Data Sharing in Mobile Apps — User Privacy Expectations in Europe

Nils Quermann
Chair for System Security
Ruhr-Universität Bochum
Bochum, Germany
nils.quermann@ruhr-uni-bochum.de

Martin Degeling
Chair for System Security
Ruhr-Universität Bochum
Bochum, Germany
martin.degeling@ruhr-uni-bochum.de

***Abstract*—Although privacy on mobile apps has increased through tougher permission guidelines and new data protection legislation, developers still make profit of sharing data of and about their users with advertisers or other services. In light of the continuous debates on privacy and mobile apps we wanted to understand users' expectations of privacy when it comes to different app types and data sharing practices. We conducted a survey on MTurk ($N = 103$) to learn more about privacy awareness and expectations. We found that 9 out of 10 users suspect that they are being tracked across different mobile applications and that they have a good understanding of what information is necessary for certain types of mobile applications. 65 % of participants also expect that, on average, 2-6 different services receive their personal information and 93 % think that at least a minority of mobile applications share their personal information without their consent.**

***Index Terms*—privacy, expectations, user study, mobile application**

1. Introduction

Around the world, billions of people use mobile applications on a daily basis. The revenue model of most apps is based on advertising for which they collect personal data that is shared with the advertising and tracking industry. In 2018, the global ad revenue surpassed $100 billion for the first time with an ever rising trend. [1] This business model is enabled by *third-party libraries* that can be used for various purposes: From analyzing software problems to collecting data to improve their service as well as behavior tracking to generate user profiles and, possibly, predict their interests or behavior. App developers embed these trackers in their applications, enabling advert and tracking services to create and enrich user profiles in the background. Previous research has shown that these libraries are often used carelessly without considering user privacy concerns [2].

For users, however, it is oftentimes not easy to understand *what* information is collected by *whom* and *why*. These questions are difficult to answer since apps do not openly communicate what specific data they collect, even if they have a privacy policy [3]. But even if privacy practices are disclosed, users often lack the knowledge to understand what information is collected about them or the time to fully read and understand apps' privacy policies. This raises the question: *what expectations do users have about their privacy regarding mobile applications?*

Past research has focused primarily on either the privacy understanding of individuals or the implementation of privacy regulations in mobile applications. Studies on user expectations are 7 or 8 years old. In this paper we focus on the awareness of smartphones users for information sharing through mobile applications to update the knowledge in this area. What are their expectations of data sharing by and between apps and how does this relate to their general privacy concerns? To answer this question we asked 103 workers on Amazon's *Mechanical Turk* (MTurk) platform to answer an online survey. We use the standardized *Internet Users' Information Privacy Concerns* (IUIPC) questionnaire to assess participants general internet privacy concern and report their agreement with IUIPC statements that are rarely reported in other studies in detail and can confirm abstract scores provided by other researchers.

Our study shows that about 9 out of 10 users suspect that they are being tracked across different mobile applications. We also highlight that users understand which personal information may be necessary for certain types of mobile applications, although, they expect certain categories (i. e., gaming, shopping, entertainment and social media apps) to collect more than the average amount of data. The majority of participants (65 %) also expects that, on average, 2-6 different services receive their personal information. Almost all participants (93 %) think that at least a minority of mobile applications share their personal information without their consent. This highlights a rather pessimistic view about mobile applications from our participants, however, there were also individuals that felt less concerned about mobile applications collecting their personal information. In fact, some of them welcomed apps showing them personal adverts, based on previously collected data. Lastly, we could not find any proof that privacy concerned participants made different app choices and, e. g., chose privacy friendly apps.

1.1. Related Work

Previous research shed light on the practices of user tracking and the collection of personal data through mobile applications. Multiple studies have analyzed mobile applications to gather information about data receivers. Razaghpanah et al. [4] developed a privacy-enhancing app called *Lumen Privacy Monitor (Lumen)*. Lumen uses a

virtual private network (VPN) to analyze mobile application's network traffic, so that users can monitor which app may leak personal information about them to some third party. With the help of all Lumen users, Razaghpanah et al. were able to identify new advertising and tracking services that have previously been unknown by popular blacklists used against tracking and advertising. Another key finding of their research is that sharing data with subsidiaries and third parties does not seem to be the exception, but rather the norm.

In another study from 2015, Zang, Dummit, Graves, Lisker and Sweeney [5] analyzed 110 popular Android and iOS applications with the goal to find out "who knows what about me". In their analysis, Zang et al. set up some network proxy which captures all the (apps') traffic. The traffic then got analyzed to see what kind of data got transmitted — and to which service. As a result, they found that many mobile apps transmitted sensitive user data to third-parties. Especially the user's location, email, and name were transmitted most frequently.

Besides this technical studies on apps, usable privacy and security researchers have studied the perception of privacy and apps. Already in 2012 Felt et al. [6] found that many users do not understand the permission model of android. Lin et al. [7] showed that expectations are an important element of privacy in regards to mobile applications and, further, that expectations of privacy practices are directly related to feeling comfortable with the data collection. Later Shklovksi et al. [8] showed that many users think tracking in apps is "creepy" and should be forbidden, but do not deal with the services they are uncomfortable with as there are few alternatives.

MTurk is a platform where tasks are assigned to, so called, workers who will work for some monetary incentive. While this platform is originally designed to distribute small tasks like annotating images to freelance workers, it has also been used in the past to gather participants working on questionnaires. Redmiles et al. [9] compared the results gathered from MTurk, a census web panel, and telephone interviews. Their questions aimed at participants' self reported privacy and security knowledge. Redmiles et al. conclude that answers gathered through MTurk are more representative than the ones collected from the census web panel. Furthermore, MTurk answers were quite similar to the general population, at least for participants younger than 50 years or with at least some college education. Lastly, MTurk workers' answers were consistent over time and did not change based on major news events.

Malhotra et al. [10] developed a framework on the dimensionality of internet users' information privacy concerns. With this framework, they developed a causal model between IUIPC and users' behavioral intentions regarding sharing personal information. Malhotra et al. analyzed the results of 742 respondents with the help of this framework. They conclude that their causal model is well suited to describe a lot of variance in behavioral intention.

2. Method

We developed a survey with the goal to understand privacy expectations of users regarding their mobile applications. To tell whether these expectations differ from their general privacy expectations, we also determined an individual baseline using the *IUIPC* questionnaire.

In the following section we detail our recruitment process and the questionnaire's structure.

2.1. Participant Recruitment

We target citizens of the European Union, as it has rather strict and consistent privacy regulation, due to the regulations of the *General Data Protection Regulation (GDPR)*. We argue this yields more consistent results than we may receive when we mix participants with different privacy concerns due to different privacy cultures.

In order for us to characterize our participant pool, we opted for *MTurk* as a tool to gather participants for our study. We defined four key requirements that workers had to fulfill to participate in our study. Workers were required to

1) live in a country of the European Union,
2) not retake the survey multiple times,
3) have a previous *approval rate* of above 97 %, and
4) have successfully completed 5 000 HITs.[1]

In order to estimate the time needed to answer all questions, we conducted a pre-test. This guided as a measure as to how much money workers should be rewarded with. In our pre-test participants took 8 minutes and 30 seconds, on average, to answer all of our questions. We used 12€ as hourly wage, thus, paid workers 1.70€ for answering our questionnaire. The survey tool *LimeSurvey* was hosted locally at our university's IT center to have full control over the data collection.

2.2. Survey Structure

The four parts of the survey deal with the participants' general privacy understanding, positions towards mobile data sharing, and expectations towards processing of their personal information. At the end of the survey, participants had to answer some demographic questions and were given the chance to leave feedback or comments.

The reason behind this structure is that we first wanted to measure the users' general privacy expectations with the help of the IUIPC questionnaire. Afterwards, we targeted our questions more towards mobile applications so that we can compare both results for possible differences.

The full questionnaire that was used for the user study is shown in Appendix A.

2.2.1. Disclaimer. In the very first part of the survey, we informed participants about the aim of the study and which information about them is collected. Additionally, we gave information about ourselves and how participants may contact us if they have any further questions or concerns. Afterwards, participants were required to answer four questions, e. g., denoted as Q1, stating that they

1) have read and understood the previous information (Q1),
2) are 18 years or older (Q2),
3) live outside of Europe (Q3), and

1. Both, the 97 %and 5 000 HITs were suggested by this blog post for academic requesters on MTurk: https://turkrequesters.blogspot.com/2012/09/tips-for-academic-requesters-on-mturk.html

4) want to participate in this research and continue with the study (Q4).

Of particular interest is Q3 as we are only interested in participants living in Europe. Thus, the negation seems odd, however, rephrasing the question requires conscious participants to select *no* as the "correct" answer. If participants mindlessly click *yes* for every question, without reading all the information, they will likely fail this kind of *attention check*.

Only if all four questions have been answered correctly, participants were allowed to answer the remaining questionnaire. Otherwise, they were navigated directly to an alternative end of the survey stating that they were sadly not eligible.

Our research institution does not have a formal IRB process. Therefore, we designed our survey following guidelines of the data protection office. For example, all participants were shown a recruitment screen that included data protection information following the GDPR transparency rules. We also minimized the personal data we stored, gave participants the option to contact us or the data protection office in case they had any concerns, and allowed them to discontinue the study at any given time. We also enabled the "prefer not to answer" option for all questions.

2.2.2. General Privacy Expectations. To measure participants' general privacy expectations, we used the most common dimension of the IUIPC framework. We will compare this to later questions and aim at telling whether participants are more concerned when it comes to mobile apps, or vice versa, maybe participants are (overly) critical with mobile apps.

The IUIPC questionnaire has been used in other studies, for example by Naeini et al. [11] and Habib et al. [12], and is suitable as a baseline to compare the privacy awareness of our participants with those of other studies. The IUIPC has also been evaluated with young German users, with respect to location based services [13].

The framework measures three different dimensions: *Control* and *awareness* over data collection as well as the concernmet of the *collection* itself. Each statement of these dimensions sets a slightly different focus on the participants' information privacy. The *control* dimension seeks to find out how much control participants want to exercise over their personal information. In contrast to the first dimension, the second one is of passive nature, that is, the extent to which participants want to be *aware* about a company's privacy practices. The last dimension estimates how concerned participants are about others possessing personal information about them relative to some perceived benefits.

Participants were presented each of the statements, along with a seven-point likert scale, ranging from *strongly disagree* to *strongly agree* with a neutral point in between. For the sake of readability, we show the exact phrasing of each statement in the Appendix (see Q5, Q6, and Q7).

2.2.3. Privacy Preferences towards Mobile Data Sharing. In this part of the questionnaire we asked participants about their perception of privacy with respect to mobile data sharing. In Table 1 we present the statements, where participants had to mark their level of agreement, again, based on a balanced 7-point likert scale.

TABLE 1. PARTICIPANTS AGREEMENT REGARDING PRIVACY OF MOBILE APPLICATIONS.

POS1	"In the past, I preferred one app over the other, if it shared less personal information."
POS2	"On average, apps request an adequate amount of permissions."
POS3	"Concerns regarding data sharing play a role when choosing an app."
POS4	"On average, apps share an adequate amount of personal information."
POS5	"I expect apps to share only personal information that are required for their purpose."

Next, we asked participants what information they think apps of certain categories are likely to collect. We used five app categories, as well as nine personal information, forming a 5×9 matrix. If participants expect that a certain category collects this kind of personal information, they had to check the corresponding box.

As app categories, we listed *Games*, *Entertainment*, *Tools*, *Education*, and *Personalization*. These categories, are based on the top five most popular app categories.[2] We ordered the categories by how many apps with more than 50 000 downloads are within the category. From this list, we selected the top five categories.

The nine categories of personal information that may be collected by apps were, partly, based on the Android permission model. These permissions are: Access to the individual's *location*, its *messages*, the smartphone's *microphone* or *Camera*, or *files stored* on it. In addition, we added some more generic information types. For example, (not unique) *user IDs*, and information about the user's *behavior*. The *General Data Protection Regulation (GDPR)* defines specific personal information as "special categories of personal information" or *sensitive* information, which are, thus, included as well.

The ninth option was that participants do not expect any of these information to be collected. See Figure 1 for the list of all categories and information.

In the following question Q10, we asked participants to estimate how many third-party services receive personal information from a single app. The results to this question could indicate how aware participants are about apps sending (personal) information to other parties. The bin sizes (i. e., "1 (only the app service)", "2-3", "4-6", "7-10", "More than 10"), are oriented on the study conducted by Zang et al. [5]

In the final question Q11 of this section, we asked participants what they believe how likely they are being tracked across different apps. Possible answers were a balanced five-point likert scale ranging from *very likely* to *very unlikely*.

2.2.4. Processing of Personal Information. This section of the survey dealt with the users' concerns regarding apps collecting and/or sharing their personal information with others.

Question Q12 was a binary *yes* or *no* question, asking if participants expect any app category to collect more

2. According to statistics about Android Play Store categories in June 2019 (https://www.appbrain.com/stats/android-market-app-categories)

personal information than others. An additional question (Q13) was shown, asking why they think that way and if they thought of a specific category of personal information, when selecting *yes*.

In Q14, participants had to estimate how many apps share information without their consent. Answers were a five-point likert scale that ranged from *none* to *all*. This question should give us an idea as to how pessimistic participants think about mobile applications and their own privacy.

Q15 asked about participants' general level of concern regarding apps sharing their personal information with third parties. With this question, we forced participants to commit to either side of the *concerned* or *unconcerned* spectrum by omitting a neutral answer. This should indicate whether participants are generally more concerned or unconcerned. In the followup question Q16, participants were asked (but not forced) to detail why they selected their answer.

In the last question Q17 of this section, participants had to list five apps they most frequently use. This result can indicate how well our set of participants match the global app popularity (cf. categories of Q9). Since this cannot accurately be measured by participants without additional tools, the results, however, can only guide as an indicator.

2.2.5. Demographics. The survey concluded with questions about demographic information and an attention check.

As demographic information we asked participants to provide their gender, age, level of education and employment status.

3. Results

The survey-task was published on MTurk in September 2019 and we gathered responses from 103 participants, of which we consider 95 as valid.

On average, workers spend 8 minutes and 17 seconds completing the survey, being slightly faster than the results of our pre-test suggested.

3.1. General Privacy Expectations

In the following, we detail the results found for each of the three *IUIPC* dimensions *control*, *awarenesss*, and *collection*, before we finally summarize our findings.

When we speak of the *mean* and *median* answer, we first mapped the answers *strongly disagree* to *strongly agree* on a equally spaced scale, ranging from 1 to 7. Afterwards, we computed the *mean* and *median* based on these values.

To simplify interpretation, we *binned* answers into the *agreeing* and *disagreeing* category, according to their level of (dis-) agreement. Participants that selected *neither agree nor disagree* are reported as *undecided* in the following sections. In order to receive a total of 100 %, we removed responses stating to "prefer not to answer".

3.1.1. Control. In the *control* dimension, participants were mainly *agreeing* with the statements, that is, 86 % to 89 % of participants agreed, with an average of 87 %. The results also show that the *median* for all questions is 6, corresponding to the answer *agree*. In Table 2 we present the (dis-) agreement for each question individually.

TABLE 2. (DIS-) AGREEMENT WITHIN CONTROL DIMENSION.

	CON1	*CON2*	*CON3*
Agreeing [%]	86	86	89
Disagreeing [%]	11	7	7
Undecided [%]	3	6	3
Mean	5.75	5.69	5.63
Median	6	6	6
Std Dev	1.48	1.42	1.32
N	93	95	95

We see that *CON1* ("[...] privacy is really a matter of consumers' right to exercise control [...] over [...] how their information is collected [...].") got one of the highest level of disagreement among all statements. With such a (comparably) high level of disagreement, this could indicate that (some) participants do not want to control the use of their data on their own. While 11 % is still a rather small portion of the participants, it is still one of the highest rates of disagreement among all *IUIPC* questions

Either participants think control is unnecessary or they disagree for other reasons, e.g. they would prefer automated over manual controls. or their disagreement could also indicate that they think privacy should be mandated and not optional.

Another interesting observation is the different distribution of agreeing answers in *CON3* ("[...] privacy is invaded when control is lost [...] as a result of marketing transaction"). Even though most participants agree, we see both a smaller mean and standard deviation. This shows that participants did not agree with this statement as strong as with the previous two. This result may support the thesis that participants would give up some of their privacy in order to receive some (monetary) incentive. Previous studies, for example [14], found that many Americans would trade personal information if they were offered a lucrative deal.

3.1.2. Awareness. In the *awareness* dimension, participants *agreed* or *strongly agreed* with all three statements on average. The level of agreement ranges between 86 % and 92 %. Interestingly, more than half of the participants stated that they would *strongly agree* (median of 7) with *AWA1* ("companies [...] should disclose the way the data are collected, processed, and used.") For all other statements the median level of agreements was *agree* (median of 6, see Table 3).

AWA2 and *AWA3* reflect whether companies should disclose their privacy practices in a clear and conspicuous way and whether it is important to participants to be aware and knowledgeable about how their information will be used. Especially for the former statement, we see high levels of agreement. However, previous studies have shown that users rarely read privacy policies. For example Utz et al. [15] reported that only 0.5 % of visitors click on privacy policies (embedded in a cookie consent notice).

A reason for this mismatch is likely the complexity of modern privacy policies. As McDonald et al. [16] have shown that internet users would spend years of their lives reading privacy policies, and more recent studies have

TABLE 3. (DIS-) AGREEMENT WITHIN AWARENESS DIMENSION.

	AWA1	*AWA2*	*AWA3*
Agreeing [%]	89	92	86
Disagreeing [%]	5	6	11
Undecided [%]	5	2	3
Mean	6.24	6.0	5.91
Median	7	6	6
Std Dev	1.42	1.49	1.56
N	93	95	95

shown that the length of privacy policies further increased. Degeling et al. showed that the length of privacy policies increased by 41 % between 2016 and 2018 alone [17].

3.1.3. Collection. Like the previous dimensions, participants mainly agreed with the statements regarding the *collection* of personal information. Just like the first dimension, the median for *all* statements is *agree* (median of 6).

In terms of agreement, however, we see that all statements, except *COL2* ("I sometimes think twice before providing [personal information]") have the lowest level of agreement across all three *IUIPC*-dimensions, as shown in Table 4 (the level itself, however, is still rather low). *COL1*("It usually bothers me when online companies ask me for personal information") has the lowest level of agreement, but also happens to have the highest level of disagreement.

TABLE 4. (DIS-) AGREEMENT WITHIN COLLECTION DIMENSION.

	COL1	*COL2*	*COL3*	*COL4*
Agreeing [%]	80	86	85	84
Disagreeing [%]	12	9	5	12
Undecided [%]	8	5	10	4
Mean	5.53	5.72	5.89	5.77
Median	6	6	6	6
Std Dev	1.63	1.52	1.47	1.69
N	95	94	92	94

Comparing *COL1* with *COL3* ("It bothers me to give personal information to so many companies"), we see another surprising tendency: While the former question asks if participants are bothered by companies *asking for* personal information, the latter one asks if participants are bothered *to give* personal information to so many companies. When we compare these two statements, we may conclude that some participants do not mind companies asking for their personal information and are generally willing to give (some) of them.

3.1.4. Summary. Across all three dimensions, we observe high levels of agreement with the *IUIPC* statements. On average, we see the highest level of agreement in the *awareness* dimension and the lowest level of agreement in the *collection* dimension (compare *Avg.* of Table 2, 3, and 4).

The differences are not very big, however, it seems participants generally want to be more knowledgeable about companies' privacy practices. At the same time they seem less concerned about giving their personal information out — as long as they are informed and therefore can retain a sense of agency. In the questions about the processing of personal information, we will see some responses that support this hypothesis, for example the concernment about data (mis-) use of third parties (see section 3.3). This result matches the study of Lin et al. [7] who observed that participants are willing to trust an mobile application more easily when the purpose of the data collection is clear.

In Table 5 we present the average response for each *IUIPC* dimension, along with the average of other studies from Emami-Naeini et al. [11] and Habib et al. [12]. Comparing the studies, we see that our participants were similarly or slightly less concerned regarding their information privacy.

TABLE 5. COMPARISON BETWEEN OTHER STUDIES.

Study	Control	Awareness	Collection
Emami-Naeini et al.	5.95	6.44	5.79
Habib et al.	5.8	6.2	5.6
This study	5.69	6.05	5.73

3.2. Positions towards Mobile Data Sharing

In the second part of the survey participants were asked about their use of mobile apps and privacy concerns implied by the usage.

3.2.1. Statements. The answers to Q8a show that nearly five of seven participants (compare Table 6) have chosen one app over the other in the past because it shared less personal information. This indicates that participants are not only privacy concerned but also act accordingly.

POS2 stated that apps request an adequate amount of permissions, on average. The opinions about this statement are somewhat divided in half.

TABLE 6. (DIS-) AGREEMENT WITH POSITIONS TOWARDS MOBILE DATA SHARING.

	POS1	*POS2*	*POS3*	*POS4*	*POS5*
Agreeing [%]	73	47	79	51	87
Disagreeing [%]	16	39	12	32	7
Undecided [%]	12	14	9	18	5
Mean	2.76	3.83	2.69	3.73	2.15
Median	2	4	2	3	2
Std Dev	1.85	1.75	1.55	1.63	1.30
N	95	94	95	95	95

79 % of the participants agree to *POS3* (data sharing concerns play a role when choosing an app). Comparing this to the results of *POS1* (participants chose more privacy friendly apps in the past), we see a significant correlation ($p < 0.01, r = 0.72$). This allows the conclusion that most participants that are concerned about data sharing also base their app choices on the amount of personal information the app shares. However, there are some individuals that are concerned, but are not influenced in their decision making.

Interestingly, we could not find any (significant) correlation between *POS2* and *POS3*. We suspect that selecting more privacy friendly apps is independent of participants' feeling whether the average app collects an adequate amount of personal information.

Similar to *POS2*, statement *POS4* split our participant pool in half: A majority of 50 %agreed with this position

while the rest either disagreed (41 %) or were undecided (9 %). This indecisiveness may be a result of participants' lack of knowledge to what extend mobile applications actually share their personal information.

We see another significant ($p < 0.01$, $r = 0.66$) correlation between *POS2* and *POS4*. This suggests that participants who think apps use an adequate amount of permissions do also think that they use only necessary information.

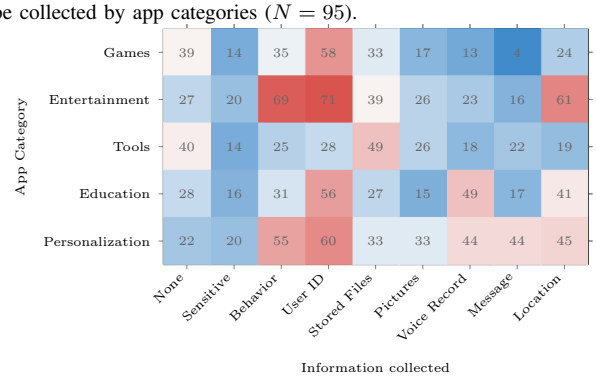
The final statement received the highest level of agreement across all five statements. Therefore, we conclude that participants generally *expect* apps to share only (personal) information necessary for their purpose.

We suspect one ambiguity with the phrase "to expect", that is, some participants answered what they thought apps *should* or *need to* collect while other participants thought what apps *do* collect.

Looking at the previous responses in the *IUIPC* dimensions, we see a significant correlation ($p < 0.01, r = 0.43$) between *CON1* and *POS1* as well as *POS3*. We recall that *CON1* stated that online privacy is a matter of consumer's right to exercise control over how their information is collected, used, and shared. Thus, we confirm our previous speculation that participants actually exercise their control by choosing mobile applications that are less privacy concerning.

3.2.2. Information collected by different app categories. The heatmap shown in Figure 1 presents participants' expectations about what personal information is collected by apps in different categories. A red tinted cell corresponds to more participants selecting it, while a blue tinted cell translates to fewer participants selecting this cell.

Figure 1. Number of participants expecting certain information types to be collected by app categories ($N = 95$).



It appears that only few participants expect apps from any of the categories to collect *sensitive* information. Further, apps from all categories, except for *Tools*, are expected to collect some *User ID*. The latter one can be explained to some extent, because a *User ID* is probably required in order to personalize the content. Users would be dissatisfied if their favorite game suddenly lost their progress or their shopping cart got empty after restarting the app. On the other hand, *tools* — like a flashlight — do most likely not need to recognize the user.

The results for the *Gaming* and *Entertainment* category are somewhat surprising. We can observe the highest peaks — on either end of the scale — for those two categories. For example, most participants expect that *Entertainment* apps collect not only some *User ID*, but also data about the user's *Behavior* and *Location*. On the other end of the scale, we see very few participants expecting *Gaming* apps to collect their private messages.

Participants expect *personalization* apps (e.g. a keyobard app) to collect the most information. While collecting behavioral data and user IDs seems necessary for personalization, accessing voice records and messages is not directly required for personalization. Given the examples for this category (*Google Keyboard* and *Samsung Launcher*), it is not clear why location data would be required for personalization, but participants might have experience with location based recommendations (e.g. in Google Maps) that are related to this category.

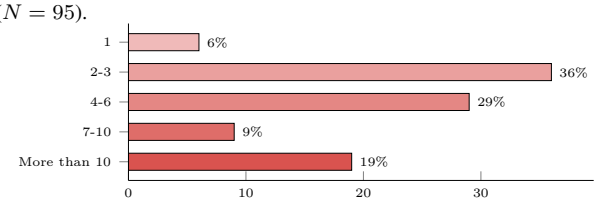
On the opposite end, participants seem to expect from *Games* and *Tools* to collect no information at all (except some *user ID* for *Games*). This makes somewhat sense, as the main purpose of these apps is to play, for example, some level of *Candy Crush* or to turn on the smartphone's flashlight. However, this trend may be again due to the ambiguity of the word "expect" (cf. 3.2.1).

Another finding is that we saw no significant correlation between participants answers of the *IUIPC* questionnaire and the amount of information collected (i. e., how many boxes were checked by participants). This suggests that participants' expectations of the amount of collected personal information is not influenced by their general privacy expectations.

Zang et al. [5] results suggest that participants expectations are quite different, especially for the *gaming* category: While participants do not expect sensitive information to be sent, Zang et al. showed that they do share this kind of information. Gaming apps do not seem to send this information to the app's service provider, but to some third party tracking service. However, participants expectations were not always wrong as gaming apps do not send location data, as expected by them.

3.2.3. Number of Information Recipients. Figure 2 shows participants' best guesses on how many services receive their personal information. The majority of participants, i. e., 65 %, expect that 2-6 services receive their personal information, on average.

Figure 2. Average number of services personal information are sent to ($N = 95$).

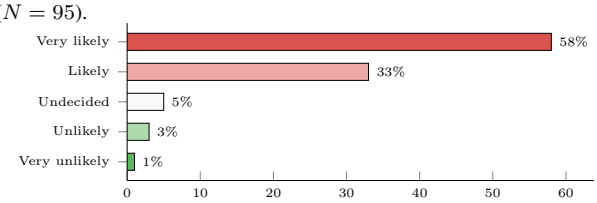


However, the data does not appear to resemble some normal distribution. A large subgroup of participants expect more than 10 services to receive their personal information. Diving deeper in the gathered data, we see that these individuals also seem to be generally more pessimistic about their information privacy supporting a view that many users have become "privacy resigned" [18].

3.2.4. Tracking across different Apps. When asked whether participants expect to be tracked across different

apps we also got rather pessimistic responses. About 90 % of the participants stating that it would be *likely* or even *very likely* that their actions on one app influence some other app (see Figure 3).

Figure 3. Expected likelihood of being tracked across different apps ($N = 95$).



We even see that nearly 58 % thought that this would be a *very likely* behavior of apps. On the other side, only about 4 % thought that this would be (very) unlikely. This may indicate that participants have a strong disbelieve in apps. Reasons for this may be personalized ads that probably a large portion of participants have already encountered in the past. Data for this personalization must be collected from somewhere, for example, their shopping app or the search history of another app.

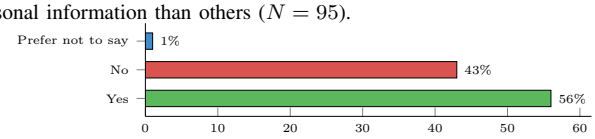
Past research has shown that end users often do not read the privacy policy of (mobile) applications [15]. This raises the question whether participants may have agreed to this transmission in the app's privacy policy without reading it. A more precise question would be "how likely are personal information *unwillingly* shared?".

3.3. Processing of Personal Information

In the second last section, we asked participants about their beliefs regarding the procession of their personal information.

3.3.1. Suspicious App Categories. We wanted to know, if participants expected apps from certain categories to collect more information than others and, if yes, from which category. The results, see Figure 4, are somewhat mixed and differ by only 12 %.

Figure 4. Expectations whether some app categories collect more personal information than others ($N = 95$).



With the previously (pessimistic) responses, it is somewhat surprising to see that many participants answered *No*. We could, however, *not* observe that participants answering *No* were generally less concerned about their information privacy.

One explanation could be that participants believed that all app categories share an equal (i. e., too high) amount of personal information, thus, not believing that some apps share more of them. On the other hand, we have already seen that participants expected apps from certain categories (*Personalization* and *Entertainment*, compare Figure 1) to collect more information than others.

Additionally, we asked participants about their reasoning behind answering *Yes* and what app category they were thinking about. In the Appendix Table 8 we grouped some of the user responses by app category.

Participants frequently stated that they expect *gaming*, *entertainment*, *social media*, and *shopping* apps to collect more information than the average app category. This slightly contradicts the previous results (compare Figure 1) where we have seen which information participants expect to be collected from specific app categories. This supports the suspected ambiguity, i. e., participants do not see a reason that these apps *should* collect more data than others — but expect that these apps still *do* collect more personal information. We see that participants' pessimistic view on those kind of apps seem to be justified. A previous study by Zang et al. [5] revealed that, indeed, *social media* and *shopping* apps send more personal information than, for example, *gaming* or *business* apps.

Some participants mentioned personalized ads and, as a result, the (companies) revenue as reasons for collecting more information. Other participants stated that certain apps may need specific information in order to deliver their functionality. This highlights that participants are aware that some apps actually need personal information for their functionality while others use their personal information for profit.

3.3.2. Transmission without Consent. It appears that participants are unsure how many apps share personal information without their consent, as Figure 5 suggests. On average, however, participants believe that *about half of them* do send information without their consent. Only one of three participants believed that only *a minority* of apps or *none* shared personal information without their consent.

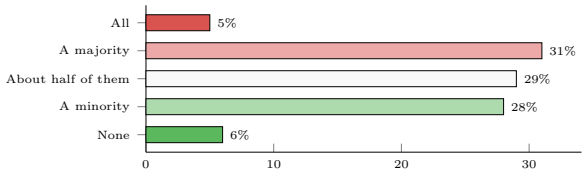


Figure 5. Expectations about data sharing without user consent ($N = 95$).

On the other side, we see an overwhelming majority of 93 % of participants that could imagine that, at least *a minority* of apps, share personal information without prior consent. This highlights once again the quite pessimistic view of the participants and indicates that — even one year after the *GDPR* took effect — participants are highly skeptical about apps obeying common law. The *GDPR* explicitly prohibits the transmission of personal information without consent of the affected person. With this in mind, participants may be questioning the effectiveness of this regulation.

In a later text response, we even find a reason why participants might be so skeptical. One participant stated:

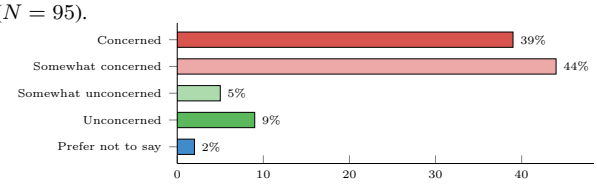
> "I think the value of the data makes it very lucrative, and there isn't enough regulation or monitoring to prevent this kind of behaviour."
> – Participant #75

These expectations seem to be inline with research that has shown that about 71 % of [19] apps do not have

a privacy policy although they process personal data and that 17 % might share data without consent.

3.3.3. Concernment. After we asked participants about the likelihood of personal information being unwillingly shared, we followed up with the question of how concerned they are about third parties receiving their personal information. In Figure 6, we present the participants' responses.

Figure 6. Level of concernment about data sharing with third parties ($N = 95$).



Even without forcing participants on one end of the scale (we omitted a neutral answer), the tendency of being concerned might still have become clear. More than four fifth stated that they were at least somewhat concerned that third parties receive their personal information. The 14 participants stating to be unconcerned, however, agreed with most of the *IUIPC* questions. This raises the question as to why participants feel less concerned when it comes to apps sharing their personal information with third parties.

In the second part of the question, we gave participants the opportunity to explain their reasoning behind their level of concernment. In the Appendix Table 9, we present some of the participants' answers.

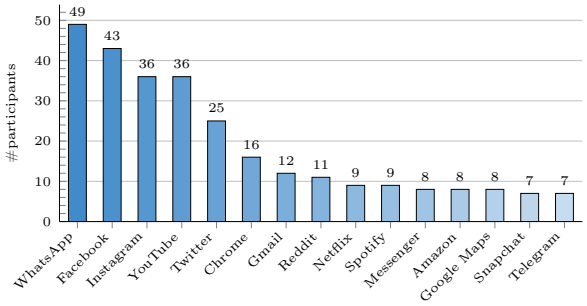
Essentially, we found four reasons why participants were either concerned or unconcerned. The main reasons for concernment were not knowing *where* and *how* the personal information is being used. In one of the *IUIPC* questions (*AWA1*, compare Table 3) we already saw that more than half of the participants *strongly* agreed with the statement that it is important for them to know how their personal information will be used.

Interestingly, however, participants did not respond that strongly in *CON1*. This may be due to the wording of the question, which also asked about how information is collected and shared — not only used. Another interpretation is that participants are generally more concerned how personal information from their smartphone is being used than other personal information (e. g., due to private photos).

The two most common reasons for participants being *unconcerned* were believing it would *cause no harm* or they simply *did not care* to share personal information. One participant, who stated to not care, even explained that personalized ads are something good for her. She preferred to *know* beforehand when her personal information is shared, however. If apps would sell them without informing her, it would be a no-go. Other participants thought that their data is not that important or valuable for third parties.

3.3.4. Popular App Choices. With all previous questions, where participant stated to be rather privacy concerned, it is surprising to see that the top three apps all belong to the *Facebook Inc.* One would expect that past (privacy) scandals would have caused serious damage to the reputation of Facebook and, ultimately, users leaving their platform(s). Especially after the *Cambridge Analytica* scandal[3], where millions of Facebook users were affected, it is surprising to see privacy concerned individuals still using apps from this very same company.

Figure 7. Self-reported most used apps by participants ($N = 95$).



We might expect more privacy unconcerned participants to use one of the top five apps (i. e., *WhatsApp*, *Facebook*, *Instagram*, *YouTube*, and *Twitter*). However, we did not find any evidence or correlation supporting this claim. The paradox that people choose these apps over their privacy concern is probably a result of the network effect of these services. Oftentimes using these apps is required to participate in social (online) life, regardless of the individual's concerns (compare section 3.3.1).

These services are among the — if not *the* — leading ones in their respective category, making it difficult for individuals to switch to more privacy friendly services, especially for messaging and social media services, when all their peers would need to switch as well (e. g., peer pressure) in order to successfully avoid them.

Interestingly, participants seemed not to use apps from the *personalization* category very frequently. Responses to Q9 showed that participants expected apps from this category to collect their location, messages, and behavioral data. It is unclear whether participants consciously decided against apps from this category, to not share these information or whether there is a bias in reporting, as participants may not be aware that they are using apps from this category frequently. For example, a custom keyboard or app drawer might be used almost all the time, but participants might have forgotten that they (or someone on their behalf) installed them.

3.3.5. Attention Check. Before asking demographic questions, we used one last *attention check* to estimate how thoroughly participants have read the previous questions. Inside a rather long task description, we hid the information that this question is an attention check and the correct answer would be *Ebay*. If participants would only read the bottom of the description, they would think that their task would be to select the company that collects the most information in their opinion.

We see that about 41 % of the participants selected the wrong answer. This is a surprisingly high number of wrong answers, thus, we assume that fatigue has occurred or the attention check is not suited for its task. Therefore,

3. https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html, as of June 24, 2020

we did not *immediately* exclude participants from the results if they answered this question wrong. Instead, we *additionally* considered other factors, for example, whether participants

- answered the previous attention check wrong,
- did not respond in text fields (or wrote some nonsense that did not suite the task),
- submitting a wrong *survey code*, or
- sped through all tasks

before marking their responses as invalid.

Combining all these characteristics, we removed nearly 8 % of the responses. This way lower acceptance rate made more sense, since one of our characteristic from our MTurk participant pool had been a high ($> 97\,\%$) approval rate from previous HITs.

3.4. Demographic

The demographic questions show, that participants were mainly male (77 %), rather young (90 % younger than 45 years), and good educated (63 % with Bachelor's degree or higher), as shown in Table 7.

TABLE 7. DEMOGRAPHIC BREAKDOWN OF OUR PARTICIPANTS ($N = 95$)

Gender		Education	
Female	23 %	Less than high school	4 %
Male	77 %	High school diploma	33 %
		Bachelor's degree	43 %
		Master's degree	18 %
		Doctorate	2 %
Age		**Employment**	
18-24	33 %	Employed full-time	42 %
25-34	42 %	Student	25 %
35-44	19 %	Self-employed	11 %
45-54	4 %	Employed part-time	8 %
55-64	2 %	Unemployed (looking for work)	11 %
		Unemployed (not looking for work)	2 %
		Prefer not to answer	1 %

We also asked about participants' employment status ("what is your current employment status?", Q22). The results show that the majority (i. e., about 75 %) of participants were either *employed full-time* or still *students*. Since we have 31 participants age 18-24 years old, it is of no surprise that we observer a relatively large group of 25 students. With the results of Vannette et al. [20] in mind, the somewhat large group of students may indicate that our questions got answered more conscious. Due to the nature of the subject, i. e., mobile applications, it makes sense to see some shift in the age spectrum towards the younger end.

With all previous questions, we could not find any significant correlation between participants answers and their demographic. Thus, privacy expectations towards mobile applications seems unrelated to, for example, the education level or age.

4. Conclusion

With all information gathered, we draw the following conclusions. Firstly we did not observe lower privacy expectations for mobile applications, compared to the participants' general privacy expectations (based on the three *IUIPC* dimensions *control*, *awareness*, and *collection*). We did, however, find a significant correlation between participants wanting the control over their personal information and exercising this control in the mobile ecosystem through conscious choices of more privacy friendly applications. There were, however, individuals that did not find control that important. In fact, they do not mind sharing their personal information as it has helped them in the past, due to interesting adverts. However, they demand knowledge to whom their information is shared with.

In general, participants had a rather pessimistic view. 93 % of them believe that at least some applications share their personal information without prior consent. Slightly less participants (i. e., 9 out of 10) believe that they are tracked across different apps. One reason for this mistrust in mobile application seems to be "the lack of regulation and monitoring", as one participant stated.

We could not answer the naturally arising question how justified participants' pessimism is. While we cannot generally proof that participants' expectations are justified, we have seen that participants estimated the number of third parties receiving their personal information correctly. Most participants answered that they expect 2-6 third parties to receive their personal information, which matches the actual number found by Zang et al. [5].

4.1. Recommendations

Our participants had a rather pessimistic view on their own information privacy regarding mobile applications. Although they were, in general, privacy concerned they did use many apps that collect more personal information than necessary and even expected apps to collect more. This privacy resignation is another call for action for developing application following "privacy by design" guidelines to regain users' trust. Privacy (and security) should not be the users' but the developers' primary concern.

Apps that follow this principles should then use transparency measures and simplified privacy controls to regain user trust. For example, apps should not only have a long and detailed privacy policy (that will most likely not be read [15]), but some sort of *quick facts* section that provides a understandable summary of the privacy policy.

Further, similar to Android's permission system, apps should ask whether they are allowed to send a certain personal information somewhere. This control should be designed to allow the transmission either temporarily or every time, offering ways to intervene in the data collection [21]. Unlike the Android permissions, this approach would be less abstract [6]. Instead, users would know exactly what the permission is used for.

Such an approach would give users the opportunity to know both, *what* information is send and *where* it is send to. Additionally, it would prevent apps from sending personal information without users' knowledge and consent. This would also give users the ability to exercise their rights given by the GDPR. There may, however, be still the problem that users are unaware of their rights. To enforce this new policy, apps could be curated before being published, similar to Apple's App Store. Apps are then curated in greater depth, revealing if they collect or share only the users' necessary personal information.

4.2. Limitations

In this study we asked participants about their privacy preferences and behavior with respect to mobile apps. Previous research has shown that the disclosure of privacy preferences is often not in line with actual privacy behavior and that these kinds of questions might result in a social desirability bias. For example, when users report that they have decided to use one app over the other for privacy reasons, they might not have actually done so. We can't rule out these inadequacies but argue that this will not change the overall direction of the answers.

4.3. Future Work

Our study detailed what expectations participants had regarding their information privacy. Naturally, future work should focus on whether current mobile applications actually meet these expectations. Especially interesting may be to see what personal information are collected by applications and to which services they are sent to. There have already been studies in the past, e. g., Zang et al. [5], that analyzed the behavior of mobile applications, however, most of these studies have been conducted prior the *GDPR* coming into force.

Conducting such a study also yields the answer if the differences between application categories, are as our participants expected (e. g., *Tools* collecting less personal information than *Entertainment* applications).

Our study focused on participants living in the European Union, as there are rather strict regulations with respect to the individuals information privacy, due to the *GDPR*. In a more general study, it could be interesting to see whether there are indeed different expectations between European and non-European citizens.

References

[1] IAB, "Iab internet advertising revenue report 2018," 2019, https://www.iab.com/wp-content/uploads/2019/05/Full-Year-2018-IAB-Internet-Advertising-Revenue-Report.pdf, as of June 24, 2020.

[2] Abraham H. Mhaidli, Y. Zou, and F. Schaub, ""We Can't Live Without Them!" App Developers' Adoption of Ad Networks and Their Considerations of Consumer Risks," Santa Clara, CA, 2019.

[3] S. Zimmeck, P. Story, D. Smullen, A. Ravichander, Z. Wang, J. Reidenberg, N. C. Russell, and N. Sadeh, "MAPS: Scaling Privacy Compliance Analysis to a Million Apps," *Proceedings on Privacy Enhancing Technologies*, vol. 2019, no. 3, pp. 66–86, Jul. 2019, publisher: Sciendo Section: Proceedings on Privacy Enhancing Technologies. [Online]. Available: https://content.sciendo.com/view/journals/popets/2019/3/article-p66.xml

[4] A. Razaghpanah, R. Nithyanand, N. Vallina-Rodriguez, S. Sundaresan, M. Allman, C. Kreibich, and P. Gill, "Apps, trackers, privacy, and regulators: A global study of the mobile tracking ecosystem," 01 2018.

[5] Z. J., D. K., G. J., L. P., and S. L., "Who knows what about me? a survey of behind the scenes personal data sharing to third parties by mobile apps," *Technology Science. 2015103001*, Oct 2015. [Online]. Available: https://techscience.org/a/2015103001/

[6] A. P. Felt, E. Ha, S. Egelman, A. Haney, E. Chin, and D. Wagner, "Android permissions: user attention, comprehension, and behavior," in *Proceedings of the Eighth Symposium on Usable Privacy and Security*, ser. SOUPS '12. Washington, D.C.: Association for Computing Machinery, Jul. 2012, pp. 1–14. [Online]. Available: https://doi.org/10.1145/2335356.2335360

[7] J. Lin, S. Amini, J. I. Hong, N. Sadeh, J. Lindqvist, and J. Zhang, "Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, ser. UbiComp '12. Pittsburgh, Pennsylvania: Association for Computing Machinery, Sep. 2012, pp. 501–510. [Online]. Available: https://doi.org/10.1145/2370216.2370290

[8] I. Shklovski, S. D. Mainwaring, H. H. Skúladóttir, and H. Borgthorsson, "Leakiness and creepiness in app space: perceptions of privacy and mobile app use," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '14. Toronto, Ontario, Canada: Association for Computing Machinery, Apr. 2014, pp. 2347–2356. [Online]. Available: https://doi.org/10.1145/2556288.2557421

[9] E. M. Redmiles, S. Kross, and M. L. Mazurek, "How well do my results generalize? comparing security and privacy survey results from mturk, web, and telephone samples," in *2019 IEEE Symposium on Security and Privacy (SP)*, 2019, pp. 1326–1343.

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

[11] P. Emami-Naeini, S. Bhagavatula, H. Habib, M. Degeling, L. Bauer, L. F. Cranor, and N. Sadeh, "Privacy expectations and preferences in an iot world," in *Proceedings of the Thirteenth USENIX Conference on Usable Privacy and Security*, ser. SOUPS '17. USA: USENIX Association, 2017, p. 399–412.

[12] H. Habib, J. Colnago, V. Gopalakrishnan, S. Pearman, J. Thomas, A. Acquisti, N. Christin, and L. F. Cranor, "Away from prying eyes: Analyzing usage and understanding of private browsing," in *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*. Baltimore, MD: USENIX Association, Aug. 2018, pp. 159–175. [Online]. Available: https://www.usenix.org/conference/soups2018/presentation/habib-prying

[13] M. Fodor and A. Brem, "Do privacy concerns matter for Millennials? Results from an empirical analysis of Location-Based Services adoption in Germany," *Computers in Human Behavior*, vol. 53, pp. 344–353, Dec. 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0747563215300066

[14] L. Rainie and M. Duggan, "Privacy and information sharing," *Pew Research Center*, 2016.

[15] C. Utz, M. Degeling, S. Fahl, F. Schaub, and T. Holz, "(Un)informed Consent: Studying GDPR Consent Notices in the Field," in *Proc. CCS*, ser. CCS '19. New York, NY, USA: ACM, 2019, pp. 973–990. [Online]. Available: http://doi.acm.org/10.1145/3319535.3354212

[16] A. M. McDonald, R. W. Reeder, P. G. Kelley, and L. F. Cranor, "A comparative study of online privacy policies and formats," in *Privacy Enhancing Technologies*, I. Goldberg and M. J. Atallah, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 37–55.

[17] M. Degeling, C. Utz, C. Lentzsch, H. Hosseini, F. Schaub, and T. Holz, "We Value Your Privacy ... Now Take Some Cookies: Measuring the GDPR's Impact on Web Privacy," in *Proc. NDSS 2019*. Internet Society, 2019. [Online]. Available: https://www.ndss-symposium.org/wp-content/uploads/2019/02/ndss2019_04B-2_Degeling_paper.pdf

[18] N. A. Draper, "From Privacy Pragmatist to Privacy Resigned: Challenging Narratives of Rational Choice in Digital Privacy Debates," *Policy & Internet*, vol. 9, no. 2, pp. 232–251, Jun. 2017. [Online]. Available: http://doi.wiley.com/10.1002/poi3.142

[19] S. Zimmeck, Z. Wang, L. Zou, R. Iyengar, B. Liu, F. Schaub, S. Wilson, N. M. Sadeh, S. M. Bellovin, and J. R. Reidenberg, "Automated analysis of privacy requirements for mobile apps," in *AAAI Fall Symposia*, 2016.

[20] D. L. Vannette, "Testing the effects of different types of attention interventions on data quality in web surveys. experimental evidence from a 14 country study," *American Association for Public Opinion Research*, 2016.

[21] T. Herrmann, A. Schmidt, and M. Degeling, "From Interaction to Intervention: An Approach for Keeping Humans in Control in the Context of socio- technical Systems," *Proceedings of STPIS'18*, p. 10, 2018. [Online]. Available: http://ceur-ws.org/Vol-2107/Paper8.pdf

Appendix

1. questionnaire

Data Sharing Practices

The record of your survey responses does not contain any identifying information about you, unless a specific survey question explicitly asked for it.

If you used an identifying token to access this survey, please rest assured that this token will not be stored together with your responses. It is managed in a separate database and will only be updated to indicate whether you did (or did not) complete this survey. There is no way of matching identification tokens with survey responses.

Disclaimer

General
This survey is conducted by the chair of System Security at Ruhr-University Bochum.

Purpose
The purpose of this study is to understand the user's expectations towards mobile data sharing practices. The answers to this survey will be compared to a later analysis of real world mobile applications in order to estimate how well they meet the user's expectations.

Participation requirements
Android users who **live in Europe** and are at least **18 years old** are eligible to participate.

Procedure
Participants have to answer some questions regarding the control over their personal information. Afterwards, they are asked questions about their position towards mobile data sharing practices. In the last part, participants may answer some demographic questions.

Data Privacy
We do not collect any personal information during the study. Your answers will be stored securely on a server hosted by by the chair of System Security at Ruhr-University Bochum in Germany.
You can find contact informations of the responsible data protection officer at https://dsb.ruhr-uni-bochum.de/.

Voluntary Participation
Your participation in this research is voluntary. You may discontinue participation at any time during the research.

Right to Ask Questions & Contact Information
If you have any questions, comments, or concerns about the study either before, during, or after participation, please contact me at nils.quermann@ruhr-uni-bochum.de.

1. I have read and understand the previous information.
- ○ Yes
- ○ No

2. I am age 18 or older.
- ○ Yes
- ○ No

3. I live outside of Europe.
- ○ Yes
- ○ No

4. I want to participate in this research and continue with the study.
- ○ Yes
- ○ No

Information Sharing Concerns

Regarding control over personal information, please rate your level of agreement.

5a. Consumer online privacy is really a matter of consumers' right to exercise control and autonomy over decisions about how their information is collected, used, and shared.
Strongly agree □–□–□–□–□–□–□ Strongly disagree

5b. Consumer control of personal information lies at the heart of consumer privacy.
Strongly agree □–□–□–□–□–□–□ Strongly disagree

5c. I believe that online privacy is invaded when control is lost or unwillingly reduced as a result of a marketing transaction.
Strongly agree □–□–□–□–□–□–□ Strongly disagree

Regarding awareness of privacy practices, please rate your level of agreement.

6a. Companies seeking information online should disclose the way the data are collected, processed, and used.
Strongly agree □–□–□–□–□–□–□ Strongly disagree

6b. A good consumer online privacy should have a clear and conspicuous disclosure.
Strongly agree □–□–□–□–□–□–□ Strongly disagree

6c. It is very important to me that I am aware and knowledgeable about how my personal information will be used.
Strongly agree □–□–□–□–□–□–□ Strongly disagree

Regarding collection of personal information, please rate your level of agreement.

7a. It usually bothers me when online companies ask me for personal information.
Strongly agree □–□–□–□–□–□–□ Strongly disagree

7b. When online companies ask me for personal information, I sometimes think twice before providing it.
Strongly agree □–□–□–□–□–□–□ Strongly disagree

7c. It bothers me to give personal information to so many companies.
Strongly agree □–□–□–□–□–□–□ Strongly disagree

7d. I am concerned that online companies are collecting too much personal information about me.
Strongly agree □–□–□–□–□–□–□ Strongly disagree

Position towards Mobile Data Sharing

Please rate your level of agreement with the following statements.

8a. In the past, I prefered one app over the other, if it shared less personal information.
Strongly agree □–□–□–□–□–□–□ Strongly disagree

8b. On average, apps request an adequate amount of permissions.
Strongly agree □–□–□–□–□–□–□ Strongly disagree

8c. Concerns regarding data sharing play a role when choosing an app.
Strongly agree □–□–□–□–□–□–□ Strongly disagree

8d. On average, apps share an adequate amount of personal information.
Strongly agree □–□–□–□–□–□–□ Strongly disagree

8e. I expect apps to share only personal information that are required for their purpose.
Strongly agree □–□–□–□–□–□–□ Strongly disagree

Mark information that you expect to be collected by apps in the specific category. Multiple marks per row/column are allowed!

9.

	Games (Candy Crush, Temple Run, ...)	Entertainment (YouTube, Netflix, ...)	Tools (Calculator, File Explorer, ...)	Education (Babble, Duolingo, ...)	Personalization (Google Keyboard, Samsung Launcher, ...)
Location (e.g., know your position)	□	□	□	□	□
Message (e.g., access your SMS)	□	□	□	□	□
Voice Record (e.g., record your voice)	□	□	□	□	□
Pictures (e.g., access your camera)	□	□	□	□	□
Stored Files (e.g., access your storage)	□	□	□	□	□
User ID (Account name, (unique) identifier, ...)	□	□	□	□	□
Behavior (e.g., search history, actions within the app, ...)	□	□	□	□	□
Sensitive (e.g., ethnicity, medical records, ...)	□	□	□	□	□
No additional Permissions needed	□	□	□	□	□

10. What would you guess, how many different services receive your personal information from a single app?
- □ 1 (only the app service)
- □ 2-3
- □ 4-6
- □ 7-10
- □ More than 10

11. In your opinion, how likely is it that your behaviour is tracked accross different apps? For example, displaying ads on CandyCrush based on your search history from eBay.
Very likely □–□–□–□–□ Very unlikely

Processing of Personal Information

12. Would you expect apps from a certain category to collect more data than apps from other categories?
- □ Yes
- □ No

13. (If previous question answered yes) Which category and why do you think that way?
_____
_____

14. What do you think, how many apps do send personal information before getting your consent?
- □ None
- □ A minority
- □ About half of them
- □ A majority
- □ All

15. In general, how concerned are you about app services sharing some of your data with third parties?
Concerned □–□–□–□ Unconcerned

16. (If previous question answered) Why do you think that way?
_____
_____

17. In your opinion, what are the five apps you most frequently use?
1. _____
2. _____
3. _____
4. _____
5. _____

General

18. The amount of data collected from mobile devices has increased significantly in the last decade. There exist several online companies. In order to test your attention and regardles of the instruction below, please select any Name starting with "E". In your opinion, which company collects the most data? Multiple selections allowed.
- □ Google
- □ Facebook
- □ Amazon
- □ Ebay
- □ Microsoft

19. How do you self identify?
- □ Female
- □ Male
- □ Other: _____

20. What is your age?
- □ 18-24 years old
- □ 25-34 years old
- □ 35-44 years old
- □ 45-54 years old
- □ 55-64 years old
- □ 65-74 years old
- □ 75 years or older

21. What is the highest degree or level of school you have completed? If currently enrolled, highest degree received.
- □ Less than a high school diploma
- □ High school degree or equivalent
- □ Bachelor's degree
- □ Master's degree
- □ Doctorate
- □ Other: _____

22. What is your current employment status?
- □ Employed full-time
- □ Employed part-time
- □ Unemployed (currently looking for work)
- □ Unemployed (currently not looking for work)
- □ Student
- □ Retired
- □ Self-employed
- □ Unable to work
- □ Other: _____

2. Text Responses

TABLE 8. APP CATEGORIES THAT ARE EXPECTED TO COLLECT MORE PERSONAL INFORMATION THAN THE AVERAGE ONE

Games
"I would expect games or fun apps to collect less."
"Free gaming apps."
Social media
"Social networks and dating apps (Facebook, instagram, inder,...)."
"Social network apps, games."
"Because Facebook, for example, needs more data than a gaming app."
"A search machine or a social network."
"I think social media apps will collect more data then other apps."
Shopping
"Shopping apps, as they collect data on what you are interested in."
"Online shopping, as they would seek to exploit the data to persuade you to buy more things."
""I think it's more plausible in apps that sell products."
Entertainment
"Entertainment and Games."
"I would expect entertainment companies to collect the most, because it is vital for their ad targeting."
Other
"It depends on the apps. Some application needs more data than other depending on their category, in order to serve you better."
"Any apps, that get their revenue mostly with ads."
"A extremely specific app may need some certain data to carry out its task to the the best of its ability. A medical app, for example, may need sensitive material."

TABLE 9. REASONS AS TO WHY PARTICIPANTS FEEL (UN-) CONCERNED ABOUT THIRD PARTIES RECEIVING THEIR PERSONAL INFORMATION

Concerned about the destination
"Because I don't know who is going to receive my data."
"Because sometimes you do not know where exactly that data ends up."
"I feel like I'm not told what data will be shared and with whom. I also don't know who is profiting from my data without my consent."
"Because I don't get to know to which third party they send them and what kind of my personal information they give them."
Concerned about the use
"Never know where or how it's going to be used, or if it will be used against you at some point."
"Third party apps could use those information for unknown purposes."
"Data can easily be used to manipulate views."
Unconcerned as it would mean no harm
"Honestly, if it doesn't harm me I don't mind them having some of my private data (but not all though)."
"I don't think sharing my personal information would damage me in any way."
"I don't like the idea of it, but I think the information they share isn't that sensitive or important."
Unconcerned as participants don't care
"I don't mind personalized adverts, in fact, it helps sometimes to see an advert for something you were looking at, but I do object to apps selling my info without me knowing to lots of different companies."
"Because, personally, I do not care what they know about me."