G., 2018. Family Health Promotion in Low-SES Neighborhoods: A Two-Month Study of Wearable Activity Tracking, in: Proc. of the Conf. on Human Factors in Computing Systems (CHI), CHI '18. ACM, Montreal QC, Canada, pp. 1–13. https://doi.org/10.1145/3173574.3173883

SAMUELSON, W., ZECKHAUSER, R., 1988. Status Quo Bias in Decision Making. Journal of Risk and Uncertainty 1, 7–59.

Schneegass, S., Poguntke, R., Machulla, T., 2019. Understanding the Impact of Information Representation on Willingness to Share Information, in: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, CHI '19. Association for Computing Machinery, New York, NY, USA, pp. 1–6. https://doi.org/10.1145/3290605.3300753

Spil, T., Sunyaev, A., Thiebes, S., Van Baalen, R., 2017. The Adoption of Wearables for a Healthy Lifestyle: Can Gamification Help? Presented at the Hawaii International Conference on System Sciences. https://doi.org/10.24251/HICSS.2017.437

Torre et al., 2016. Fitness trackers and wearable devices: how to prevent inference risks? Presented at the 11th EAI International Conference on Body Area Networks.

Vitak, J., Liao, Y., Kumar, P., Zimmer, M., Kritikos, K., 2018. Privacy Attitudes and Data Valuation Among Fitness Tracker Users, in: Chowdhury, G., McLeod, J., Gillet, V., Willett, P. (Eds.), Transforming Digital Worlds, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 229–239. https://doi.org/10.1007/978-3-319-78105-1_27

Weiss, G.M., Timko, J.L., Gallagher, C.M., Yoneda, K., Schreiber, A.J., 2016. Smartwatch-based activity recognition: A machine learning approach, in: IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI). Presented at the 2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI), IEEE, Las Vegas, NV, USA, pp. 426–429. https://doi.org/10.1109/BHI.2016.7455925

Wolf, G., 2010. The Data-Driven Life. The New York Times.

Yang, H., Yu, J., Zo, H., Choi, M., 2016. User acceptance of wearable devices: An extended perspective of perceived value. Telematics and Informatics 33, 256–269. https://doi.org/10.1016/j.tele.2015.08.007

Yang Liu, Z.L., 2018. aLeak: Privacy Leakage through Context-FreeWearable Side-Channel. Presented at the INFOCOM'18.

Zimmer, M., Kumar, P., Vitak, J., Liao, Y., Kritikos, K.C., 2020. 'There's nothing really they can do with this information': unpacking how users manage privacy boundaries for personal fitness information. Information, Communication & Society 23, 1020–1037. https://doi.org/10.1080/1369118X.2018.1543442

8 Appendix

A. Interview protocol

Interviews on Privacy and Utility Perceptions of Fitness Tracker Users - Protocol

November 2020

Research Team
Kavous Salehzadeh Niksirat - kavous.salehzadehniksirat@unil.ch
Kevin Huguenin - kevin.huguenin@unil.ch
Laura Mazuryk - laura.mazuryk@unil.ch
Lev Velykoivanenko - lev.velykoivanenko@unil.ch
Mauro Cherubini - mauro.cherubini@unil.ch
Noé Zufferey - noe.zufferey@unil.ch

Interviewers
Lev Velykoivanenko, Laura Mazuryk, Noé Zufferey (backup)

Setting
The interview will be held in the HEC-LABEX between [insert dates]. The questions are related to users' perceived utility and privacy risks with regard to their fitness trackers. The participants should be invited to talk freely and to explain why they think and feel the things that they do.

Instructions for the Interviewers
This protocol contains instructions and multiple sets of questions arranged into blocks. The interviews will be conducted in a semi-structured way. Hence, the interviewer should follow the protocol in a non-strict manner. The questions are intended as a guiding checklist. The interviewer needs to ascertain all the required information. Deviations from the questions and their ordering are possible, for example: the number and/or order of the questions may change, some questions may be discarded, and new questions may be added. The interviewer should, instead of collecting opinions and abstract ideas, attempt to get experiences and anecdotes to explain the answers. The participants should be invited to explain why they think or feel something. Interviewees should be given ample time to think about their answers instead of immediately being pressured with probing questions.

Research Questions

RQ1. Privacy-Preserving Strategies (PPSs): Which PPSs would users be willing to adopt and why?
Do users use any "manual" PPSs in their everyday life?
Do users consider these strategies as effective ways to protect their privacy?

RQ1.1. **Data Minimization:** Which data minimization strategies would users be willing to adopt and why?

RQ1.2. **Removal:** What is the relationship between privacy and removing the bracelet?

RQ2. Impacts of Privacy-Preserving Strategies on Utility: How do different Privacy-Preserving Strategies impact the users' perceived utility?

RQ3. Inference and Privacy Risks: Which types of data are users most worried about being inferred, and why?

RQ4. Mental Models: What are the mental models of different types of users (about fitness trackers), and what are the relationships between users' technical understanding of how fitness trackers work, privacy concerns, and utility?

RQ5. Design Aspects: What should designers of fitness trackers take into account to improve their design both from a privacy and a utility perspective?

Interview Algorithm Syntax

General
Normal text: something to say.

- Bullet points: questions to ask.
    - Indented bullet points: indicate follow-up questions to the parent bullet point.

Logic
Custom Blue: General Conditional logic
Dark Green 1: Positive Conditional Logic
Dark Red 1: Negative Conditional Logic

Placeholder texts
Dark Yellow 1: placeholder for something specific that the interviewer will have to say
XXX: is to be replaced with participants' specific answers to the survey
[...]: instructions for where/how to get the placeholder text

Interviewer Directions
: interviewer's action
Italics: *probing question*

Misc.
// : inline comment

Interview Questions Definition

Total estimated time: **70 min**

QID	Exit Survey Reference QID	Question text	Goal
Introduction (~5min)			

// Introducing the interviewees to the project:
Hello, thank you for your participation in our research.
I am [add description of interviewer(s)].
The main purpose for this interview is to better understand [add description of the study].
 	
// Explaining the interview's conditions:
The session is going to take around 60 to 90 minutes.
During this session we will ask you general and specific questions about fitness tracking technology and privacy.
There are no right or wrong answers. The more you share your thoughts and opinion, the better.
 	
Feel free to ask any questions you have in mind, and remember that you are free to end this interview at any time, for any reason.
 	
The session will be recorded (audio + video). It's only to help me create a transcript of the interview. The recording will not be shared and it will be deleted after transcription. If you want to pause the recording for parts of the discussion, it's perfectly ok, just tell me and I'll pause it. I need you to give your consent for the participation in the experiment, including the recording.
 	
Do you have any questions?
 	
During the interview we will also ask questions related to the fitbit bracelet that was given to you.
We may refer to it by calling it as: "tracker" or "bracelet; instead of always saying "fitness tracker".
 	
// Getting their consent + financial form:
Before we begin, please sign this consent and financial form.
# Give them the consent form so that they can sign it. 				

Block A - Warm Up Questions			

A1		// About their technology knowledge:
		• What technology do you use in your everyday life?
		  ○ *smartphone?*
		  ○ *social media?*
			RQ3
			~3min

		#EndIf • Have you ever tried changing the privacy settings? ○ *Such as changing your password, your friend list, your achievement, etc.?* • Why? Now I would like to ask you about what you did after the end of the experiment. • Are you still wearing the bracelet? • After the experiment finished did you try changing privacy settings? • Did your bracelet use/wearing behaviours change after the experiment ended? • Did you revoke access to the Unil app that was used to collect your data during the experiment?	
		Block C - Privacy Concerns	

C1	D2	In the survey you filled in before, you mentioned that you would or wouldn't be worried if the following information was inferred accurately using data collected by your fitness tracker, including: [only say 2-3 data types].
		
		// Worried data types: Moderately to Extremely worried
		For the worried data types:
		You had said that you were worried about XXX:
		• Why are you worried about this type of information being inferred?
		• Do you think there could be negative repercussions from the inference?
		  ○ *Such as discrimination?*
		
		// Non-Worried data types: Not at all to Slightly worried
		For the non-worried data types:
		You had said that you were not at all worried about XXX.
		• Why are you *not* worried about this type of information being inferred?
		  ○ *Do you not think that you could be discriminated against based on this information?*
			RQ3
			~2min

C2	D1	In the survey, we asked you about what type of information you believe can be inferred accurately, based on the data collected by your bracelet.	RQ3 ~3min

		You said that you believe this type of information (XXX) can be inferred accurately, and this type of information (YYY) cannot: • Why do you think so? ○ *Why do you think some types of information can be inferred more easily than others?*	
C3	A2	**If participants said in survey they gave access to their Fitbit account to 3rd party app/social media apps:** You mentioned giving other apps or social media access to your Fitbit account: • Which ones? • Why? • Did you have any privacy concerns with doing that? / Do you have any privacy concerns with having done that? **If participants said in survey they did not give access to their Fitbit account to 3rd party app/social media apps:** You mentioned **not** giving other apps or social media access to your Fitbit account: • Why? ○ *Was it because of privacy concerns?*	RQ3 ~2min

C4	F1, F2, F4, F5	**If the participant said in the survey that they removed their bracelet:**
		You said in the survey that you removed your bracelet in certain situations.
		• Can you tell me about a situation in which you removed your bracelet?
		  ○ *Can you elaborate about why you removed the bracelet in that situation?*
		• Can you tell me about other situations?
		  ○ *What about XXX?*
		
		If the participant said in the survey that they removed their bracelet during the day:
		• *Were there specific situations that occurred **during the day** where you removed your bracelet?*
		  ○ *Why always during the day or at night?*
		
		If the participant said in the survey that they removed their bracelet at night:
		• *Were there specific situations that occurred **at night** where you removed your bracelet?*
		
		If the participant said in the survey that they did not remove their bracelet:
		You said in the survey that you kept your bracelet on at all times, except for charging it.
			RQ1.2, RQ3
			
			~3min

		• Can you elaborate about why you never removed the bracelet?	
C5	F1, F3, F4, F6	Could you try to recall a situation in which you removed your bracelet due to privacy concerns or due to being in a sensitive situation. • Were there any such situations? **If the participant says they can recall such a situation:** • What sort of situation was it? • Could you please describe it? • *What or who prompted you to remove it?* **If the participant says they did not remove their bracelet:** # Compare with survey responses • **If consistent with survey responses skip this question** • **If not consistent with survey responses, ask about inconsistencies** ○ *In the survey you had said XXX. What made you change your mind?* ○ *Why?*	RQ1.2, RQ3 ~2min
C6	F1, F3, F4, F6	• Do you think that removing the bracelet on some occasions can be useful to protect your privacy? ○ Why?	RQ1, RQ3 ~1min
Block D - Disabling Sensors			
D1	F7, F8	**If the participant said they would disable a sensor if it was possible:** In the survey, you said that if it was possible, you would disable the following sensor(s): XXX. • Why would you disable them? • Why not disable the others? • For how long would you disable it? (All the time, sometimes,...) • How would you like to be able to disable it? ○ *Physical button or software button?* ○ *Software "switch" or hardware killswitch?* • Can you give me a scenario in which you would disable them? **If the participant did not say that they would disable a sensor if it was possible:** In the survey, you said that you wouldn't disable any sensor(s) even if it was possible. • Why are you not interested in disabling sensors? • *Why would you not be willing to disable the sensors?*	RQ1, RQ5 ~4min

D2	F7, F8, F9	• In which situation(s) do you disable sensors (e.g., GPS) on your smartphone?
		  ○ *Could you please describe the situation?*
		  ○ *Why?*
		• Have you ever been in a situation where you disabled sensors on your smartphone because of privacy reasons?
		  ○ *Could you please describe the situation?*
		  ○ *Why?*
		• When switching off a sensor (e.g., GPS) on your smartphone, do you believe it is really disabled or do you believe it still runs in the background somehow?
		• Would that perception change if the switch to disable the sensor was different, for example if it was a physical switch (similar to a light switch)?
			RQ1, RQ3, RQ4, RQ5
			
			~4min

D3	F9	In the survey, we asked if you think that disabling certain sensors could be useful to protect your privacy. **If participant said useful in the survey:** • Why do you think that disabling sensors is useful for protecting your privacy? **If participant said not useful in the survey:** • Why do you think that disabling sensors isn't useful for protecting your privacy?	RQ1, RQ4 ~2min
Block E - Privacy Preserving Strategies (PPS) Exploration			
E1	E2, E3	In the survey we asked about which platforms you use to check steps and sleep data. **If participant did use the website to track steps or sleep data:** You said that you sometimes/often/always [select the frequency participant reports in survey] use the website. • Why do you check your data on the website? • Would it feel less useful if you could not check your data on the website? • Do you think you would not be able to achieve your fitness goal without this function? **If participant did not use website to track steps or sleep data:** You said that you rarely/never [select the frequency participant reports in survey] use the website. • Why do you rarely/never check your data on the website? • Would it feel less useful if you could not check your data on the website?	RQ1, RQ2 ~5min

		• Do you think you would not be able to achieve your fitness goals without this function?	
E2	E1, G3	**If participant check their steps data on a second smartphone, on a tablet, or via the Fitbit website:** In the survey you said you often checked your steps data on [list platforms other than bracelet and device that they use]. • Would you be happy if you could see only your total step count per day on these platforms? Imagine that your phone uploads your daily step count only at the end of the day (i.e. after midnight). • Would you be happy with having a delay for seeing your step count on other platforms?... ○ *Why?*	RQ1, RQ2 ~4min
E6	G3	• When you look at the step data for the day, what do you usually want to see? ○ *Total for the day, steps for every 15min, steps for every minute?* # Compare answer with their survey answer **If their answer was different from the survey:** In the survey you had said XXX. • What made you change your mind?	RQ1, RQ2 ~2min

E3	G1	Now I would like to hear your opinion about the accuracy of the data collected by the bracelet.
		
		// Establish baseline
		• Do you think the data recorded by the bracelet is accurate?
		○ *Step data?*
		○ *Heart rate data?*
		○ *Sleep data?*
		• Do you care about the accuracy of the device?
		○ *For which data types do you care about the accuracy?*
		○ *How accurate does the device have to be?*
		// A lot of papers have found that users don't care about accuracy, our results contradict those findings.
		
		If participant said that they care about the accuracy:
		In the survey, we asked you [add question G1], and you answered XXX.
		• What do you feel you would be missing if your data was less accurate?
			RQ1, RQ2
			~6min

		○ *Do you think it would be more difficult for you to achieve your fitness goal if the accuracy was lower?* **If participant said they do not care about accuracy:** In the survey, we asked you [add question G1], and you answered XXX. ● Can you explain your answer (in relation to accuracy)?	

E4	E2, E3, G1, G2, G3	● Have you ever deleted your data (e.g., content you have shared, your entire account, location history, etc…) from social media websites (e.g., Facebook, Instagram), Google, or fitness tracking apps/websites (e.g., Fitbit, Strava, MyFitnessPal)?
		
		if Yes:
		● Could you describe the sort of situation in which this occurred and the type of data that was deleted?
		○ *Why did you delete it?*
		● Were there any situations where you deleted data due to privacy concerns?
		○ Could you please describe the situation and type of data that was deleted?
		
		// Establish their technical understanding
		● How sure are you that your data was deleted from their servers?
		● Would having the assurance that your data is really deleted be valuable to you?
		endif
			RQ1, RQ2, RQ3, RQ4, RQ5
			~6min

E5	D2, E2, E3, G2, G3	# Look at what they were the most worried about being inferred from their fitness data // The following distribution is to make it seem believable, if we only say things that they are worried about they may be doubtful # Pick 1-2 things that they are the most worried, 1 thing moderately worried about, 1 thing not worried about # If they question how, say: using learning based approaches it can be done, but we cannot discuss it in detail due to it being used for a research paper that has yet to be published. In our lab we have been able to accurately infer XXX, by using a combination of step data, heart rate, and sleep data. Imagine that you would only be able to see your data on your bracelet & smartphone, but not on another device, nor the website. ● Do you think you would still use the device/find it useful? ○ Why or why not? ● As an alternative to synchronizing your data through Fitbit's servers, would you find it convenient to be able to export your data to a cloud storage (e.g., Google Drive, Nextcloud, Dropbox) and then import it on a new device? This is how WhatsApp does it's chat backup.	RQ1, RQ2, RQ3, RQ5 ~3min

		○ Why or why not?	
		Block F - Mental Model	
F1	Block C - Privacy - Mental Model	# Show participants their mental model drawing. Here is the drawing we asked you to do in the survey. It is meant to explain to us how you think your [select appropriate one: steps or sleep data] are processed and transferred on the internet. I am going to ask you a few questions to understand it better. # Give some time to the participant to look at their drawing. ● Can you explain your drawing to me? ● Which part(s) of the drawing do you think you use the most? // If they say that they don't know how to answer the question, tell them "parts" refers to arrows showing data flow. ○ *Do you think this can change in the future?* ● Which part(s) of the drawing do you think can cause privacy risks? ○ *What risks?* ○ *Why?* **If they drew something about third parties getting data:** ● Who do you think these [use the same term as they did for third parties] are likely to be? ● Do you have any concerns about them getting the data? ○ Why or why not?	RQ1, RQ2, RQ3, RQ4 ~8min
			~70min

B. Consent Form

UNIL | Université de Lausanne
HEC Lausanne
Décanat

FORMULAIRE DE CONSENTEMENT POUR EXPÉRIENCE DÉDOMMAGÉE (KODI-I_Bank)

Vous êtes invité à participer à une étude nommée « KODI-I ». Nous vous remercions d'avance pour votre participation. Cette étude est dirigée par le laboratoire de sécurité de l'information et de vie privée de l'Université de Lausanne. Cette étude est financée par Prof. Kévin Huguenin (Fond 26056639).

L'objectif de cette étude est de comprendre l'expérience, les perceptions ainsi que les opinions des personnes ayant porté le « fitness tracker, » Fitbit Inspire HR, durant la période de mai 2020 à septembre 2020 (expérience KODI). De plus, il faut que ces personnes aient complété un questionnaire au début et à la fin de la période de recherche.
Nous désirons comprendre quelles stratégies de protection de la vie privée les utilisateurs seront prêts à adopter et pourquoi, les inquiétudes par rapport à leur vie privée que les fitness tracker posent, et les modèles mentaux concernant le fonctionnement des fitness trackers.

En prenant part à cette étude, vous allez être convoqué à une interview qui durera, au plus, un total de 90 minutes. En contrepartie de votre participation, vous recevrez une compensation financière de 35 CHF. Vous serez payé par virement bancaire avant la fin du mois de février 2021.

Vous serez interviewé dans les salles de l'Université de Lausanne qui seront désinfectées, comme préconisé par les normes sanitaires en place dans le campus, afin d'éviter la propagation du virus Covid-19. Si cela n'est pas possible, les interviews seront faites à travers Zoom. Les interviews seront audio-enregistrés afin que les chercheurs puissent procéder à une retranscription et une analyse qualitative, élément essentiel à la génération des données qui serviront à la compréhension des phénomènes étudiés dans ce projet de recherche. Il est toutefois possible que nous recourions à un enregistrement vidéo additionnel si vous décidez d'illustrer vos propos à l'aide d'éléments visuels. Dans le cas d'un enregistrement visuel, nous ne filmerons que l'élément démonstrateur sans élément permettant de vous identifier. Toutes les données collectées durant l'interview seront anonymisées et uniquement utilisées à des fins de recherche. De plus, vos données personnelles ne seront jamais transmises à une entité tierce ou externe. A la fin de l'étude, les données collectées (comprenant les fichiers audios, les vidéos et les transcriptions) seront conservées durant 2 ans, puis seront finalement supprimées.

Votre participation à cette étude ne vous expose à aucun risque. Vous êtes libre de refuser de répondre à n'importe quelle question posée par la personne vous interviewant, si celle-ci (la question) pourrait vous porter préjudice ou des ennuis de quelque manière qu'il soit. Vous avez le droit, à n'importe quel moment, d'accéder aux données personnelles que vous avez fourni préalablement.

Votre participation est sur une base de volontariat, et vous pouvez vous retirer de l'étude à tout moment. Dans le cas où vous décideriez de ne plus prendre part à celle-ci, après l'avoir

UNIL | Université de Lausanne
HEC Lausanne
Décanat

accepté préalablement, vous pouvez annuler votre choix à tout instant en contactant le responsable de l'étude, sans besoin de fournir une justification. Cependant, dans le cas d'une annulation ou d'une non-participation, vous ne pourrez bénéficier de la compensation mentionnée auparavant.

Pour nous communiquer votre accord de participation à cette étude, nous vous prions de remplir et signer ce formulaire de consentement au préalable de votre participation à celle-ci. Votre signature garantit que vous avez reçu toutes les informations nécessaires et que vous exprimez votre volonté de participer, de votre plein gré.
En cas de questions concernant cette étude ou l'utilisation de vos données personnelles, nous vous prions de contacter Lev Velykoivanenko au moyen de l'adresse suivante : lev.velykoivanenko@unil.ch

Votre participation est anonyme et votre nom ne pourra en aucun cas être relié à vos réponses et vos données collectées qui resteront confidentielles. Si les résultats de cette recherche venaient à être publiés, ils le seraient de manière anonyme et/ou agrégée. L'identité des participant·e·s ne serait en aucun cas divulguée.

En signant le présent formulaire, vous attestez avoir lu les informations précédentes et avoir rempli vous-même les champs mentionnés ci-dessous. Enfin, vous confirmez **ne pas recevoir sur l'année civile plus de CHF 2'300.- de dédommagement, toute expérience confondue, de la part de l'UNIL.** Si la totalité des salaires versés par l'UNIL et des dédommagements reçus est supérieure à CHF 2'300.- par année civile, votre dédommagement sera alors soumis aux charges sociales et versé selon le calendrier des salaires.

J'accepte de participer à l'étude:

Nom, Prénom:

Date de naissance (jj.mm.aaaa):

N° d'étudiant (le cas échéant):

Date: Signature:

C. List of codes

All codes	Total participants
A general idea is precise enough	5
Advanced sportsperson need high precision	2
Anonymity protects privacy	1
As precise as possible	6
At ease using technology	3
Authoritarian country	4
Aware of privacy risks but not worried	3
beginner sportsperson don't need high precision	2
cannot escape data collection	9
Cannot give concrete examples of negative consequences	4
Could do without history	1
Data is not public anymore	6
Data protection law	1
Deletion should be a right	1
Disabling GPS	7
Disabling no sensors because don't care about it	1
Disabling no sensors because they are useful	4
Disabling sensors because they are useless	7
Disabling sensors is useful to protect privacy	1
Disabling sensors is useless to protect privacy	6
Discrimination	2
Distrust of companies	6
Distrust of how fitbit manages personal data	4
Example of negative consequences	3
Feeling close to technology	1
Fitbit data is sensitive	1
Fitbit data is informative	3
Fitbit data is not sensitive	11
Fitbit is for sport beginner	2
Fitbit is not necessary to reach personal goals	2
Fitbit is not precise enough	3
Future worry	2
Give fake information about themself	1
Give little information about themself	3
Having precise numbers is better	5
High precision for medical need	1
Inferable because habits	1
Inferable because input in app	3
Inferable because physiological data	7
Insurance	1
Interest in fitbit data	10
Intervals are not precise enough	4
Intervals can motivate to do sport	1
Lack of interest in bracelet data	5
Lack of technology knowledge	2

1

Larger intervals are not precise enough	8
Larger intervals are useless	2
Less precise so perceived as less effective	1
Limitation of social media usage	2
Lost control of personal data	5
Lost interest in fitbit	4
Love-fear feelings for technology	1
Fitbit as a watch	1
Motivation to do sport	2
Need for high precision	5
Need for low precision	7
Need for reference points in sport data to make comparison	5
Negative consequences are unlikely	8
Never deleted	3
Not inferable based on fitbit data only	4
Not inferable because not physiological data	3
Not inferable because too far fetched	4
Not really deleted	11
Not worried about data collection	1
Not worried about privacy protection	3
Not worried about sport related data	2
Not worried about targeted ads	3
Not worried because data is not sensitive	6
Not worried because it is a commonly shared data	4
Not worried because many people use it	1
Not worried because not a target	7
Not worried because nothing to hide	4
Permission management	4
Personal data should be kept private	9
personal data should not be shared	9
Personal data should not be sold	1
Personal data should not leave the phone	2
Physical button	2
Possible negative consequences	9
Powerful algorithms	1
Precision for relevant data	8
Precision need varies with different usage	4
Privacy protection against family members	4
privacy protection on social media	5
privacy protection strategy	3
Privacy setting of fitbit modified	1
Really deleted	1
Relevant data for sport activity	1
Removal is useful to protect privacy	4
Removal is useless to protect privacy	2
Removing for a long time can help to protect privacy	3

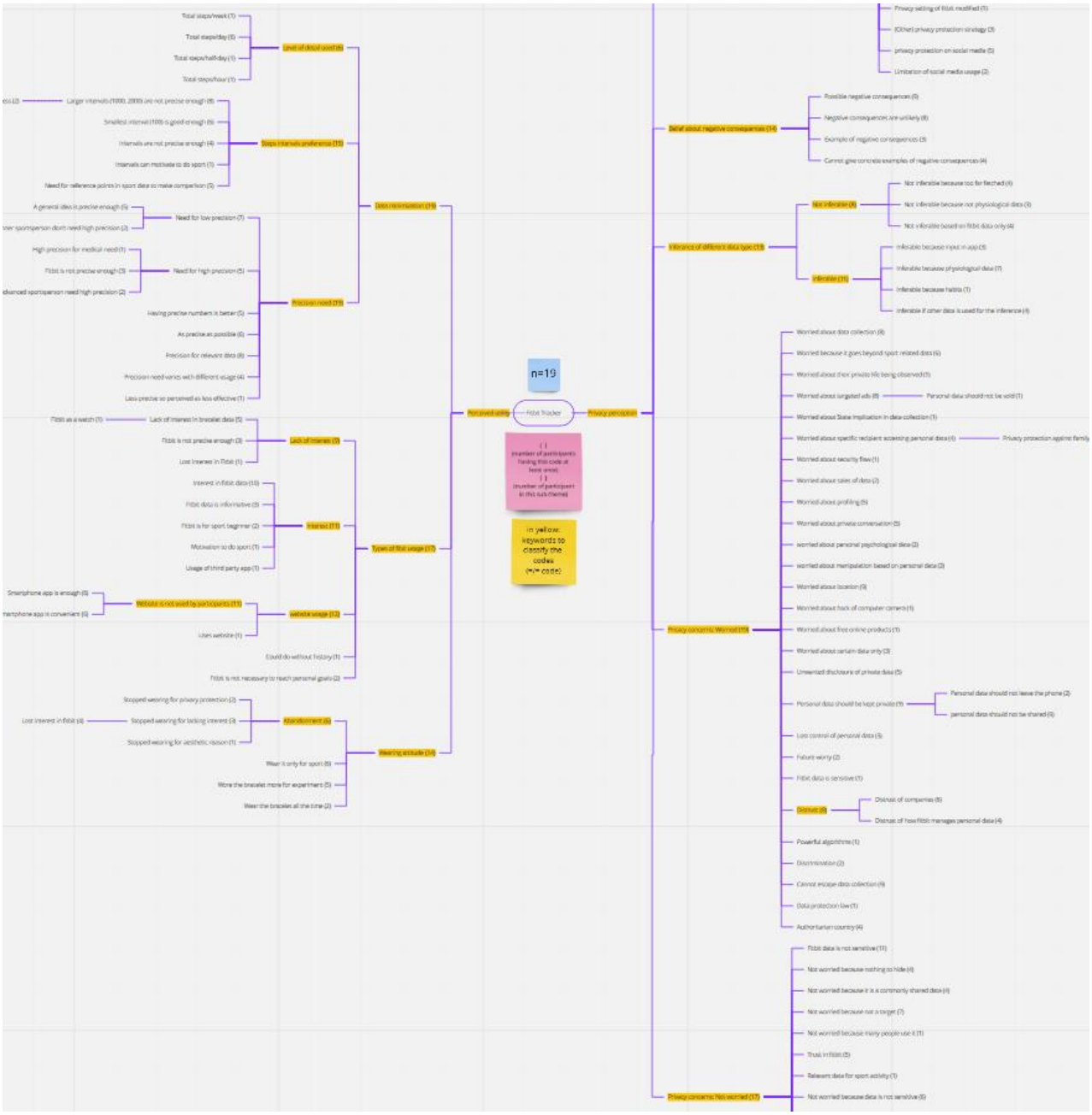
2

Sensors are not really disabled	5
Sensors are really disabled	4
Smallest interval is good enough	6
Smartphone app is convenient	6
Smartphone app is enough	6
Software button	6
Stopped wearing for aesthetic reason	1
Stopped wearing for lacking interest	3
Stopped wearing for privacy protection	2
Technology dependency	1
Technology is a need	2
Technology is useful	3
Think before posting	4
Total steps/day	6
Total steps/half-day	1
Total steps/hour	1
Total steps/week	1
Trust in fitbit	5
Unwanted disclosure of private data	5
Usage of third party app	1
User type: Advanced	1
User type: Beginner	2
Uses website	1
Using technology is a normal daily activity	3
Want to keep control of personal data	4
Wear it only for sport	6
Wore the bracelet more for experiment	3
Worried about certain data only	3
Worried about data collection	9
Worried about free online products	1
Worried about hack of computer camera	1
Worried about location	9
worried about manipulation based on personal data	3
worried about personal psychological data	2
Worried about private conversation	5
Worried about profiling	6
Worried about sales of data	2
Worried about security flaw	1
Worried about specific recipient accessing personal data	4
Worried about State implication in data collection	1
Worried about targeted ads	8
Worried about their private life being observed	1
Worried because it goes beyond sport related data	6
Wear the bracelet all the time	2
Inferable if other data is used for the inference	4

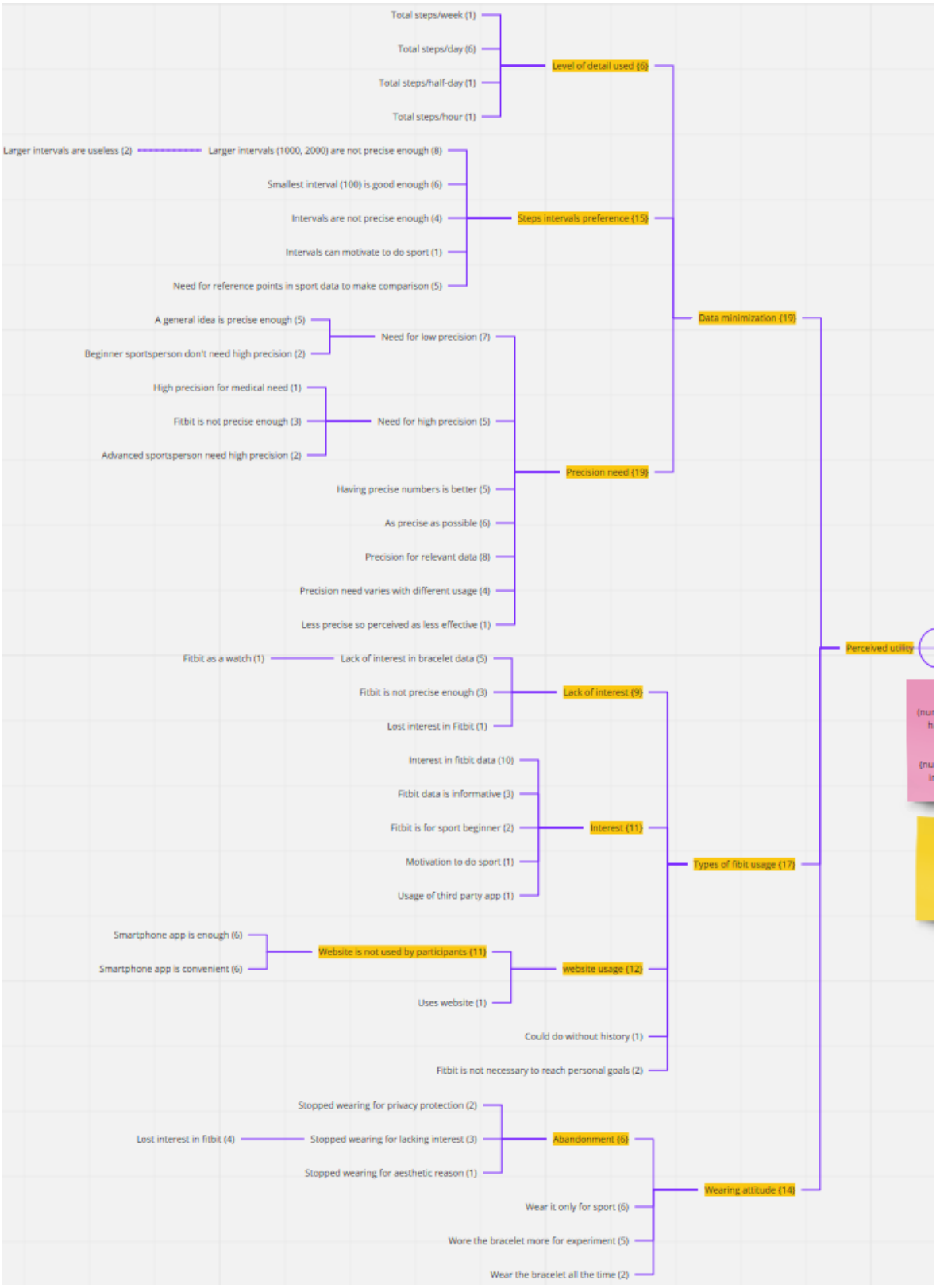
3

D. Miro Tree organizing all the codes

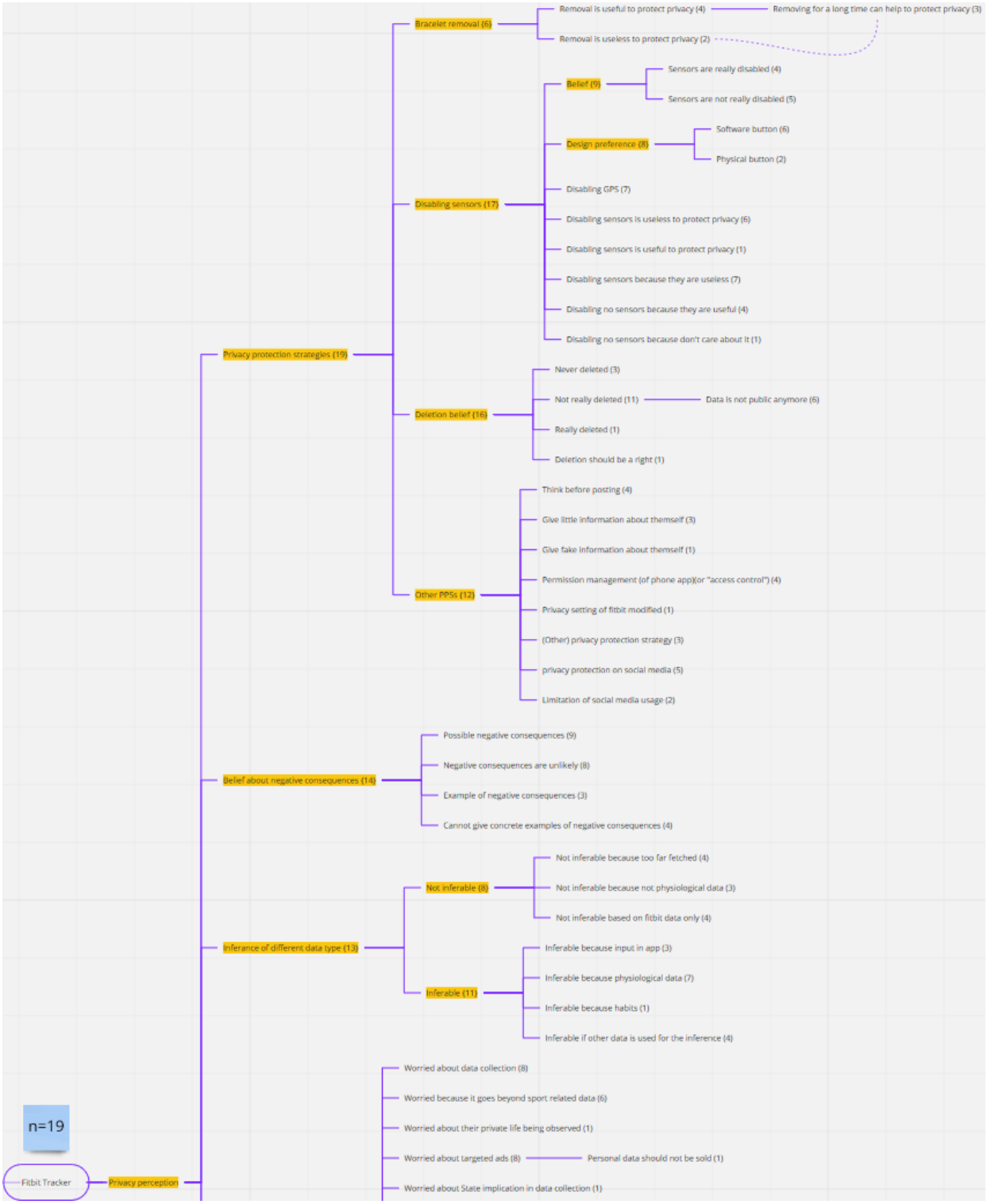
a. Miro all tree: overview



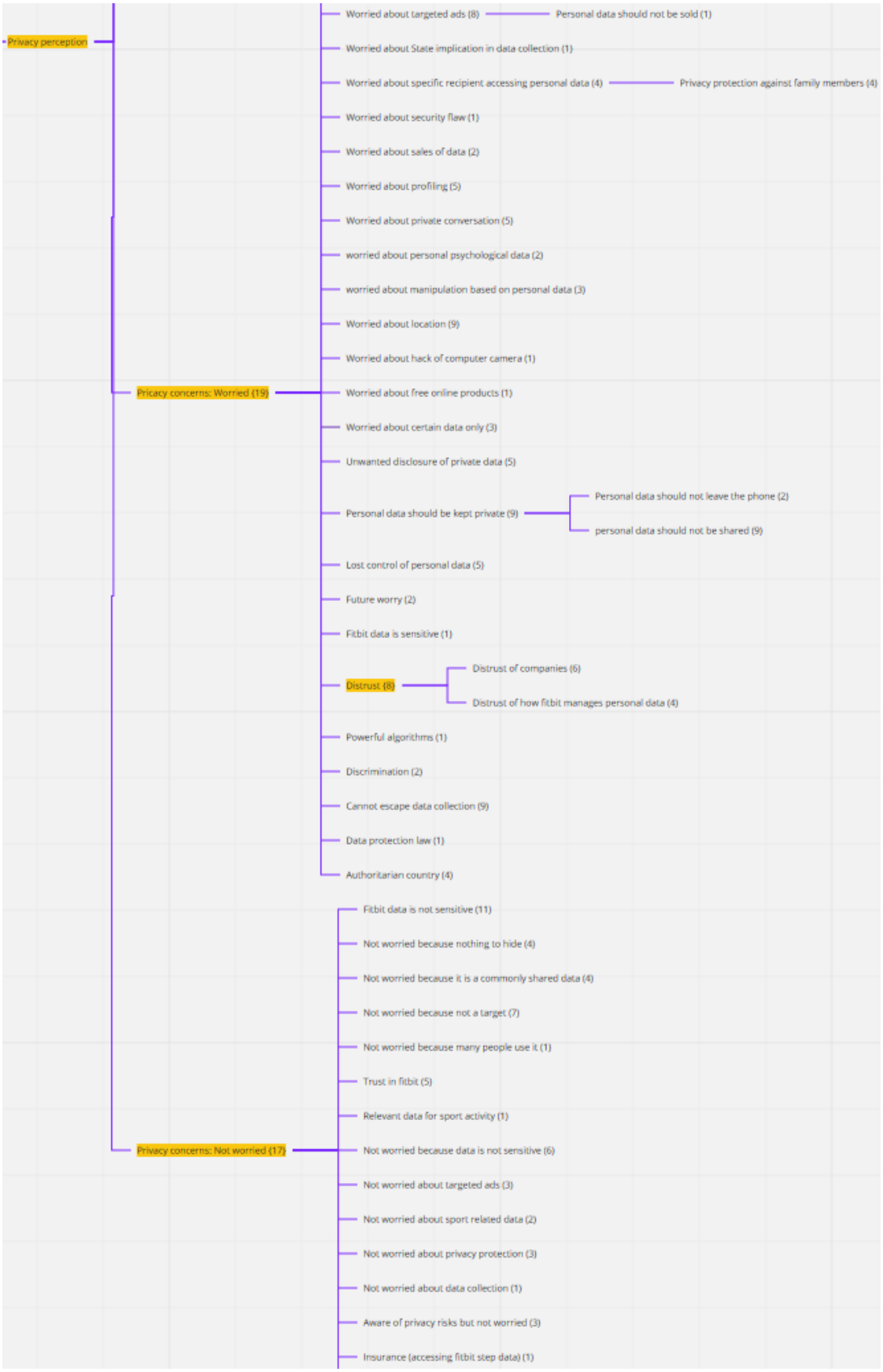
b. Miro tree: left side



c. Miro tree: right side part 1



d. Miro tree: right side part 2



E. Participants' concerns and perceived likelihood of inference of religion and personality

Participant	Level of concern		Belief about inference accuracy	
	Religion	Personality	Religion	Personality
P1	Not worried	Extremely worried	Moderately precise	Not precise
P8	Not worried	Not worried	Not precise	Slightly precise
P11	Not worried	Slightly worried	Moderately precise	Moderately precise
P15	Not worried	Slightly worried	Not precise	Moderately precise
P17	Not worried	Not worried	Not precise	Extremely precise
P18	Not worried	Not worried	Not precise	Moderately precise
P21	Moderately worried	Very worried	Moderately precise	Moderately precise
P23	Not worried	Not worried	Not precise	Moderately precise
P28	Moderately worried	Very worried	Not precise	Moderately precise
P32	Not worried	Not worried	Not precise	Slightly precise
N=10	**Not: n=8** **Moderately: n=2**	**Not: n=5** **Slightly: n=2** **Very: n=2** **Extremely: n=1**	**Not: n= 7** **Moderately: n=3**	**Not: n=1** **Slightly: n=2** **Moderately: n=6** **Extremely: n=1**