To Be, or Not to Be Notified

Patrick Murmann, Delphine Reinhardt, Simone Fischer-Hübner

To Be, or Not to Be Notified

Eliciting Privacy Notification Preferences for Online mHealth Services

Patrick Murmann[1], Delphine Reinhardt[2], and Simone Fischer-Hübner[1]

[1] Karlstad University, Sweden
[2] University of Göttingen, Germany

Abstract Millions of people are tracking and quantifying their fitness and health, and entrust online mobile health (mhealth) services with storing and processing their sensitive personal data. Ex post transparency-enhancing tools (TETs) enable users to keep track of how their personal data are processed, and represent important building blocks to understand privacy implications and control one's online privacy. Particularly, privacy notifications provide users of TETs with the insight necessary to make informed decision about controlling their personal data that they have disclosed previously. To investigate the notification preferences of users of online mhealth services, we conducted an online study. We analysed how notification scenarios can be grouped contextually, and how user preferences with respect to being notified relate to intervenability. Moreover, we examined to what extent ex post notification preferences correlate with privacy personas established in the context of trust in and reliability of online data services. Based on our findings, we discuss the implications for the design of usable ex post TETs.

Keywords: privacy, transparency-enhancing tool, usability, personas, mhealth

1 Introduction

According to Cisco Systems, the worldwide number of wearable devices connected to the Internet will climb from 325M in 2016 to a projected 929M in 2021 [16]. This means that an increasing number of people track their health using personal activity trackers like fitness bracelets or smart watches. Such wearables allow their users to collect statistics about a plethora of physiological characteristics, and optionally enrich the data with location data and information about one's lifestyle. Hence, potentially seamless information are collected about a person's health, pinpointing it in time and space.

However, according to [7], the consequences of how mhealth data processed by data services are not fully transparent to the users of such services. Data subjects often lack the information necessary to make informed decisions about managing their personal data and to exercise their right of intervenability, especially in scenarios that involve third parties. Conversely, the EU General Data

Protection Regulation (GDPR) [2] stipulates that users of data services must be able to control their personal data, and grant them a right of ex post transparency and intervenability rights to delete, rectify, block or export data or to object to data processing (GDPR Art. 15–20). However, exercising these intervenability rights implies that data subjects are aware of how their data are processed, and therefore depend on processes that are transparent and comprehensible as mandated by GDPR Art. 12–15. *Ex post transparency-enhancing tools* (TETs) facilitate transparency by informing about how a data subject's personal data have been processed by online services, e.g. by the means of privacy notifications [17]. As it has been shown in our previous work, however, existing privacy indicators of TETs often lack transparency themselves in that their settings are not always verifiable or customisable [11].

Seeking to infer viable predictors for the design of usable TETs, our contribution is to investigate the notification preferences of users of mhealth services in terms of (1) how data processing scenarios can be grouped contextually, (2) to what extent these preferences can be predicted by means of privacy personas, and (3) how intervenability relates to notification preferences. Ultimately, our goal is to help designers of TETs to provide users with default settings for receiving privacy notifications based on a user's predisposition. Receiving privacy notifications that are tailored to their individual needs will allow users of mhealth systems to make informed decisions about controlling their personal data they have disclosed previously. We postulate the following hypotheses:

H1. Users of online mhealth services have different notification preferences depending on the contextual cue underlying the notifications.
H2. There is a correlation between a user's privacy persona (Sect. 3.2) and her notification preferences.
H3. The ability to intervene with the processing of one's personal mhealth data has an impact on one's notification preferences.

Our paper is structured as follows: Sec. 2 discusses related work. Sec. 3 describes the methodology applied in our online study. Sec. 4 presents the results, while we discuss our findings in Sec. 5, before concluding this paper in Sec. 6.

2 Related work

Our work shares similarities with [3,4,6,12,19]. Similar to [19], our study relies on the concept of personas and how they relate to behaviour and their consequences, but investigates the outcome of preferences for privacy notifications instead of behavioural intent. It is related to the work of Knijnenburg et al. [6] in that it accounts for multiple dimensions that lead to data subjects being grouped in terms of their privacy attitude. However, we do not seek to segment subjects according to their disclosing styles, but to establish a correlation between their disposition in terms of privacy and their notification preferences. Like Emami-Naeini et al. [12], we envision a privacy assistant that provides its users with customised notifications about personal data processing. However, the measure

that reflects the independent variables in our study are not constituted solely by discomfort, but by the overall values captured by a particular privacy persona.

Related to ex post transparency, Harkous et al. [4] touched upon privacy indicators by providing insight about potential consequences based on history-based insight about data processing. Our work complements these indicators in that we aim to establish the circumstances under which users of TETs want to receive such notifications. Moreover, our previous work published in [3] indicates that participants have different notification preferences depending on whether they could intervene, i. e., do something about how their data were processed. Hence, we follow up this research by investigating whether and to what extent intervenability has an effect on users' notification preferences.

In summary, our study follows a new direction in terms of (1) how scenarios dealing with ex post personal data processing can be grouped conceptually, (2) how privacy personas established in the literature [9] relate to a user's preferences of being notified about such scenarios, (3) to what extent intervenability has an impact on these decisions, and (4) what implication these findings have for the design of usable TETs.

3 Methodology

Our study was implemented using an online questionnaire that consisted of three parts, which are addressed in Sec. 3.1, 3.2 and 3.3, respectively.

3.1 Demographics and usage behaviour

In addition to demographics, the first part collected information on the types of devices our participants owned, what they were using them for, and with whom they shared their data. These insights helped us better understand how our participants reflected the intended target audience of users of mhealth services.

3.2 Privacy personas based on privacy statements

The second part dealt with privacy statements, which reflected our participants' privacy personas according to Morton et al. [9]. We considered but ultimately disregarded alternative models established in the literature, such as Dupree et al.'s [1] segmentation based on qualitative research, as well as Westin's [18] tripartite, mostly linear classification of privacy personas whose conception predates the advent of the Internet age.

We chose to segment users of online mhealth services according to their privacy attitude based on 15 statements described in Morton et al.'s study [8]. Test subjects assigned a total of 70 points as weights of 0–10 among 15 statements, which map to five triples that reflect the dominant factors of each of the five personas. The reason for choosing their segmentation over alternative approaches is twofold: Firstly, the methodology suggested by Morton et al. is based on quantitative research that does not require manual post-processing.

Secondly, the narrative factors and themes identified by Morton et al. are generic in that they capture the notion of a user's trust in online services, which potentially map similarly to scenarios encountered in mhealth environments. We have therefore slightly adapted the original statements to reflect the particularities of the mhealth context. By doing so, we tried to capture the original meanings that reflect the five personas established by Morton et al.: Security Concerned (SC), Organisational Assurance Seekers (OAS), Crowd Followers (CF), Benefit Seekers (BS), and Information Controllers (IC). SC seek the use of technological means to ensure the security of their personal data. OAS look for formal indicators, such as privacy policies, that warrant their trust in a service. CF value the reputation of a service and heed the recommendations of trusted peers. BS are after useful benefits and are willing to give up personal data in exchange. IC seek to control the collection, access to and use of their personal data. The Flesch reading ease and Flesch-Kincaid grade level [5] were not affected by our changes and remained stable at 59.1 and 8.8, respectively. To mitigate the effects of cognitive fatigue and habituation, we randomised the order in which the statements were displayed. We have made available supplementary material about the study design and results on a dedicated website.[3]

3.3 Categories of notification and notification scenarios

The third part of the study covered notification preferences, which describe what kind of scenarios related to personal data processing data subjects want to be notified about. We distinguished three categories of privacy notifications:

Breaches refer to "a breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, personal data [...]" according to GDPR Art. 4 (12). Hence, breaches cover both accidental incidences and deliberate misappropriation of personal data by the data controller or by affiliated parties.

Consequences. This category seeks to clarify the consequences that arise for a user of an mhealth service due to the processing of her data. It covers consequences based on actual facts as well as hypothetical outcomes given the circumstances at hand. Consequences differ from breaches in that respective outcomes are the result of personal data processing that is compliant with the privacy law, or that pertain to a possible occurrence in the future.

Practical tips refer to customised guidance for a user intended to improve her online privacy. They are customised in that they pertain to personal situations, and therefore address matters to which users can relate. Practical tips may suggest a concrete change of behaviour or motivate action by notifying her to consider alternatives that would improve her data privacy.

We constructed 17 hypothetical scenarios with themes related to mhealth (Table 1). We relied on plain language and avoided technical terms that might

[3] `https://murmann.hotell.kau.se/notification-preferences/` [10]

Table 1: Notification scenarios

Breaches	1. Your data are stored longer than is specified in the privacy policy that you have agreed to. 2. Your data are processed differently from what is specified in the privacy policy that you have agreed to. 3. Your data are shared with parties not covered in the privacy policy that you have agreed to. 4. Your mhealth provider used software in which a critical software bug was detected, which made the service vulnerable to hacking and unauthorised access. 5. Your data got lost on their way to your mhealth service provider. 6. Your mhealth service provider was attacked by an unknown party on the Internet, which succeeded in copying some of the data. 7. One of your mhealth provider's partners has access to data not intended for them.
Consequences	8. Recording both your location data (GPS) and your health data (like your pulse and blood pressure) allows someone with access to these data to know where you performed your activities, like the trails you hike or ride most frequently and how you performed along the way. 9. Recording both your health data (like your pulse and blood pressure) and the time allows someone with access to these data to learn about your general life style, like how fit you are, your health risks and diseases. 10. You receive customised advertisements about healthy food, sports products and insurances based on the data recorded using your device. 11. Your mhealth provider shares your data with other companies for the purpose of profiling (analyse your data for patterns). 12. Your mhealth provider or their partners reside outside of Europe. 13. Someone with access to your location data (GPS) and the location data of other users may learn when you have spent time together and what type of activities you have been performing. 14. Your mhealth provider changed their privacy policy twice since you started using their services a couple of years ago. Each time, the policy stated that by continuously using their service users will agree to the terms and conditions.
Tips	15. Your current mhealth service provider shares your data with an online marketing company and an insurance company. There is a different service provider that offers you the same level of service quality and device compatibility but that does not share your data with third parties. You have the option to have your archived data transferred to the new provider and have them erased from the current one. 16. Your mhealth device senses your pulse rate using the highest resolution possible. The device has an option to inform you that by using this setting you collect more data than is necessary to track your health reliably. 17. You have stopped tracking your health and switched off your mhealth device. You have the option to be notified that you can download and/or erase your health data that are currently stored online by your mhealth service.

be misinterpreted by lay persons. The overall Flesch reading ease of the scenarios was 54.4 and the Flesch-Kincaid grade level was 10.8.

Each scenario was displayed in random order on a dedicated screen that showed the narrative and two questions. In the first question, we asked the participant whether she wanted to be notified about it. Possible answers were 'Yes,' 'No' and 'I don't know'. If a participant selected the latter option, a secondary set of options appeared. This set captured the respondent's uncertainty and offered four follow-up options: "I don't understand the scenario," "I can't relate to the scenario," "I need further details to make an informed decision" and 'Other.' We deliberately excluded a free text field in lieu of 'Other' to prevent users from disclosing sensitive information.

The second question, "Does your ability to object to the processing of your data affect your choice above?" aimed at capturing the impact of intervenability on the choice to be notified. For the sake of comprehensibility, we substituted the verb 'to intervene' by 'to object' even though objection does not holistically reflect intervenability. Available options were 'Yes,' 'No' and 'I don't know,' the latter triggering the following secondary options: "I don't understand the question" indicated ambiguity of the task itself. "I don't know what it means to object in this context" meant that a respondent understood the concept of intervening, but felt unable to apply it, "I wouldn't know how to object" implied inability to exercise the legal right, and 'Other' covered everything else.

3.4 Online survey

We implemented an online questionnaire and hosted it on a web server located in Germany. The answers of the participants were stored using pseudonyms, so that they could later be linked back to a participant ID assigned by the crowd sourcing platform for payment purposes. Once the participants had been paid for their work, these IDs were removed for the purpose of data minimisation. Before publishing the questionnaire, we conducted six independent user tests to evaluate its usability and to fix minor issues. The study was approved by the ethics committee of the University of Göttingen.

3.5 Recruitment

We recruited our participants through the crowd sourcing platform Prolific Academic Ltd[4] because all their data processing was conducted within the EU and its workers were reviewed as being reliable [13,14]. The population of workers were screened using three criteria: (1) 18+ years of age, (2) reside within the borders of the EU, the European Economic Area, or an European country in which data protections laws similar to the GDPR applies, and (3) own a mhealth device in the form of a fitness tracker. The high percentage of workers from the UK [15] mirrored the population primarily reflected in the studies conducted by Morton et al. [8,9]. Our test subjects finished the questionnaire in roughly

[4] `https://prolific.ac/`

20 minutes. Considering European standards for minimum wages, we paid the workers €8.4/hour, i. e., €2.8 for 20 minutes. The study was published between August 17 and 18, 2018, during which time 300 submissions were gathered.

4 Results

4.1 Demographics and usage behaviour

82% of our participants were from the UK. The second largest groups were from Portugal and Spain with 3% each, and Italy with 2%. The rest hailed from all over Europe. 69% of our participants were female and 31% male, one participant identified as 'other.' Their age distribution was similar to the one published for the total population of workers available on Prolific [15].

The majority of our participants owned mainstream devices, such as fitness bracelets or smart watches. Only few owned breast belts or headbands. The predominant purpose reported was to track their fitness and motivate them to exercise. 47% used their device to track their geographic location. This implies that their devices are capable of processing GPS signals, a feature usually found only in premium price segments or in combination with mobile phones. Less than half our participants shared their data with relatives, and one third shared them with acquaintances denoted as friends. More than one third did not share their data at all. Roughly 5% of our participants were not using mhealth devices (anymore/yet). The majority were mid-term or long-term users who had been using their devices for four months or longer [10].

4.2 Privacy persona segmentation

Composition analysis. Overall, the distributions of most statements were similar in terms of their means and quartiles. However, statement 10 was a noticeable exception in that its mean and quartiles differed significantly from the ones of statements 11 and 12. This peculiarity motivated a deeper in-between analysis of the triples that constituted each of the five personas established by Morton et al. [8].

In many cases, we observed a lack of coherence between the three statements of the triples, and in some cases we even detected negative correlations. The dataset contained a considerable number of extreme values, i. e. weights of zeros and tens assigned to individual statements. For statements 7, 9 and 10, zero was the weight assigned most frequently. Overall, 3.3% of our participants assigned their points exclusively to the two scores 0 and 10, and 9.7% relied on four or fewer different patterns to allot their 70 points among the privacy statements. In 17.3% of all 1500 (300 participants $\times$ 5 triples) sets of triples, two statements were very high (≥ 9), whereas the corresponding third one was very low (≤ 1). These patterns seemed random and contradicted the supposedly high coherence of the statements constituting the triples. More importantly, it questioned the explanatory power they posed in terms of indicating privacy personas. The high amount of variance was also reflected in noticeably low values of Cronbach's α for all five triples ($\alpha = \{0.15, 0.25, 0.31, 0.30, -0.02\}$).

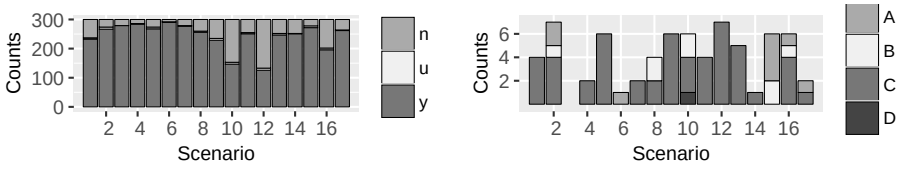
Table 2: (a) Classification of the participants ($n = 300$) according to Morton et al. [8]. (b) Number of personas including duplicates.

(a)

	SC	OAS	CF	BS	IC	Σ
Incl. duplicates	79	60	37	44	45	265
Excl. duplicates	37	24	19	25	23	128

(b)

#Personas	0	1	2	3
Count	109	128	52	11



(a) (b)

Figure 1: (a) Counts of notification choices ($n = 300$) for scenarios 1–17: **y**es, **u**ndecided, **n**o. (b) Counts of reasons for being undecided: **A.** Scenario unclear, **B.** Cannot relate to scenario, **C.** Need further details, **D.** Other.

Classification. Irrespective of the incoherence, we carried out the classification according to Morton et al. [8] to ascertain the privacy personas of our participants (Tab. 2a). We noticed that 109 respondents could not be classified uniquely, and that an additional 63 were ambiguous in terms of being classified as more than one persona (Tab. 2b). Morton et al. designated such cases as 'unclear,' and specified clearly specified personas as cases with but a single dominant triple.

We therefore tried to establish alternative personas based on the privacy statements using both k-means analysis and principal component analysis. However, neither method yielded satisfactory results. In the latter case, even the Kaiser-Meyer-Olkin and Bartlett's tests failed due to the composition of the underlying data.

We repeated both analyses on an adjusted dataset, in which we removed cases with fewer than four different patterns used for weighing the 15 privacy statements. On average, it took our participants 218 seconds to allot their 70 points. We removed cases in which a respondent had spent less than 90 seconds on this sub task, which left us with a total of 240 cases. However, the outcomes of the analyses did not change and we therefore rejected the hypothesis that the noise inherent in the data was the result of superficiality on the part of some of the participants.

4.3 Notification preferences

Per-scenario analysis. Most participants chose either 'yes' or 'no' as answers on the questions of whether they wanted to be notified (Fig. 1a). On average,

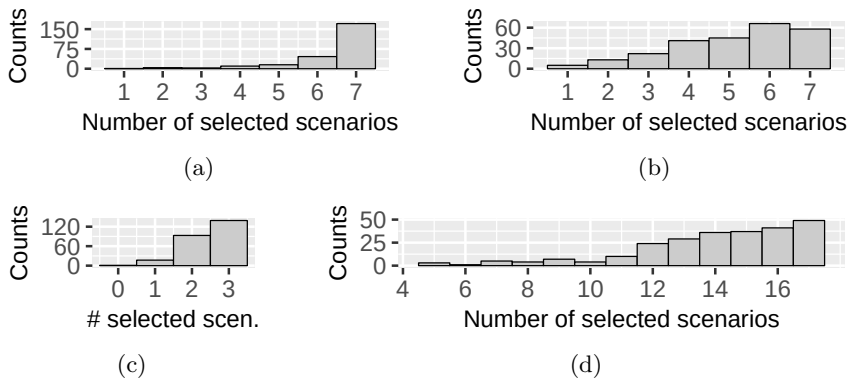
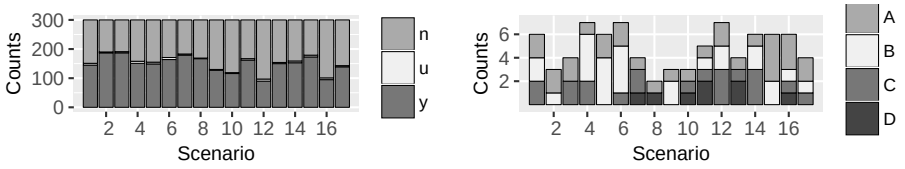


Figure 2: Counts of sums of positive notification choices per category ($n = 250$): (a) Data breaches, (b) Consequences, (c) Practical tips, (d) Overall.

less than five respondents per scenario chose 'I don't know' (Fig. 1b). The reason for doing so indicated most frequently was that the respondent needed further details to make a decision (option C). In most cases, the scenarios seemed to be comprehensible (few counts of option A), yet not always personally applicable (few counts of option B). Scenarios 10, 12 and 16 registered a noticeably low amount of positive choices. As for accidental events, scenarios dealing with software vulnerabilities (scenario 4) and cyber attacks (6) registered the highest counts overall, whereas data loss (5) registered slightly fewer counts and roughly at the mean of the breaches category. Scenarios about hypothetical data processing (8, 9, 13) registered average amounts. Scenarios related to location data (8, 13) registered counts that were slightly above the means of the categories of consequences and tips.

For the subsequent analysis, we removed from the dataset those participants who answered 'I don't know' on any scenario, which resulted in 250 out of the original 300 cases. As the cardinalities of the three categories of scenarios varied (7 breaches, 7 consequences, 3 tips), we introduced a metric of normalised sums for each group. They represented a uniform measure of the amount of positive choices irrespective of the cardinality of that group. Overall, 83% of our participants wanted to be notified about the circumstances described in the scenarios. 92% wanted to be notified about privacy breaches, 74% about consequences, and 83% about tips.

Fig. 2 shows the sums of positive choices for (a) breaches, (b) consequences, (c) tips, and (d) throughout all scenarios. The counts reflect how many different scenarios of each respective category our respondents wanted to be notified about. As regards breaches, roughly two thirds (171) wanted to be notified about all seven scenarios, whereas few wanted to be notified about five or fewer scenarios. For consequences, any combination of four to seven out of seven scenarios accounted for 84% of the cases. For practical tips, we observed that 93% wanted



(a) (b)

Figure 3: (a) Counts of whether intervenability affects choice for scenarios 1–18: **y**es, **u**ndecided, **n**o. (b) Counts of reasons for being undecided: **A.** Task unclear, **B.** Intervenability unclear, **C.** Unclear how to intervene, **D.** Other.

to be notified about at least two out of three scenarios. Overall, more than 86% of the respondents wanted to be notified about any 12 out of 17 scenarios.

Hence, we conclude that hypothesis **H1** holds in that our participants reported distinctively different patterns of notification preferences for each of the three categories of notifications.

Correlation with Morton et al.'s privacy personas. We conducted a frequency analysis of positive choices across the scenarios. Overall, all participants had high demands of being notified about privacy breaches. The only noticeable deviation was when the retention period of personal data storage was exceeded (scenario 1), which subjects classified as Crowd Followers and Benefit Seekers were less interested in. As for consequences, scenarios 10 and 12 registered noticeably low values for all personas, but especially among Benefit Seekers whose interest for being notified about consequences were generally low. Conversely, Organisational Assurance Seekers were among the ones who had the highest demands in this category. Getting the best service possible (scenario 15) seemed important for all personas.

We applied logistic regression to establish a model that helped us predict the notification settings of a subject based on her privacy persona, but failed due to high collinearity between the regression coefficients, which either resulted in high residuals and unusable models. Moreover, we investigated the statistical relationship between privacy personas and notification preferences by relying on general linear models. For multiple combinations of categories and personas both the Levene's test of homogeneity of variances and the post hoc tests failed. In most combinations, the models thus established were not significant for either all personas or just the ones clearly classified.

We therefore conclude that hypothesis **H2** does not hold in that we were unable to establish a holistic model that describes a relationship between notification preferences and privacy personas established by Morton et al. [8].

Intervenability. Our participants' opinion on whether their ability to intervene with the processing of their mhealth data impacted their choice for being notified is depicted in Fig. 3a. On average, barely 50% answered positive on the impact

of their intervenability, which was broken down into 55% breaches, 46% consequences, and 45% tips. Scenarios dealing with deviating processes (scenario 2) or nameable affiliates (3, 7, 11, 15) registered high positive counts. Similar to notification, scenarios 10, 12 and 16 registered the fewest positive counts. On average, roughly five respondents per scenario chose 'I don't know' (Fig. 3b). The reason indicated most frequently was (A) "I don't understand the question" (35%) and (B) "I don't know *what it means* to object" (30%). Not knowing *how* to object (C) was registered in 22% of the cases. Only few respondents selected 'Other' for either notification or intervenability as their reason for being undecided. This suggests that the three other options captured the reasons for their hesitation satisfactorily.

To analyse the effect of intervenability on our participants' notification preferences, we cleared the dataset of any ambiguous cases related to either notification or intervenability, which resulted in 225 unambiguous cases with values of 'yes' or 'no' for both variables. We analysed the choices in each of the 17 scenarios using a cross tabulation of notification preferences and impact of intervenability [10]. The residuals of all scenarios were positive for identical choices, meaning the observed counts exceeded the expected counts. It indicated that the counts of identical choices (yes/yes or no/no) were high compared to the counts of deviant choices (yes/no or no/yes), which suggested that the two variables correlated.

We therefore conclude that hypothesis **H3** holds in that intervenability had an effect on our participants' notification preferences, even though the absolute positive counts were low.

5 Discussion

5.1 Segmentation of ex post transparency preferences

Since we were unable to establish a correlation between notification preferences and the privacy personas established by [8], we further seek to investigate the segmentation of notification preferences in the context of ex post transparency, which would allow us to derive alternative personas for this context.

For this purpose, we conducted a principal component analysis of the unambiguous data set ($n = 250$) obtained in Sec. 4.3. Using three factors [10], we found one cluster capturing scenarios for privacy breaches and for consequences in the form of privacy-related events (not classifying as a breach) that had actually taken place (scenarios $1, 2, 3, 5, 6, 7, 11, 14, 15, 17$, Cronbach's $\alpha = 0.75$), one cluster capturing scenarios that covered 'hypothetical' consequences that may occur in future (scenarios $8, 9, 12, 13, 16, \alpha = 0.62$), and one cluster capturing solely the scenario of targeted marketing-related privacy risks (scenario 10). Despite the overall relatively low alpha values, these results show promise in that privacy personas for ex post transparency based on privacy notifications may be established into personas, namely personas for those primarily interested to be notified about privacy-critical events that took place, those primarily interested in hypothetical privacy risks, and those primarily interested to be notified about consequences related to direct marketing and unsolicited messaging.

We seek to further investigate this research question in future work, in which we intend to refine the segmentation process based on notification preferences coded as ordinal variables rather than dichotomous variables.

5.2 Design implications for TETs

It follows from Sec. 4.3 that the majority of our participants preferred to be notified over not being notified. This means that in cases when a user's preference cannot be determined equivocally, acting upon an event and sending a notification will be a reasonable default setting for a TET, which is in line with the data protection by default principle (Art. 25 GDPR). Nonetheless, we suggest that additional settings for notification clusters be available, which correspond to privacy personas for ex post transparency (Sec. 5.1). Depending on the outcome of a refined user segmentation in future work, these could, e. g., be presets related to targeted marketing, privacy breaches, incidents that have already taken place, and to hypothetical privacy risks that may occur in future.

Most respondents chose unambiguously in that they selected either 'yes' or 'no' for notification and intervenability. However, the few ambiguous choices we registered indicate that there is room for improvement in how notifications are framed, and in the amount and type of information they should contain to satisfy the demands of the target audience.

With respect to notifications, the ambiguity that was registered most frequently was a lack of details regarding the circumstances described in the scenarios. Consequently, respective respondents needed additional or more specific information to make informed decisions. To mitigate cognitive load, TETs will have to rely on multiple levels of detail to convey the full picture of sophisticated data processing scenarios. One way to accomplish this might be to start with a coarse-grained overview and provide details upon request [3], a gap that has been detected in the literature for many TETs [11]. The variety of preferences expressed by our participants supports our previous findings in that notification settings should be transparent and customisable [11].

As for intervenability, both the concept itself and how to leverage respective rights has been unclear for some respondents. Hence, TETs will have to provide such knowledge upon request. Ideally, TETs will guide users in exercising their legal right to manage their personal data and information privacy. For this, privacy notifications should be coupled with context-specific guidance on how users can react by easily exercising their intervenability rights, preferably electronically.

5.3 Limitations

Changing Morton et al.'s original statements [8] was carried out with great care, but we did not validate whether the original meanings have been preserved and carried over to the context of mhealth, nor whether the respondents interpreted

them the way they were intended. Hence, the deviation from the original statements and the altered usage context may have had an impact on individual statements, and thus the privacy personas.

All data provided by our participants were self-reported information, both in terms of the filter criteria used to recruit them via Academic Prolific, and the statements made in the online survey. All data were examined in terms of plausibility by the first author before the participants were paid.

Technical terms, such as 'privacy policy' and the concept of what it means to intervene with the processing of one's personal data, were briefly described in the study. However, we did not verify whether the respondents had actually understood the legal concepts underlying these terms.

6 Conclusion

We have conducted an online study that aimed at assessing the notification preferences of users of online mhealth services in terms of being notified about privacy breaches, consequences and tips. Our objective was to investigate to what extent ex post TETs should be individualised and what are suitable defaults that provide their users with meaningful notification settings.

We have found that notification preferences can be grouped as distinctive categories related to the contextual cues underlying the privacy notifications. Moreover, the legal right of intervenability had an impact on our participants' choices to be notified, which implies that guidance of users in exercising this right in response to notifications should be investigated further in the future.

We have, however, been unable to ascertain a correlation between notification preferences and privacy personas as described by Morton et al. [8]. Nonetheless, our first statistical analyses showed that it is possible to elicit new privacy personas for ex post transparency scenarios based on notification preferences. This is a direction of research that we will further pursue in the future.

Acknowledgements

This research has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 67573 and the SSF project SURPRISE.

The authors thank Dan Larsson and Erik Wästlund for advice on the study design, and advice on conducting and interpreting various statistical analyses.

References

1. Janna Lynn Dupree, Richard Devries, Daniel M Berry, and Edward Lank. Privacy Personas: Clustering Users via Attitudes and Behaviors toward Security Practices. In *Proc. of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2016.

2. The European Parliament and the Council of the European Union. *Regulation (EU) 2016/679 of the European Parliament and of the Council*, 2016.
3. Simone Fischer-Hübner, John Sören Pettersson, Julio Angulo, Jessica Edbom, Mia Toresson, and Hendrik Andersson. D:C-7.3 Report on end-user perceptions of privacy-enhancing transparency and accountability. Technical Report D37.3, A4Cloud Project, 2014.
4. Hamza Harkous, Rameez Rahman, and Karl Aberer. Data-Driven Privacy Indicators. In *Proc. of the Symposium on Usable Privacy and Security (SOUPS)*, 2016.
5. J Peter Kincaid, Robert P Fishburne Jr, Richard L Rogers, and Brad S Chissom. Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel. Technical report, Institute for Simulation and Training, University of Central Florida, 1975.
6. Bart P Knijnenburg, Alfred Kobsa, and Hongxia Jin. Dimensionality of information disclosure behavior. *Int. Journal of Human-Computer Studies*, 71(12), 2013.
7. Byron Lowens, Vivian Genaro Motti, and Kelly Caine. Wearable Privacy: Skeletons in The Data Closet. In *Proc. of the IEEE International Conference on Healthcare Informatics (ICHI)*, 2017.
8. Anthony Morton. *Individual Privacy Concern and Organisational Privacy Practice – Bridging the Gap*. PhD thesis, University College London, 2015.
9. Anthony Morton and M Angela Sasse. Desperately Seeking Assurances: Segmenting Users by their Information-Seeking Preferences. In *Proc. the IEEE Annual International Conference on Privacy, Security and Trust (PST)*, 2014.
10. Patrick Murmann. Supplementary material. `https://murmann.hotell.kau.se/notification-preferences/`, last visited November 13, 2018.
11. Patrick Murmann and Simone Fischer-Hübner. Tools for Achieving Usable Ex Post Transparency: A Survey. *IEEE Access*, 5, 2017.
12. Pardis Emami Naeini, Sruti Bhagavatula, Hana Habib, Martin Degeling, Lujo Bauer, Lorrie Cranor, and Norman Sadeh. Privacy Expectations and Preferences in an IoT World. In *Proc. of the Symposium on Usable Privacy and Security (SOUPS)*, 2017.
13. Stefan Palan and Christian Schitter. Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, 17, 2018.
14. Eyal Peer, Laura Brandimarte, Sonam Samat, and Alessandro Acquisti. Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70, 2017.
15. Prolific Academic Ltd. Prolific. `https://www.prolific.ac/demographics`, last visited August 27, 2018.
16. Statista. Number of connected wearable devices worldwide from 2016 to 2021. `https://www.statista.com/statistics/487291/`, last visited June 28, 2018.
17. Isabel Wagner, Ying He, Duska Rosenberg, and Helge Janicke. User interface design for privacy awareness in ehealth technologies. In *Proc. of the IEEE Annual Consumer Communications & Networking Conference (CCNC)*, 2016.
18. Alan F Westin. Social and political dimensions of privacy. *Journal of social issues*, 59(2), 2003.
19. Allison Woodruff, Vasyl Pihur, Sunny Consolvo, Lauren Schmidt, Laura Brandimarte, and Alessandro Acquisti. Would a privacy fundamentalist sell their DNA for $1000... if nothing bad happened as a result? The Westin categories, behavioral intentions, and consequences. In *Proc. of the Symposium on Usable Privacy and Security (SOUPS)*, 2014.