Lost Data

Research Report

Dr. Astrid Carolus

Prof. Dr. Frank Schwab

Dipl.-Psych. Michael Brill

Ricardo Münch B.Sc., Catharina Schmidt B.Sc. & Florian Schneider B.Sc.

Julius-Maximilians-University Wuerzburg, Germany

Content

Conclusion and Summary

Our research project offers a psychological perspective on the value of data stored on our smartphones. Under the central assumption that the use of smartphones is closely related to the fulfilment of fundamental human needs, our last research report proposed smartphones taking on the role of a digital companion to us[1]. From this perspective, smartphones are relevant for a range of important psychological processes actually exclusive for the human-human-relationship (e.g. personal involvement, importance, closeness).

To gain deeper insights into the meaning of smartphones, this project focuses on the data stored on our phones. After all, without the data on it, our phone is a mere portable device: valuable because of its material value. However, what seems to constitute our digital companion is not (only) its fancy exterior but also its inner values. The data stored on it, all the photos and messages sent to and by our beloved ones, the job-related documents and mails or simply the music or apps. What cyber criminals already know gradually dawns on us: **Yes, the hardware is expensive. However, it is the data stored on that hardware that makes it precious.** If we think of what makes the phone important, will we think of the data as the essential element? Do we (already) know this? And, even more important from a psychological perspective: Do we feel it?

Emotions are indicators of relevance. If an object or event is relevant to us, it will cause an emotional reaction or emotional relation. Following theoretical work, emotional reactions are a result of multiple cognitive appraisals (e.g. novelty, goal relevance, coping skills) leading to reactions of different components of the affective system (e.g. facial expressions, bodily changes). Therefore, measuring and understanding emotions in the context of smartphone data has to address both aspects: **cognitive appraisals and resulting emotional reactions**. Two studies were conducted providing empirical insights. First, we asked for the value participants ascribe to the different categories of data (cognitive appraisal). And second, we assessed the bodily and emotional reactions to the loss of this data (physiology and expressions).

Study 1 addresses the cognitive component or the cognitive appraisal, which we defined as the monetary value of data. Participants were asked to report the amount of money they would pay for data recovery in case of smartphone data loss. Then, they distributed this sum of money to the particular data categories stored on their phone, so less important categories could be assigned less money, and more important data could be assigned more money. Afterwards, they were made an offer: They were to get money for the deletion of their data. More precisely: They were offered to be paid for every category of data if they agreed to delete the data from their phones. This decision needed to be made for each data category. From these decisions, we derived the most important as well as the most valuable categories of data stored on smartphones: **the most important categories are photos (of oneself, family and friends), contact information, personal emails as well as documents**. In sum, participants valued private data higher than job-related data. Concerning the ascribed money, the **most valuable categories are: financial and payment details, contact information, personal documents, photos (general and family/friends) and passwords**. However, the average amount of money ascribed to these most valued categories was rather small, ranging from 13.33 € (financial details) to 8.75 € (passwords).

Study 2 addresses emotional processes that occur during data loss. The study goes well beyond questionnaire measures and focuses on (1) **the physiological component** (electrodermal activity and facial thermal activity) and (2) **the expressive component** (facial expressions) when confronted with data loss. These physiological and behavioral measures are interpreted as non-reactive, hard to manipulate indicators of emotions. Therefore, the analysis of these two components offers **deeper insights into the emotional processing of data loss, or the value of data, respectively**. Based on the insights gained in study 1, we simulated two scenarios of data loss: In a condition of severe loss of personal data, participants were led to believe that all personal data on their phones were being deleted--with emphasis on photos and contacts, the most valued data in study 1. As a control condition, participants were led to believe in an identical way that just their call history was being deleted. This not only checks potential effects of an error message by itself, it can also be seen as a very strict control condition, because it also is a negative event of data loss. Differences between the conditions could thus be directly interpreted as the extent of data importance. All three objective measures pointed in the supposed direction: The simulated loss of more important data also led to stronger physiological reactions. These trends

[1] http://media.kaspersky.com/pdf/Carolus-et-al-DigitalCompanion-ResearchReport.pdf

were not statistically significant, so interpretation requires due caution. However, it can be concluded that there is an effect of data importance, since the findings from study 2 seem to be in line with study 1. While surprisingly low monetary values had on average been assigned to data categories in study 1, there are also rather weak physiological reactions in study 2; both could be interpreted as related to a low subjective value of data. Further, categories that were more important in study 1 by tendency elicit stronger physiological reactions in study 2, so both measures of data value seem to be converging.

In sum, this report introduces a psychological perspective on data value, or the appraisal of data loss, respectively; this is achieved by focusing on cognitive, physiological and behavioral aspects of an emotional reaction. Regarding theoretical frameworks on emotions, these different domains can be interpreted as measurable and observable indicators of rather unobservable emotional processes. Summarizing the results, our studies encourage to draw a more complex multi-level picture of the emotional appraisal of stored or lost data. There might be profound arguments for considering data as valuable or even precious. However, roughly summarized, our studies reveal a quite differentiated picture:

Results from physiological as well as behavioral data are consistent. Consequently, because data is considered as rather inexpensive with an appraisal of low relevance, the reactions to data loss are rather weak. Cognitive appraisal and physical reactions are two sides of the same coin: If something is less valuable, its loss will not evoke excessive affective and bodily reactions.

In addition, interviews with our participants point to the fact that there is a differentiation with regard to the kind of data loss. A mere loss of data, e.g. accidentally deleted or due to device failure, is generally perceived as a totally different story than incidents with a criminal background, as in e.g. blackmailing.

Thinking of conclusions, we recommend a reconsideration. The first step might be a step back. People - at least up to now - rarely assume their data to be valuable. They need to be told and they need to understand. However, the value of data is hardly tangible. We know that from various allegedly free online platforms: People might rationally know that they “pay” for these services with their data, but these costs are still hard to calculate or specify. The concrete value of data remains vague. Therefore, we assume a rational, economic approach with rational arguments stressing the mere monetary value as less promising. We suggest to focus on the heartfelt value and address people emotionally: Pointing out what data stands for, what it means to oneself personally. People need to understand, or even feel, that e.g. photos are not merely pictures and contacts are not merely addresses. These data categories are rather their most valuable life memories and their representation of social connectedness and affiliation. The value of data needs to be communicated. Only then may people realize the preciousness of their data.

Theory

All my photos are gone :(

Smartphones are wonderful. Everything we love, everything important for us is stored on our phone: contact information, photos, emails, messages, and all the memories. It is so cool to carry them around, to always have them available! It is an easy way of organizing your life and staying on top of things.

However, from light comes darkness. Unlit darkness in this case: data loss. Do you know the feeling when you realize you have just deleted all the photos you took on that day? Then you realize that – once again – you did not carry out adequate backups. As a result, the photos are effectively lost. Gone forever. Or have you ever experienced that an update or a reset did not work? When you slowly realize that you have just lost contact information and your address book is incomplete, now? Or that the messages which were sent via the mobile messenger during the last month or so are gone? All the nice little messages your beloved ones sent. In most cases, all these examples of losses do not mean monetary loss. However, it sometimes feels even worse. And the data which is lost feels even more precious.

This study focuses on the perceived value of data stored on the smartphone. From a psychological perspective we will ask: What do people think and also feel about the data stored on their phone? Thinking about "value" this study will ask: What is the subjective appraised monetary value of the data? And: Will people sell their data? Furthermore, we will ask: How will people react if confronted with data loss? Will they be emotionally affected? What will their body tell?

The psychology of smartphones and data stored on it.

Back in the days, data was easy to handle. All the important goods were stored in a shelf or in a cupboard or file cabinet. And the most important stuff was locked: files with important certificates and documents in them, photo albums, records of a bank, records, videos, etc. Nowadays, data has become digital and the amount we need to handle has multiplied. The closet has been replaced by several electronic devices: e.g. important documents are stored on the PC at work and on the notebook at home, photos and favorite music are handled via our tablet and movies are stored on the smart TV. However, even if there is a variety of devices, one device is special: our smartphone.

Since Apple's first iPhone was launched back in 2007, smartphones have evolved quickly. More and more functions and applications helping us to communicate, to organize our lives or to entertain ourselves have been developed. These various modes of operation are perhaps the main reason for the enormous popularity of smartphones. During the past years, smartphones have become the most popular electronic device. In 2016, more than 60% of the population in Western Europe owned a mobile phone[2]. Taking age differences in account, the simple picture is that in 2015 in the age group 18 to 29 (depending on the specific statistics) nearly everybody owned a smartphone. And although the older population (65 plus) is markedly behind in terms of smartphone ownership, this gap is closing[3].

Taken together, our **smartphone is more than just another device:**

1. It is popular. Smartphones are the most popular devices and are way ahead of all the others. Today, nearly everybody owns a smartphone - and counting.
2. It is mobile. It is the only device we carry around with us 24/7.
3. It is versatile. As a consequence of its various functions (e.g. calling, texting, browsing the internet, taking photos, listening to music, emailing) a smartphone contains various kinds of data relevant to various situations or aspects of life.

---

[2] Statista (2017): Smartphone user penetration as percentage of total population in Western Europe

[3] Pew Research Center (2015): The Demographics of Device Ownership

Combining these three aspects, our smartphone offers multiple data formats from multiple sources for multiple challenges of everyday life. You witnessed something exciting? The smartphone captures the moment within a photo. You forgot to answer work emails? The smartphone offers mail access to do it on your way. Calling your partner, children, or family? The smartphone offers contact information. Reminding you of personal or work-related tasks throughout the day: your phone is a constant and capable as well as supporting companion.

Based on usage statistics as well as psychological theory, **smartphones** are likely to be more than ordinary portable computers. In a previous report, we introduced the term **"digital companion"** to stress the guiding idea that smartphones have long ceased to be mere technical equipment to us. We establish some kind of relationship to our phone resulting in a feeling of connectedness to our phones[4]. Smartphones keep us linked to the world, providing us with a variety of objective and useful features as well as gratifying essential human needs. Furthermore, we trust our phones in terms of keeping our data safe, no matter what happens. We store more and more important information and memories on the phone - confident that our phones will provide us with the information whenever we need it. Consequently, smartphones keep all the goods in safe custody, which used to be stored in multiple repositories. Thus, **our smartphone is a combination of repositories.** It is a shelf with e.g. pictures in it we like to show around. However, it is also a file cabinet with some files in it we would rather only show to selected people. Finally, it is also a vault with sensitive data we want to keep secret.

In sum, smartphones have become essential to (modern) living by offering important information for various private and job-related aspects of life. This data is either stored on the phone itself or the phone offers access to information stored online e.g. in web servers. Consequently, it is not only the hardware which is valuable but also the data stored on it. However, while the value of the phone itself is clearly priced, the value of the data stored on it is not. We depend on this data everyday, but are we aware of its "price"? Are we able to quantify the value of data? Furthermore, is the value of data reflected in our reactions to data loss?

These two **main questions** will guide us through this research report:

1. **What is the monetary value of data stored on our phone?**
2. **Is the value of data reflected in our reactions to data loss?**

How valuable is it: the monetary value of data.

People endeavor to satisfy different needs, e.g. affiliation, entertainment, information, communication or self-realization. Driven by our different needs, we use different kinds of tools to satisfy them. Smartphones are also need satisfying tools: If you want to be entertained, you will turn on music stored on your phone. You want to recall the dream vacation you did last year? There's only one thing you need to do: open the photo app on your phone and swipe through all these memories. Plenty of needs can be managed and fulfilled by our phone in a fast, easy and economic way. Once the information is stored, we can use it again and again. Thus, we think of the information on our phones in terms of "**externalized representations**" **of important aspects of our lives.** Furthermore, we think of it as a representation of the objects of our basic human needs. Accordingly, or computers are like extensions of our cognitive system. Richard Dawkins, an evolutionary biologist, coined the term "extended phenotype" which refers to the capacity of living creatures to modify their environment. Dawkins argues to widen the perspective on phenotypes by including not only the organism's body and behavior but also the environment this organism modifies. To give an example: The beaver builds his dams (organism modifies environment) which is regarded as extended phenotype (Dawkins, 1982). To transfer this to human life: Our electronic devices are cognitive tools similar like the beaver's dam. Our smartphone seems to be a collection of subjectively important goods (photos, contacts, messages). Therefore, phones can be regarded as **"extensions" of our body or an "extended identity"** (Schwan & Hesse, 2004). We know from our evolutionary

[4] http://media.kaspersky.com/pdf/Carolus-et-al-DigitalCompanion-ResearchReport.pdf

history that the possession of certain tools or weapons but also the possession of symbols of status is deeply rooted in our human history and belongs to our human phenotype (extended phenotype, Dawkins, 1982). Those artifacts are part of our nature, like the spider and its web or the snail and its house.

This perspective on human functioning offers an approach to understand the emotional importance of data. By adopting a psychological perspective on human functioning, we explain how modern technology fits into this hundreds-of-thousands-year-old puzzle of human evolution. Being a product of evolutionary processes our psychological mechanisms have adapted to our ancestors' world. A world in which we were adapted to tools, indeed - but also a world without digital devices, internet and modern technologies. Consequently, we were adapted to a life where we interacted face-to-face. A world in which we communicate computer mediated - respectively smartphone mediated - is new to us, at least from an evolutionary perspective. Due to the fact that we are still adapted to our ancestors' world, we transfer evolved habits to modern life and modern technologies. From the cave to the cupboard to the smartphone we collect real and digital artifacts related to our basic - often archaic - needs. Most of them are emotionally charged. Most of them should look like a digital footprint of our evolved mind addressing the same things a hunter-gatherer would be fond of (wonderful landscapes, beautiful men and women, pop music, things for leisure time, contact to our beloved and respected, signs of status and so on).

Muslukhov and colleagues (2012) interviewed a group of heterogeneous users regarding the sensitivity and value of different types of data. As a result, passwords are considered to be very sensitive data and partly valuable depending on applications that manage the passwords and reduce the likelihood of a loss. Some of the participants defined photos and videos as both valuable and sensitive, but they couldn't tell exactly which pictures and videos were considered to be of great importance. Music and the events in the calendar were the only types of data that were never mentioned as being sensitive or valuable. A study conducted by Kaspersky Lab asked participants to assess the importance of the data stored on their phones (see table 1). In contrast to the interviews conducted by Muslukhov and colleagues, results reveal that photos and videos were most important (private and sensitive, of myself, my children, and other people). Passwords, scans of certificates and documents came second. Work-related and personal emails as well as personal messages (SMS, IM) were considered less important.

Table 1: importance of data stored on participants' phones (study by Kaspersky Lab)

data category	**importance (in %)**
private and sensitive photos and videos of myself	48.80
photos and videos of my children	39.30
private and sensitive photos and videos of other people	38.90
passwords (including auto-logging in websites and apps)	25.10
photos and videos of travel	23.10
scans of passport, driver's license, insurance and other sensitive scans	22.40
financial and payment details	21.10
address book / contact information	18.40
other photos and videos (not sensitive)	12.20
personal notes and documents	11.80
work-related documents	11.60
personal emails	09.80
personal messages (SMS, IM)	08.80
work-related emails	08.50

These studies offer interesting insights into the value of data. However, we need to take into consideration the origin of these numbers. How was "importance" or "value" defined? What was the participants' exact task? To rank data or to use a given scale (e.g. 1 to 10)? Both, ranking and scaling is rather not everyday activity. If you ask for value, people would rather not answer in terms of scales or rankings. Reported studies yield results and numbers, but also leave room for interpretation. As a consequence, our study aims for a more "common sense" operationalization of data value: Living in a society where nearly everything is monetarily prized calls for prizing data, too. According to prior studies, we assume different categories to be valued differently.

And off they go: The threat of data loss.

While the first part of our study asked for the monetary value of data stored on the smartphone, this second part does not only ask - it observes reactions to loss of this data. Furthermore, we widen our methodological approach. As self-reports are prone to manipulation by participants (intentionally and unintentionally) not telling the truth (social desirability) we do not only ask participants, but implement an objective measure: we focus on bodily reactions which people are less able to control.

As we said before, smartphones are not only valuable because of the hardware (a new Iphone costs up to 1,119 €[5]) but also because of the data stored on it. In 2012, nearly 50% of young Americans and 20 % of older adults reported their phone to be lost or stolen[6]. Taking care of your phone as well as various opportunities of theft insurances protect against the monetary loss. However, what about the data stored on your phone? Are we aware of cyber criminals who do not limit themselves to the data on your pc or notebook? Or do these criminals care more about what is inside your phone? Is your mobile data more precious to them than to you?

Cyber criminals are a hazard for the information stored on our devices and online (server, cloud). A stolen phone could mean access to the data stored on it - unless it is adequately locked via a PIN or password. However, it is far from impossible to get hold of a PIN or password: 93% of the participants of one of our earlier studies gave away their PIN when asked for it by the researcher. Good to know for cyber criminals: Just pretend to work for a university and people will give you what you want.

Malware, phishing attacks or mobile spyware unintentionally installed by downloading e.g. apps are serious threats. Furthermore, what about your precious goods stored on your phone? The hardware could be replaced in case of emergency. But what about your beloved photos or love letters sent and received via instant messages? What about passwords, pins and certificates? If they were stolen or deleted (accidentally or with evil intention) how would you feel?

Feeling the loss: Emotional reactions

Talking about our data being exposed to a risk or even loss seems to evoke emotional reactions or feelings. Reacting emotionally to data loss seems to be obvious. Consequently, we need to analyze these reactions to gain deeper insight into the psychological effects of data loss. However, from a scientific point of view the reliable and valid measurement of emotions is a bold venture (Scherer, 2005). Most researchers in this scientific discipline consider emotions as a multilevel phenomenon that affects our mental system and physiology on more than one level. Thus, there is not only one way to measure emotional reactions. The adequate approach is to address multiple aspects of emotional responding as we are handling a multi-level phenomenon.

Table 2: Emotions as a multilevel-phenomenon (see Merten, 2003)

component	measurement (examples)
cognition	questionnaire, interview
(neuro-)physiology	e.g. EDA, skin temperature
motivation	behavior, behavioral intentions
expression	facial behavior, gesture, proxemics
feeling	questionnaire, interview

[5] iPhone 7 Plus 256GB, German Official Apple Store: http://www.apple.com/de/shop/buy-iphone/iphone-7/5,5%22-display-256gb-silber#01,12,22

[6] Statista, 2017b: 45% of the younger adults (18 to 24 years) have experienced cell phone loss or theft, about 20 % of cell phone owners aged at least 65 years.

To give an example: By asking someone how he/she feels e.g. using a questionnaire, you address the conscious and verbalizable aspect. However, here are further components of our affective system to be assessed: expressions of face and body (e.g. clenched fist, saturnine look), more or less visible bodily functioning (e.g. heartbeat, sweating, blood pressure) and motivation (e.g. anger or fear to fight or flight). Up-to-date approaches in emotion research (e.g. Merten, 2003; Scherer, 2001, 2005, 2009) describe emotional reactions along five components.

For the purpose of this study, we will go beyond survey methods and also focus on objectively observable correlations of emotion in physiology and expression. While most of such measures primarily describe unspecific aspects, such as a person's level of arousal, the combination of such measures and the corroboration with survey data nevertheless allows for a more complete picture of emotional processes.

In a nutshell:

Focusing on the emotional part of data loss is more than just asking "How do you feel now?". A multi-faceted understanding of emotions offers researchers a more appropriate approach when examining this complex phenomenon: Cognitive, physiological, and expressive aspects of the human affective system should be taken into account. By adopting established psychological theorizing and methods, we can widen our understanding of data value and gain more detailed insights into related processes of motivation, cognition, and emotion.

Research Questions

Two projects were conducted to shed light on these questions. These laboratory experiments were developed to give an impression of the value people attribute to the data stored on their phone. The first study considers the monetary value of different kinds of data. The second study analyzes physiological reactions to loss of data. Physiological information is interpreted as an indicator of emotions offering deeper insights into psychological processes which are not accessible by questionnaires. Two sets of research questions guided our projects:

Experiment I: mind talks.
the monetary value of data.

A **How valuable are (different kinds of) data on our smartphones?**

B **What categories of data would be deleted for money?**

Experiment II: body talks.
physiological and expressive reactions to data loss.

How does the apparent loss of data stored on our smartphones affect our emotions, as measured by physiological and expressive reactions?

A **How does data loss affect electrodermal activity?**

B **How does data loss affect skin temperature?**

C **How does data loss affect facial expressions?**

Experiment I: mind talks.

the monetary value of data.

Research Methodology - overview

Here we will present a summary of the methodological approach presenting the essentials. Chapter "Procedure and instruments" offers more detailed information. In study 1, participants were introduced to the scenario of data loss to find out the value of personal data. They were confronted with (1) the theoretical case of data loss and (2) the possibility to "sell" data.

(1) The examiner welcomed participants to the laboratories and explained the procedure of the study. Afterwards, the study began with a **survey,** which was answered at a computer. Among other things, participants were asked which data they store on their smartphone. Furthermore, they were prompted to imagine the case that all the data stored on their phone (including backups) is lost. However, a company would be able to recover the data for money. Participants were asked: ***How much money would you pay for the recovery of your data?***

(2) After the questionnaire had been finished participants were introduced to a second setting. A croupier welcomed them to some kind of **gambling scenario**. They took a seat at a table which was illuminated in a darkened room. A so called "security expert" was sitting in the back surrounded by his computer equipment. The only reason he was involved: Creating the illusion of an expert who is able to delete data from the participants' phones.
Based on the information we had gained by the questionnaire the participants had to **ascribe monetary value to the data on their smartphone** (e.g. photos, contacts, calendar). Afterwards, the croupier made an offer: **deleting the data for the exact amount of money** which had been ascribed before. Participants acted in good faith that the security expert could effectively delete the data (from phone and from cloud storage) when they decided which categories of data could be deleted and which could not.

Sample

Over a period of 2 weeks (December 2016) we recruited via online advertisements (e.g. Ebay classifieds), social media platforms (e.g. Facebook) and mailing lists. The resulting overall sample consisted of **53 participants** ranging in age from **18 to 68 years** (*mean age* = 31.62, *standard deviation* = 11.15)[7]. **Female and male respondents were equally represented. The overall level of education was high** with a majority of students and employees with a university degree.

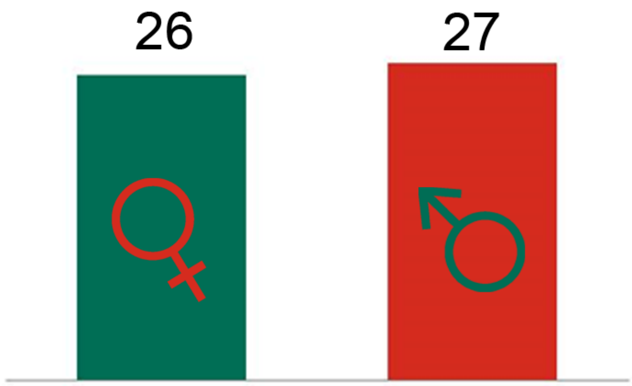


Figure 1: Number of participants by gender

The participants' age groups.

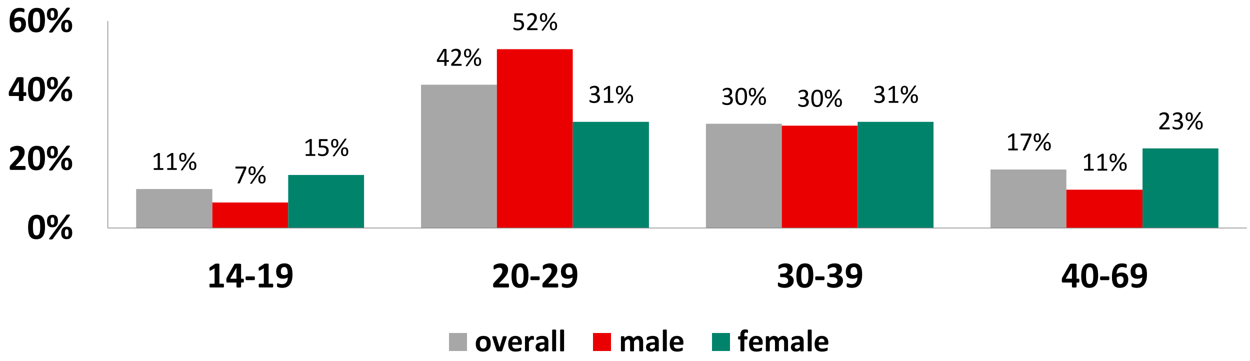
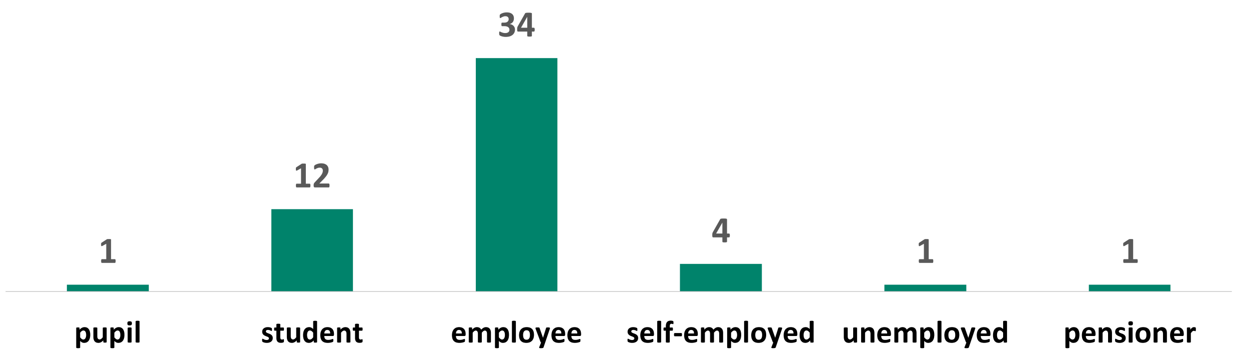


Figure 2: Percentage of participants by age group

As far as their occupation was concerned, most participants were employees.



[7]In the following, mean values will be denoted by *M*, standard deviations by *SD*.

Figure 3: Number of participants by occupation

Procedure and instruments

Participation in the experiment was entirely voluntary. The study followed core ethical principles based on the Declaration of Helsinki. Participants were invited to professional laboratory facilities of the University of Wuerzburg (Germany) to pass through **two settings**: **(1) survey and (2) gambling scenario**. Participants were guided by a researcher, who followed a detailed storyboard to ensure that each participant was addressed similarly and received identical instructions. After the participants were welcomed and introduced to the procedure they were **asked to give away the PIN for unlocking their smartphone** which was important for the credibility of the gambling scenario later on. The following segment describes the two scenarios in detail:

1. Survey - categories of data and overall value

The survey was implemented for two reasons. Participants were asked for (a) the different categories of data they store on their smartphone (pictures, calendar, music) and whether it was stored locally or using cloud services. This was an easy way to know which data participants had saved on their phone. They could have made mistakes, of course, e.g. reporting the wrong categories, forgetting others. However, this was not relevant for our experiment later on. We only needed the data participants were aware of. Afterwards, participants were asked for (b) the value of the data stored on their phone. Therefore, they were prompted to imagine a scenario in which all the data stored on their phone (including backups e.g. cloud storage) is lost. They could not get it back except with the help of a company which would be able to recover the data for money. Participants were asked: How much would you pay for the recovery of your smartphone data? This information was also relevant for the following experiment. As a result, we knew the categories of smartphone data relevant for the participant as well as the monetary value of this data. Furthermore, we increased the participants' awareness for data used via a smartphone.

2. Experiment - gambling scenario

After the questionnaire had been finished, a second researcher was introduced as "the croupier" who would guide the participant through the second part of the study. A gambling scenario was staged and the participant entered a darkened room with an illuminated green velvety table in it. The croupier invited participants to take part in a game offering the possibility to win money by gambling with your data. To take part they needed to give away their phone to a so called "security expert". This expert was sitting in the back surrounded by his computer equipment typing and looking at screens. The only reason he was involved: Creating the illusion of an expert who is able to delete data from the participants' phones as well as from all connected cloud services. When participants agreed and handed over their phone, he pretended to verify this by plugging it in and pretending to check a process on his screens, confirming that he was in fact able to access all data. (The phone was connected to a cable, which in fact was not connected to anything).

Afterwards, the croupier and the participant sat down at a table. The croupier laid down cards symbolizing the categories of smartphone data and explained that these were the categories the participants had reported in the survey. Then, he opened a suitcase and took out a packet of money, which he handed over to the participant while explaining that this was the exact amount of money the participant had reported to be willing to pay for the recovery of lost data. Now, the participant was asked to distribute the money to the different categories.

> *"The categories of data you have reported to have stored on your phone are symbolized by these cards [croupier laid down every card, one by one, slowly]. You have reported that you would pay XXX for the recovery of lost data. However, this was an overall sum for the restoring of all these data categories. I would like you to be a bit more precise by ascribing a value to each of these data categories. Would you please distribute the money to the different categories?"*

After having completed and reconsidered the distribution, the croupiers made his offer to delete data for the amount of money the participant had ascribed to it:

> *"We would like to delete your data in exchange for the amount of money you allocated. This would be the procedure: I will go through the categories one by one and ask you if would delete the data for the allocated money. This would be our data security expert's job. [....] To avoid misunderstandings: Data would be irrecoverably lost. [...]. We know from the information you provided in the survey which data is stored on your phone and which data is saved online. Both storage locations would be deleted."*

Consequently, participants acted in good faith that they were paid for the loss of data. The croupier walked the participant through the game, starting with the least precious category that had money allocated to it. If the participant decided to sell, the security expert took note of the category and the croupier went on. The game lasted until every category was sold/not sold. The participant was told that his/her data would be deleted at the end of the game.

The experiment ended with the solution and a final task. Participants were informed that this was only a game. Data had not been deleted. However, they did not win any money, either. One last task was to ask the participants if they would change the total amount of money for recovery and the distribution in terms of the data categories with the benefit of hindsight.

Results

1 Memory location and anti-virus software

There are a few categories of data, which nearly 100% of all participants store on their smartphone, with apps, call history, photos and text messages being the most common. Predominantly, **data is stored on the smartphone (locally) and not in the cloud.**

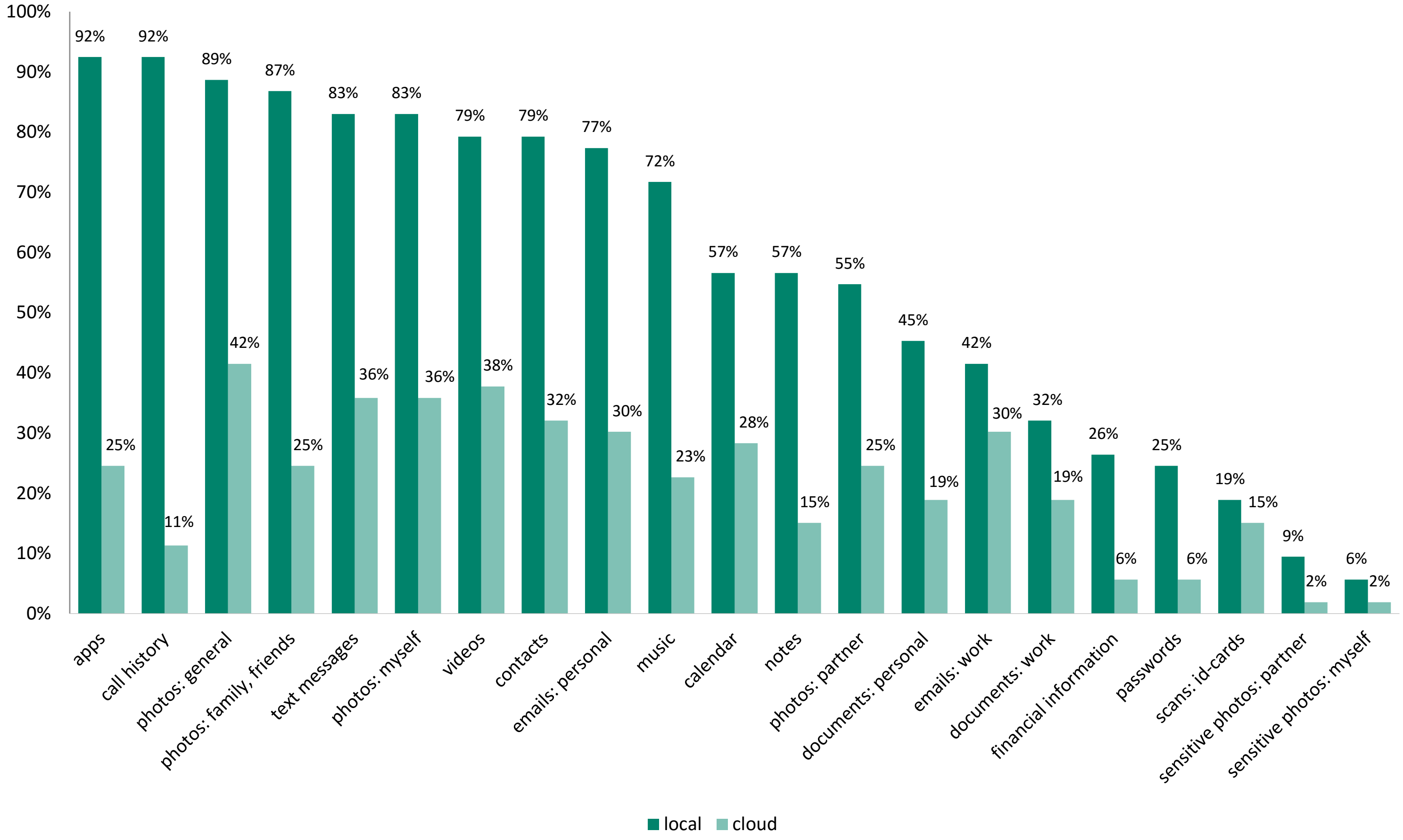


Figure 4: Diagram of the percentage of participants who store various data categories on their smartphones; the values are listed according to the type of storage starting with the most frequent (local vs. cloud)

40% of female and male participants use anti-virus software on their smartphones.

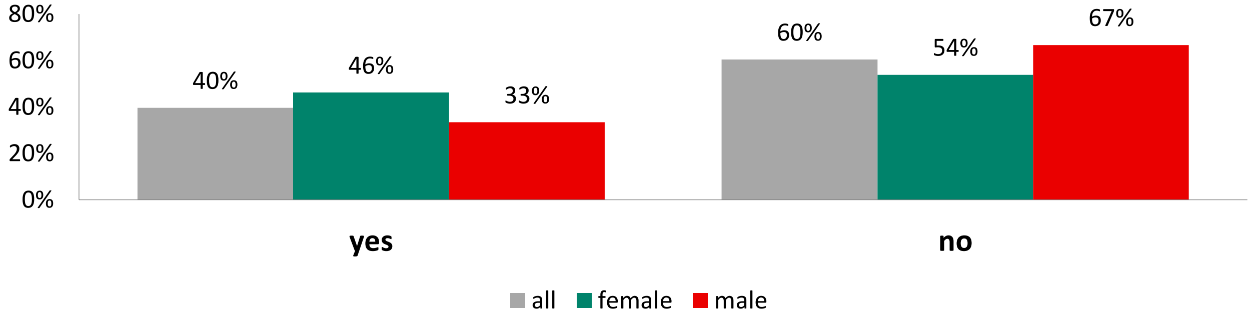


Figure 5: Percentage of participants who have an anti-virus software on their smartphone by gender and operating system

2 How much would you pay for the recovery of data: total value

To assess the overall monetary value of personal data participants were asked how much they would pay for the recovery of their data. Participants reported to pay a minimum of 1€ and a maximum of 5.000€ for data recovery, resulting in an **average value of 373.36€ (SD = 977.891) with men spending slightly more 395.78€ (SD = 989.76) than women (350.10€; SD = 984.44).**

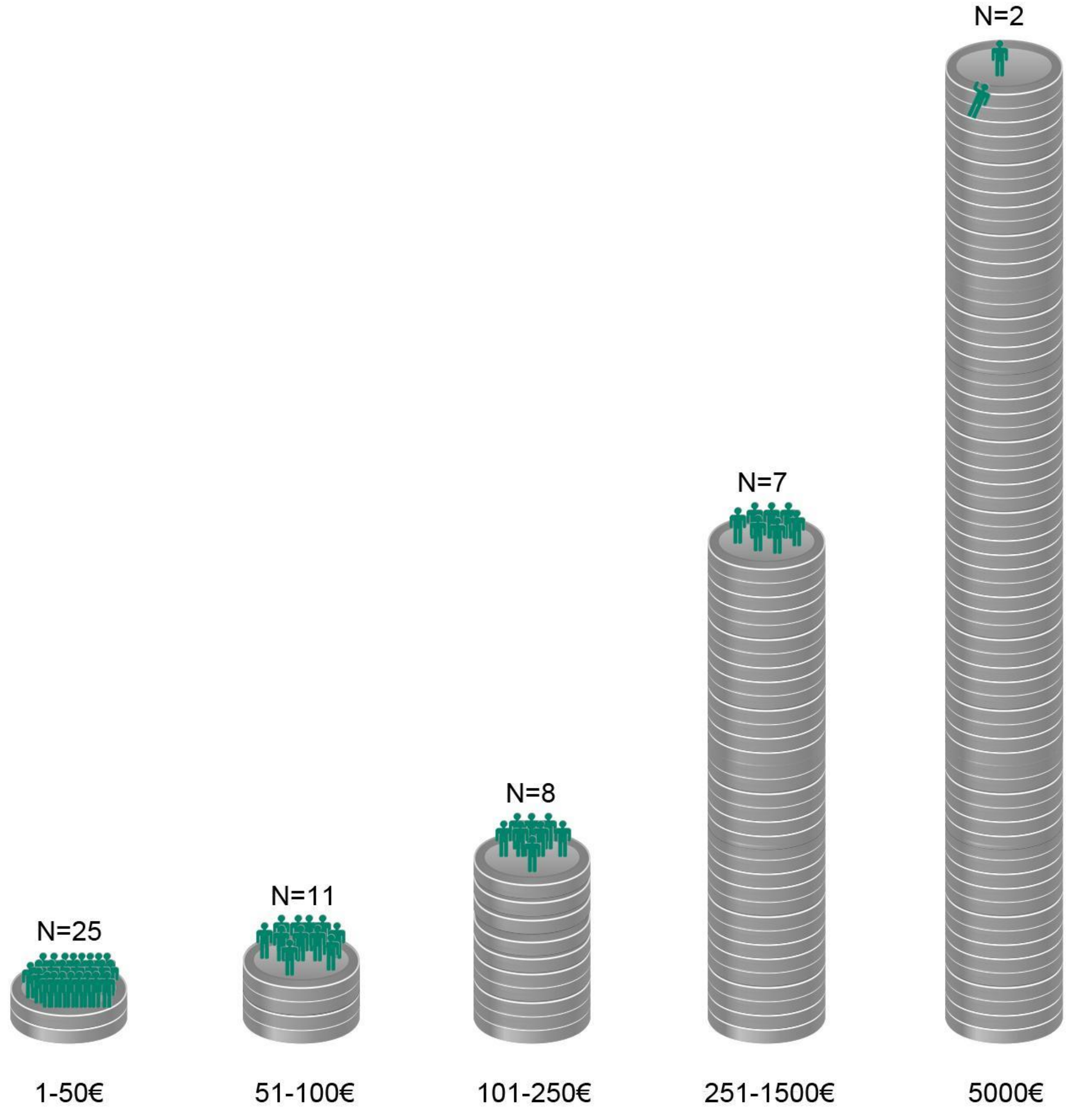


Figure 6: Number of participants categorized by how much they would spend to restore data

However, mean values do not represent the sample properly. Basically, there is a large gap between participants: **a small amount of people are willing to spend a lot of money** and a lot of people are willing to pay only a small amount to restore their data. While the majority reports rather small amounts of money, an outlier analysis reveals nine participants influencing the mean value by reporting values above 500 € (two would pay 5000€, another seven 500 to 1500€). Consequently, we analyze the monetary value in a more detailed way. Excluding the two participants reporting 5000€ lowers the average data value to 191.92€ (SD = 323.589). Additionally, **excluding nine participants who are willing to pay 500€ to 5000€ lowers the average to 74.73€** (SD = 52.46).

3 *Important data: How valuable is the data?*

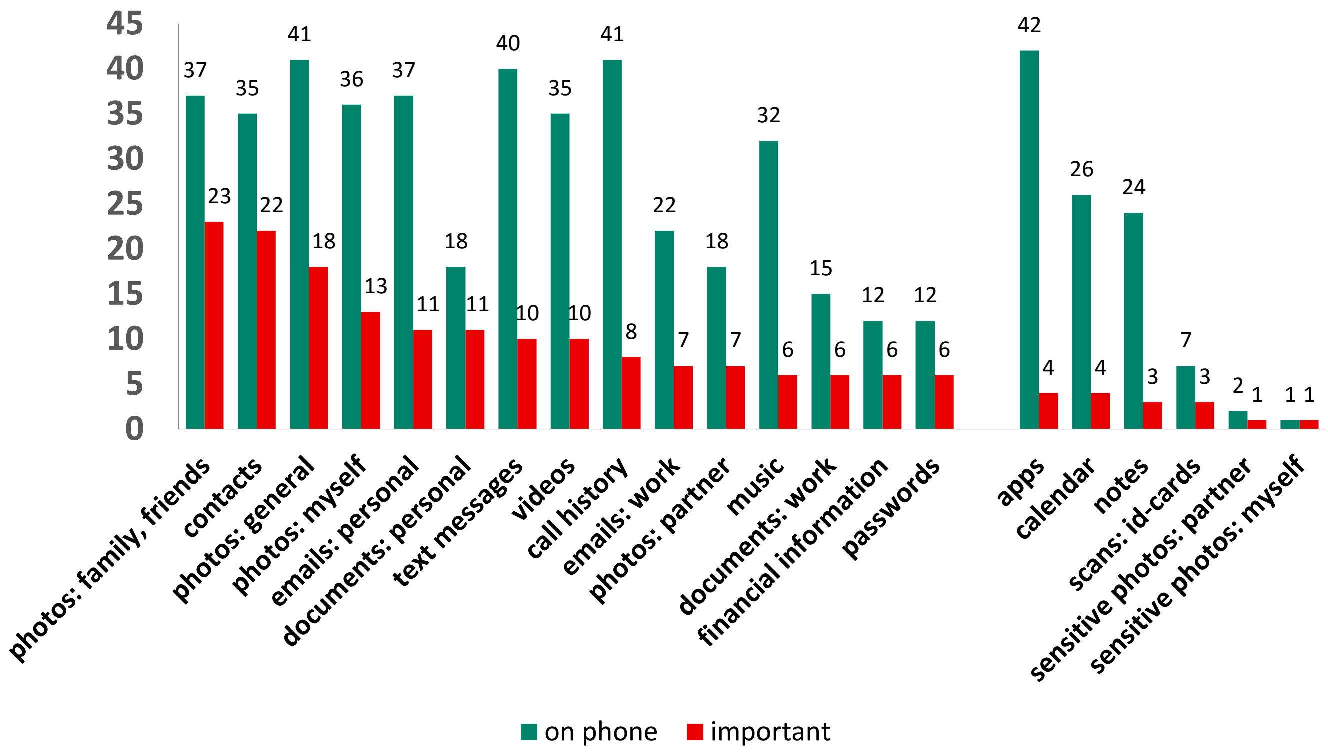


Figure 7: Data categories: existing on phone vs. considered to be important

Photos of family and friends, contact information and general photos were most often considered important (with no regard to the number of participants who reported the data category to exist on their phone). Data categories are defined as being "important" if participants allocated money to them. Several categories were only rarely considered important (5 times or less) or did not even exist on the phones often enough (see right side of figure 7). Consequently, these categories (apps, calendar entries, personal notes, scans of id-cards, sensitive photos of one's partner and of oneself) were excluded. Thus, the following analysis refers to the remaining **15 important categories**.

Figure 8 focuses on the 15 important data categories showing the percentages of participants who reported the categories to (a) exist on their phones and (b) to be important.

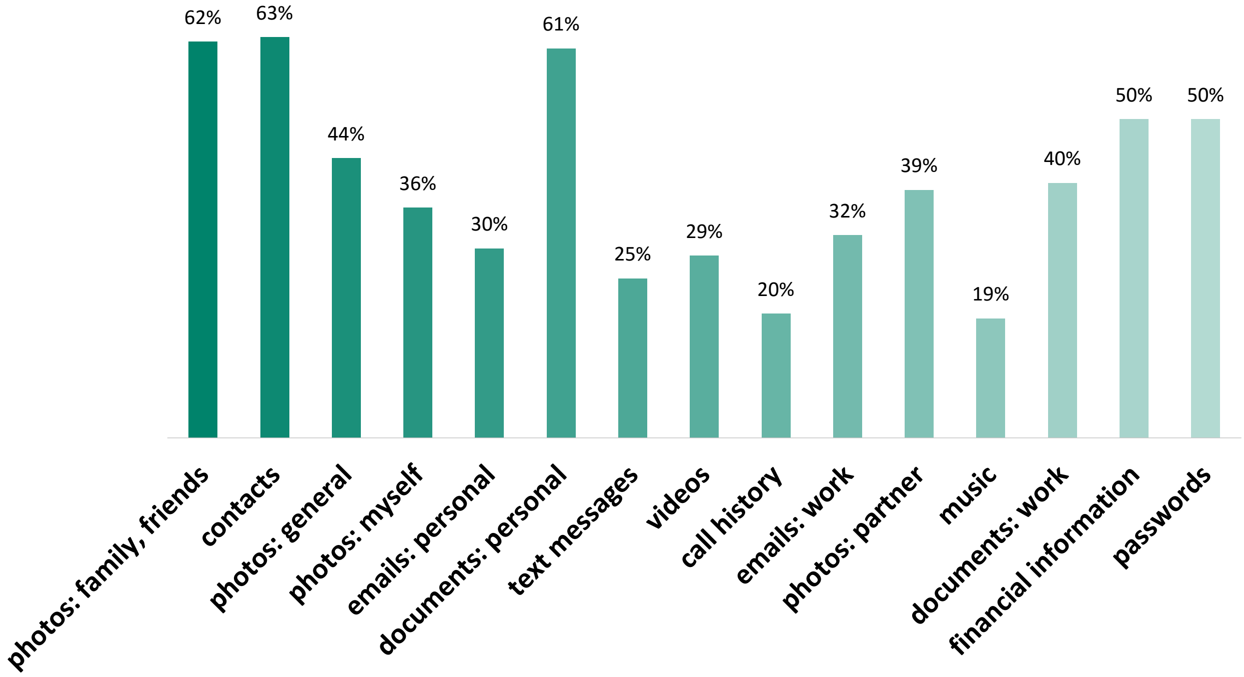


Figure 8: Diagram shows the percentage of data categories existing on the phone, which were considered important (15 most important data categories).

Figure 9 shows the assessed value of these important categories and reports the mean monetary value in € allocated to each category. On average, the category most money was distributed to was **(1) financial and payment details** with 13.33€ (but only considered important by *N* = 6), **(2) contact information with 11.89€** (considered important by *N* = 22) and **(3) personal documents** with 10.56€ (*N* = 11), **(4) general photos** with 10,37€ (*N* = 18), **photos of family and friends** with 9,05€ (*N* = 23) and **passwords** with 8,75€ (*N* = 6).

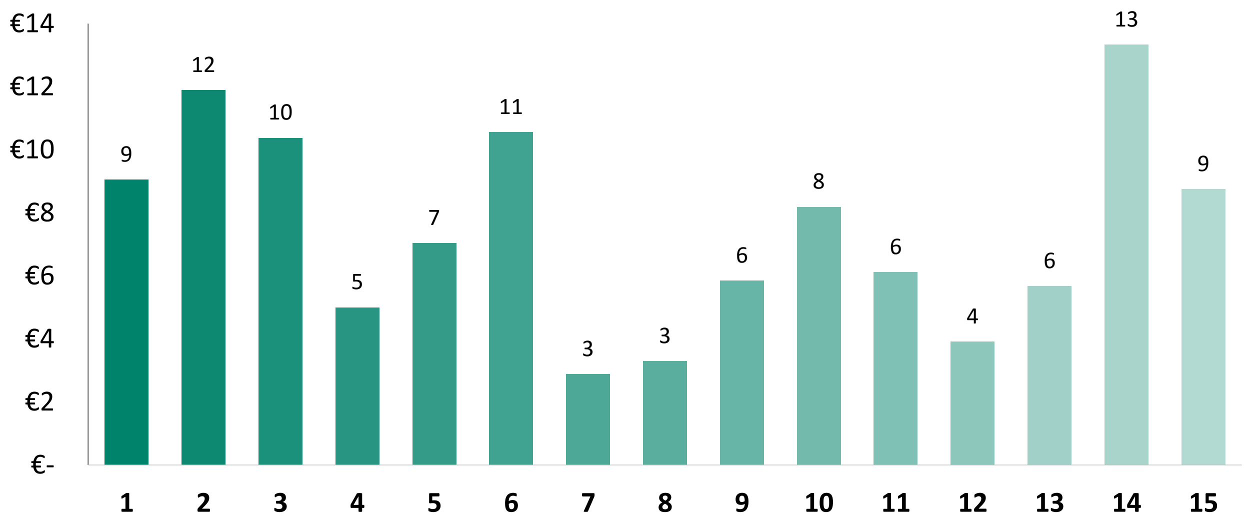


Figure 9: Mean monetary value in € of the 15 most important data categories

4 Data stored on the phone: percentage value of each category

While figure 9 shows the total monetary value of data categories, figure 10 visualizes the amount of money distributed to each of the 15 important data categories relative to the total amount of money participants had at their disposal. **As a reminder**: The amount of money participants could distribute was a result of the sum of money they had reported to pay for recovery of their data lost (survey answered before gambling session). This amount was to be distributed to the categories the participants valued the most.

Expressed as percentage (proportion: money per category - total amount of money), passwords, contact information and financial details were considered most valuable.

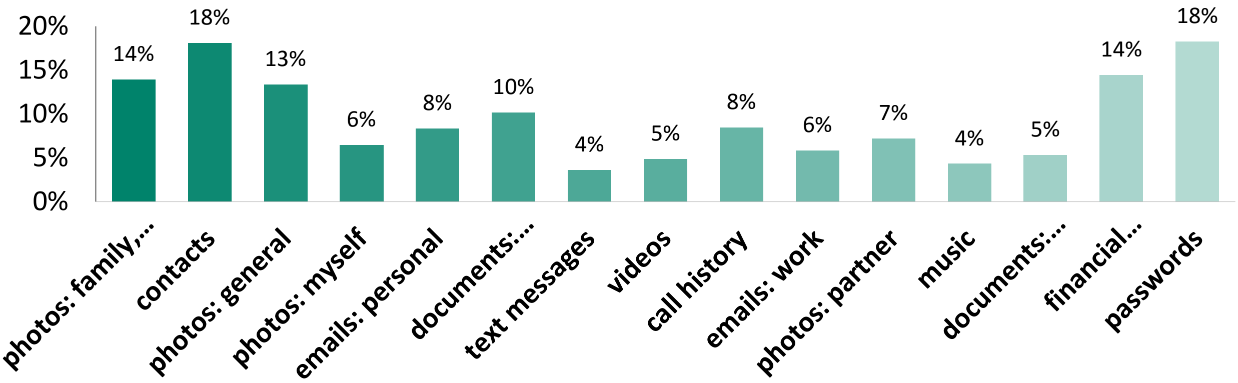


Figure 10: Mean monetary value allocated in % to the 15 most important data categories.

Figure 10 reveals that when looking at these measures, passwords rank first with an average 18.27% of the total amount of available money distributed on this category. Second is contact information with an average 18.10% of total money distributed on it. Third is financial details with an average 14.46%. The "Top 5" of money distributed relative to the total amount available concludes with photos of family and friends with 13.93% and general photos with 13.37%. This ranking is very similar to the ranking of mean monetary value (see figure 9), but differs somewhat in terms of which category occupies which place. The only category missing now is personal documents, which is replaced by password in this "Top 5". Still, it seems like there are some categories that prove to be more valuable to participants than others no matter what measurement is used.

5 Important data: deleting data for money

After having decided which data category is important by allocating money, participants were asked to approve if categories could be deleted in exchange for money. The amount of money was equal to the previously determined monetary value of each category. Figure 11 shows how frequent each category was considered important by allocating money and how frequent participants approved that these categories could be deleted for the amount of money previously allocated.

Photos of family and friends, personal documents and photos of oneself were most often approved to be deleted for their monetary value.

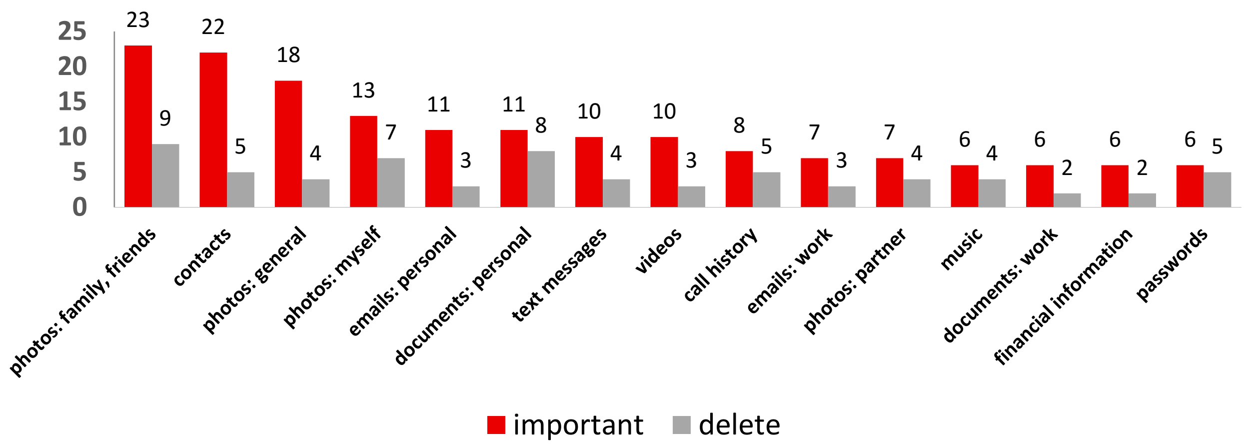


Figure 11: Frequency of the 15 most important data categories relative to the number of times allegedly deleted for money.

Looking at absolute frequencies, photos of family and friends were most often approved to be deleted ($N = 9$) followed by personal documents ($N = 8$), photos of oneself ($N = 7$), passwords and contact information (each $N = 5$). It is important to note that the proportional frequency of "approval to delete" relative to the number of times the category was considered important differs drastically between these data categories. Considering this, passwords were approved to be deleted most often (83.33% of times), whereas photos of family and friends (most often approved in absolute frequencies) were only approved to be deleted in 39.13% of times.

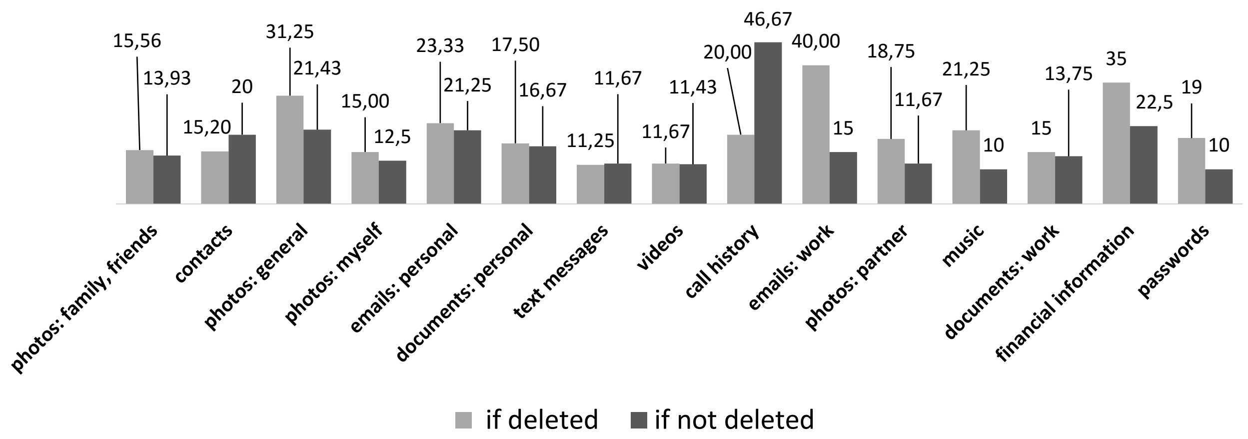


Figure 12: Mean monetary value in € of the 15 most important data categories relative to the times allegedly deleted or not deleted for money

Another aspect to look at is the amount of money the participants anticipated to get by approving to delete their data. This is illustrated in figure 12, which displays the mean monetary value in € of each of the data categories relative to the question whether the data category could allegedly be or not be deleted by our security expert allegedly.

Figure 13 provides an outline of the 15 most important data categories comparing how frequent participants 1.) stated that these categories existed on their phone, 2.) considered them important and 3.) approved of a deletion.

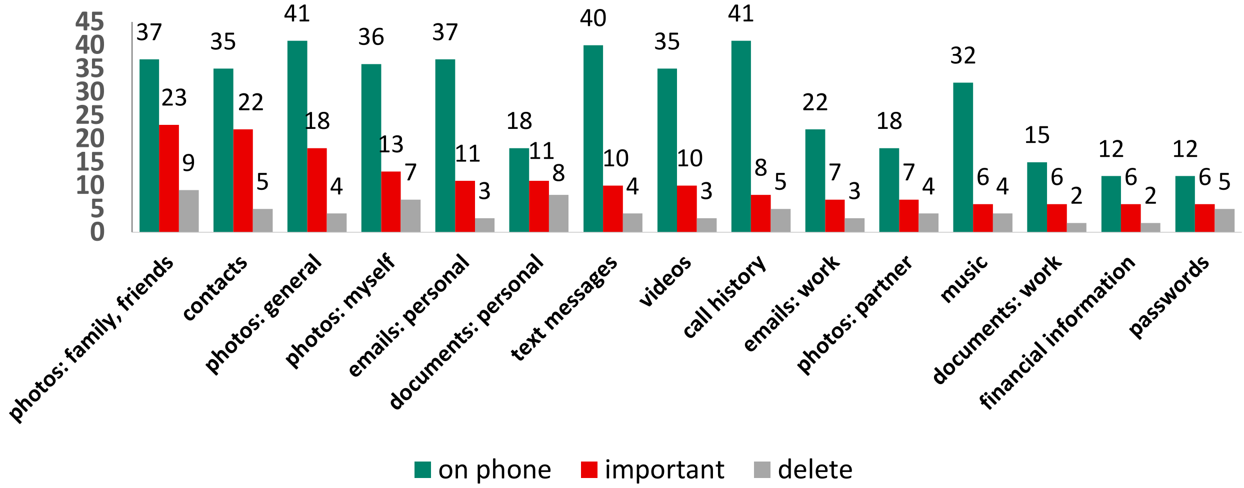


Figure 13: Comparing the frequencies of the 15 most important data categories regarding "on the phone", "important" and "delete for money"

Additional Notes

Figure 4 as well as figures 6-13 all depict the overall sample (N=44). In addition, we also checked for group differences between male and female participants as well as between anti-virus users and non-anti-virus users. There are, however, no statistically significant differences between those groups as far as importance, value and deletion of data categories is concerned.

As mentioned before, towards the end of the experiment, participants were asked whether they would redistribute the money they allocated to various data categories or even raise the total amount of money to better represent the value of their data. However, most participants chose not to change anything and those that did only made marginal changes not worth reporting.

Experiment I: Summary/Take-away

Predominantly, data is stored on the smartphone (locally) and not in the cloud.
Participants would pay an average value of 373.36€ (*SD = 977.891)* for data recovery.
However, … excluding the nine participants willing to pay 500€ to 5000€ lowers the average to 74.73€.

Photos of family and friends, contact information and general photos were the most important data.
Passwords, contact information and financial details were the most important data relative to the total amount of money disposable.
Photos of family and friends, personal documents and photos of oneself were sold most for the monetary value.

In sum, participants valued private data higher than job-related data

Experiment II: body talks. physiological and expressive reactions to data loss.

Research Methodology - overview

Sample

Similar to study 1, we recruited over a period of 2 weeks (January/February 2017) via online advertisements (e.g. Ebay classifieds), social media platforms (e.g. Facebook) and mailing lists. The resulting overall sample consisted of **52 participants** ranging in age from **18 to 62 years** (*M* = 30,50, *SD* = 10,63). **Female and male respondents were almost equally represented. The overall level of education was high** with a majority of students and employees with higher education entrance qualification.

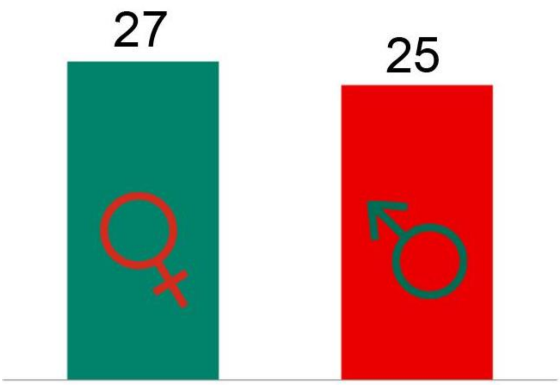


Figure 14: Number of participants by gender

The participants' age groups.

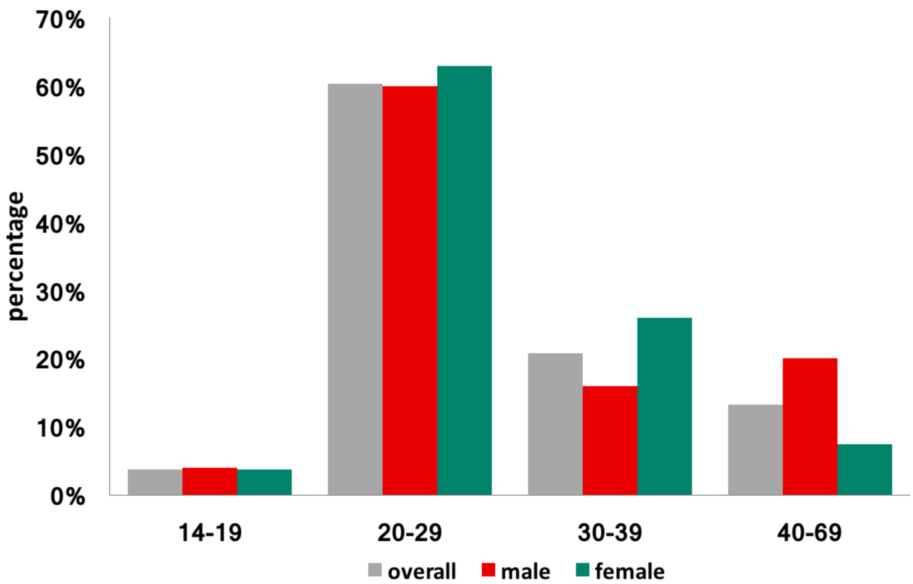


Figure 15: Percentage of participants by age group

Most participants were students or employees

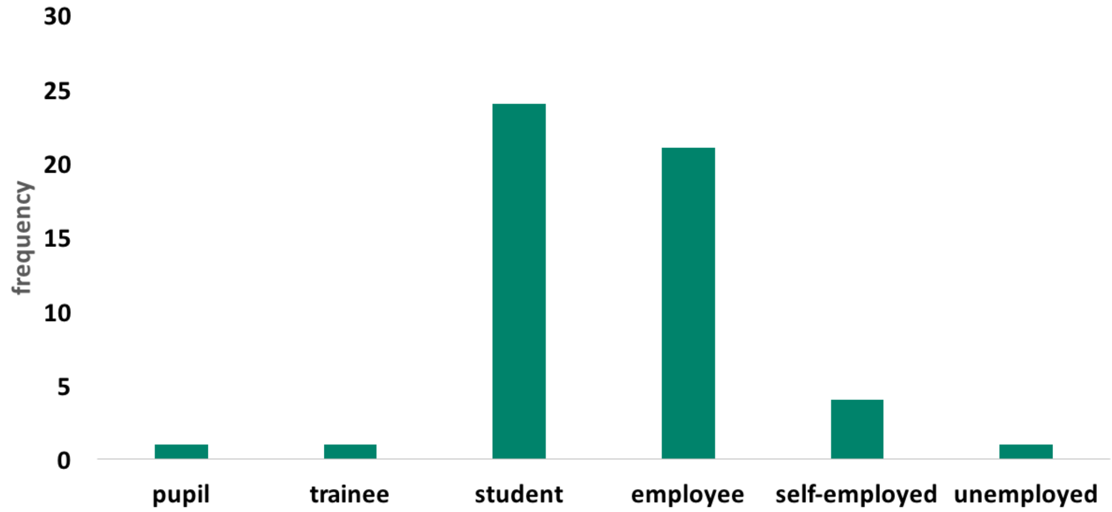


Figure 16: Number of participants by occupation

Procedure and Instruments

This study aimed for the measurement of emotional aspects of data value. Physiological and behavioral reactions were measured and interpreted as indicators of emotional processes. However, to evoke these emotional responses we needed to change perspective in terms of the definition of data value. While study 1 focused on the cognitive component (appraisal) asking for the monetary value of different data categories, study 2 took a more indirect path. **We simulated data loss on the participants' own smartphones and analyzed their physiological and expressive reactions to it**. To avoid real damages we needed to simulate this loss, of course. Therefore, the participants were told a cover story leading them directly to the loss of data. After the participants had been welcomed and introduced to the researcher and the "technical expert", they were told a **cover story about the apparent purpose of the study**: our institute had developed and evaluated a mobile web-app for improved search of online information; the alleged purpose of the study was to test this app.

Following this, the participant sat down at a desk with a wooden lectern on it. The researcher explained the various measurements that were allegedly about to be taken. Besides electrodermal activity data and video recordings (which were actually recorded), participants were told that their eye movements would be tracked. This was done in order to further legitimize the cover story. After the technician had finished placing the electrodes and sensor of the EDA system, the researcher took over again. To avoid any intrusions during the experiment and guarantee consistent experimental conditions, notifications were turned off on the participants' phones, and the display was set to a constant, appropriate brightness. They were then asked to open the URL of the web-app in their mobile browser which they were allegedly about to evaluate. **The app** consisted of:

Page 1-2: **Short overview** over the alleged purpose of the study and a mandatory-to-accept **disclaimer**, in which the researchers rejected any liabilities for damages to the phone.

Page 3: Last instructions and a **button starting** a 10 second countdown

Within these 10 seconds participants placed their **phone on the lectern**. To avoid any manual interaction during the experiment the researcher then placed a clear plastic shield over the phone (see picture below).

Page 4: An automatically changing sequence of texts and images as alleged learning stimuli: Four text snippets and one picture for each of three neutral topics; duration: 256 seconds

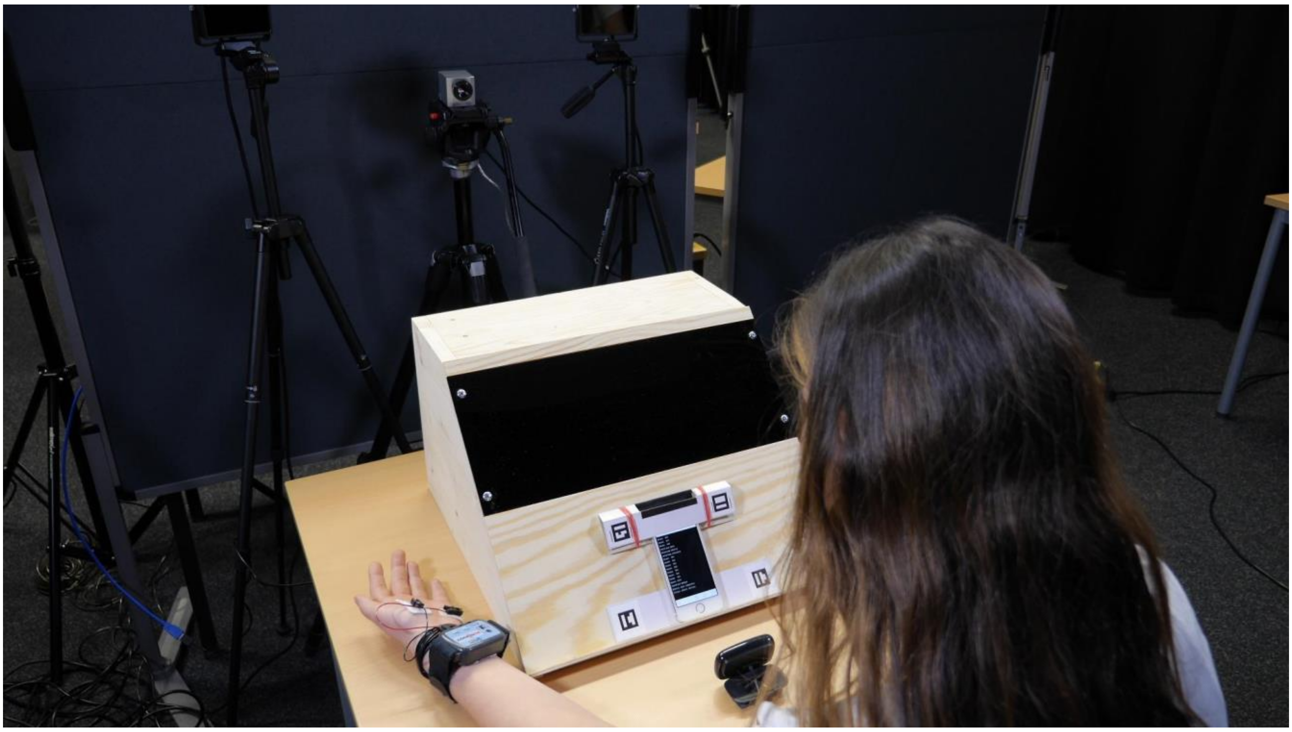
Page 5 **Black "crash screen"** with command prompt style error messages; duration 59.5 seconds (see stimulus)

Now, participants needed to be prompted not to leave the place they were sitting. This was important to allow a sufficiently long period for data acquisition as well as to ensure comparable conditions. Therefore, the researchers commented on the screen events with predefined, timed phrases as soon as the app crashed:

10 seconds researcher told the participant to "remain seated, because we are looking into it". In case the participant had addressed the researcher earlier, they were told to "wait a moment".

next 60 seconds similar short, neutral statements

Afterwards, the researcher explained that the crash was simulated and assured them that at no time, personal data was deleted or even accessed. Participants were then asked to answer a questionnaire. Finally, they were asked concluding questions, were informed about the true purpose of the experiment, and were shown the data that had actually been acquired.



Stimulus

The stimulus was designed as a website hosted on Firebase. The first part of the alleged web-app presented information and instructions on the study, and required user actions for proceeding. The second part followed a previously defined time schedule, thus not requiring user actions, ensuring comparability across participants, and minimizing measurement interference. In the first 256 seconds of this automatic phase, the web-app showed text snippets and images about neutral topics, including the history of potatoes, the history of paper and meteorological phenomena. This part was the same in both the experimental and the control condition. After 256 seconds, a black screen appeared and command prompt style error messages started to scroll on the screen, again following a predefined timing. The wording, number of lines, and outer appearance of the error message was identical for the two conditions, except for chosen manipulations of the message content:

1. In the experimental condition, the phone allegedly tries and fails at writing to the physical memory of the phone. A reset operation of the whole phone memory is being simulated and **the messages "deleting data", "deleting photos" and "deleting contacts" is being displayed repeatedly**. Based on findings from the previous study, this was designed to signal loss of the generally most valuable personal data on the phone.
2. In the control condition, the phone allegedly fails to write to the physical memory of the phone on the first try, but succeeds on subsequent tries. A reset operation of the device cache is being simulated and the message **"deleting call history" is being displayed repeatedly**.

Based on findings from the previous study, this was designed to signal loss of one of the least valuable categories of personal data.

As has already been outlined in a previous section, emotions are a multilevel phenomenon, and can be assessed using various measurement approaches. For the purpose of this study, we used three objective measures for emotional processes: Electrodermal activity, facial thermal activity, and analysis of facial expressions.

Electrodermal activity

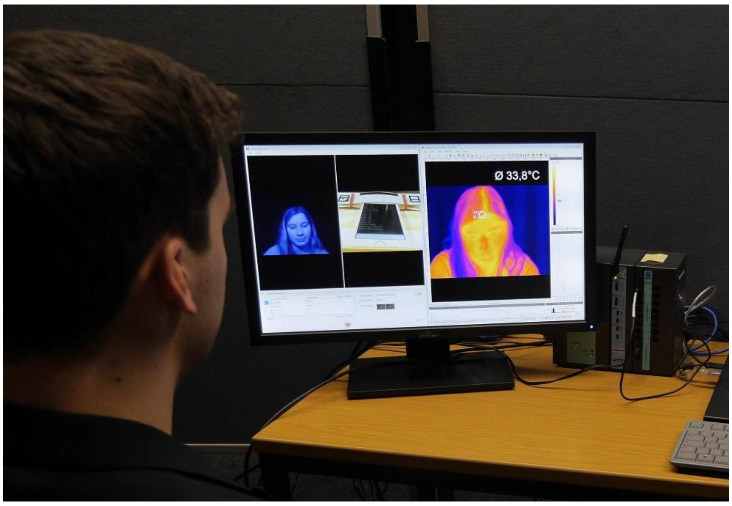
Electrodermal activity (EDA) is based on measurement of the skin's conductivity; it is directly related to activity of the sympathetic nervous system. EDA is a well-established measure for arousal and is frequently used in emotion research, e.g. on fear and anxiety. Research findings show that more stressful stimuli result in greater skin conductance responses compared to neutral stimuli (Butler et al., 2007; Khalfa, Peretz, Blondin & Manon, 2002).

We expected greater skin conductance responses for the simulated loss of valuable personal data, compared with simulated loss of trivial personal data.

For the acquisition of the EDA data, a BIOPAC system with a wireless transmitter was used. Two Ag/AgCl electrodes were placed on the palm of the participants' non-dominant hands. Boucsein (2012) suggests a sampling rate of at least 20Hz for the recording of electrodermal activity, the present study used a rate of 1kHz to achieve the best possible accuracy.

Facial thermal activity

Skin temperature of certain areas e.g. in the face is another objective measure for arousal. As the level of arousal increases, e.g. when experiencing an unpleasant situation, the activity of the nervous system causes peripheral blood vessels to constrict, so more blood is available for relevant organs. During this process, blood perfusion in e.g. hands, feet, and facial areas decreases, leading to a drop in skin temperature at these locations. To realize an accurate, but at the same time unobtrusive measure, we chose measurement via a thermal imaging system (Merla & Romani, 2007).



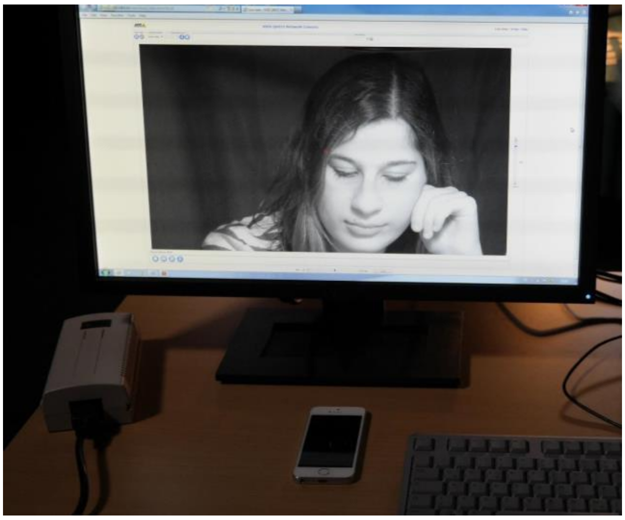
Compared with loss of trivial personal data, the simulated loss of valuable personal data should lead to more arousal and thus lower nasal temperatures.

The thermal activity in participants' faces was recorded with an Optris PI 160 infrared camera by Optris GmbH, Berlin, Germany. Radiometric videos were recorded with an optical resolution of 160x120 pixels and 10Hz, with a thermal sensitivity of 0.08K. For the analysis of recorded data, two measurements were taken from the onset of the crash screen and the end of the whole error sequence 60s later. For both moments, a rectangular measurement area was defined around the tip of the nose between the outer ends of the nasal wings. The minimal temperatures in these areas were extracted, so the change in nose tip temperature during the app's error sequence could be calculated for each participant.

Facial expressions

Emotion theories see our expressive systems as another domain of emotional effects. Facial expressions can easily convey our internal emotional state. For an experience of loss, sadness would be a reasonable response to expect.

For the simulated loss of valuable personal data, we assumed more expressions of sadness as an appropriate loss-related emotion, again compared with loss of trivial data.



The participants' faces were recorded with a Microsoft LifeCam Studio in a 720p resolution at 30 frames per second. To obtain footage from a viewing angle suitable for automatic face recognition, the camera was placed inside the lectern, right above the location of the participants' smartphones. To record the videos without the participants noticing, the camera's internal infrared filter had been removed and the camera was placed behind an opaque, but infrared translucent sheet of acrylic glass, allowing for near-infrared recordings of the face. The video material of the 60 seconds following the onset of the crash was then automatically analyzed with the facial recognition software SHORE, developed by the Fraunhofer Institute for Integrated Circuits in Erlangen, Germany. SHORE evaluates facial expressions along the dimensions of sad, happy, surprised and angry, and thus allows for an evaluation of the expressive valence in the course of time.

Results

This section reports the results for the three objective measures of the second experiment. Due to the technologically complex nature of the measurements, not all measures could be analyzed for all participants, e.g. when participants showed no measurable electrodermal activity, or when automatic video analysis was inconsistent due to participant movements. The number of participants involved in each analysis is reported in the analyses for each measure.

A Sweat-soaked data loss? How data loss affects electrodermal activity.

Losing important data causes stronger EDA, but overall only a small difference

Electrodermal activity (EDA) is related to activity of the nervous system and the skin's sweat glands, and is therefore often used in emotion research to measure fear responses. Stronger negative stimuli in general elicit more or stronger skin conductance responses (SCR). Consequently, we expected higher electrodermal activity, revealing itself as more and stronger SCRs, when loss of more important personal data is simulated.

Electrodermal activity was this study's most sensitive, but also most delicate measure. As in all studies with EDA, a certain percentage of participants had to be excluded from analysis because they showed no measurable EDA reactions. This phenomenon is not yet completely understood, but is well known in literature. It is assumed that about 10% of the population can be considered as hypo-responsive or non-responding in terms of electrodermal activity (Braithwate, Watson, Jones & Rowe, 2013). In our sample, five participants were classified as non-responders; additional reasons for exclusion from analysis were problems on the side of technology (3 participants due to equipment failure) and on the side of participants (1 participant consumed alcohol prior to the experiment, several participants did not comply with instructions and corrupted data by moving their hands). In addition, one participant's mobile device was too old and not suitable for stimulus presentation.

The resulting data of 39 participants was filtered and processed to calculate the so-called EDA Positive Change, an algorithm explicitly designed to offer a robust measure for longer lasting stimuli (Leiner, Fahr & Früh, 2012). Basically, the algorithm sums up all positive changes in EDA over a period of time, thus covering every arousal-related increase in the measure.

Loss of more important personal data tends to lead to stronger electrodermal activity (M = 4.77, SD = 3.55) than loss of trivial data (M = 3.90, SD = 2.24).

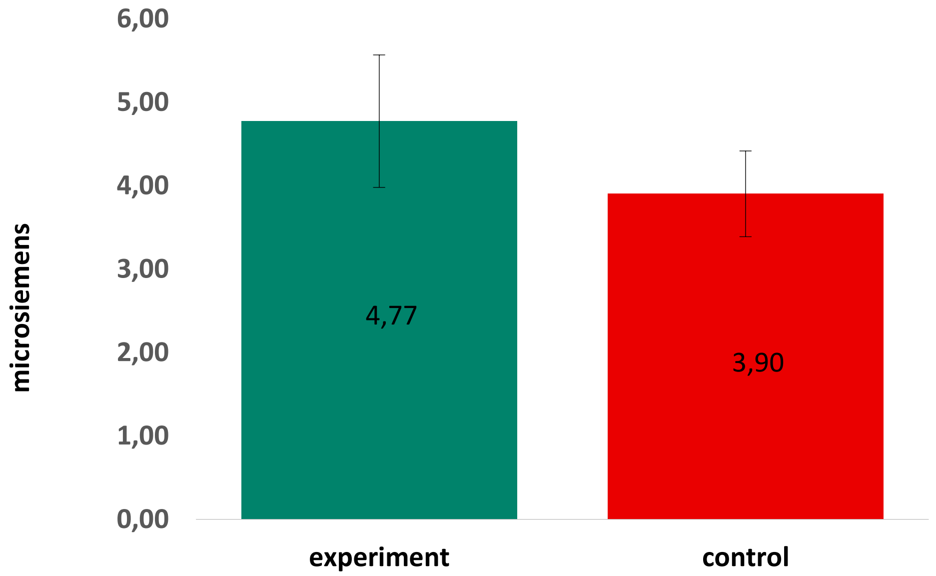


Figure 17: Positive changes in electrodermal activity, measured in microsiemens. Bars represent group means of EDA positive changes during the 59.5 seconds of error presentation (experiment: loss of important data; control: loss of trivial data). Whiskers represent ±1 standard error.

Statistical analyses suggest that the observed effect is too small to be reliably tested with our sample size. Comparing the groups with a t-test shows no statistically significant results ($t(37) = -0.91$, $p = .18$, *mean difference* = -0.87, *standard error of difference* = 0.95, *effect size d* = -0.29).

B The cooling effects of data loss - How data loss affects skin temperature.

Losing important data causes stronger thermal reactions, but overall only a small difference

Using an infrared thermal imaging system, temperature of the nose tip was used to assess arousal during the simulated data loss. A colder nose tip in this context indicates greater arousal, because constriction of peripheral blood vessels reduces blood perfusion in the nose, among others. Consequently, losing important data should lead to lower temperatures than losing trivial data.

Of the complete sample, 57 participants were included in analysis of facial thermal activity. Following suggestions from literature, participants had been given adequate time to get used to the environmental conditions in the laboratory prior to the measurements. Nevertheless, analyses show that participants in general showed increasing nasal temperatures over the course of error presentation, probably a side effect of outside winter temperatures. However, in line with hypotheses, participants who were confronted with the more negative stimulus tended to show a less pronounced increase, potentially due to the hindering influence of emotional arousal.

Loss of more important personal data tends to lead to lower nose tip temperatures (average temperature increase M = 0.07K, SD = 0.49) than loss of trivial data (average temperature increase M = 0.17K, SD = 0.39).

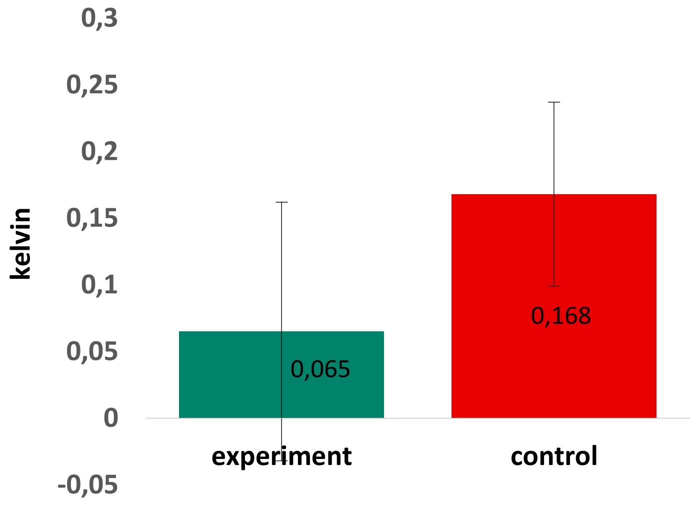


Figure 18: Change in facial thermal activity, measured at the tip of the nose (kelvin). Bars represent group means of temperature change during the 59.5 seconds of error presentation (experiment: loss of important data; control: loss of trivial data). Whiskers represent ±1 standard error.

The observed group difference was not statistically significant ($t(55) = -0.88$, $p = .19$, *mean difference* = -0.10, *standard error of difference* = 0.12, *effect size d* = -0.23), possibly also due to the rather short duration of the error message in face of slower reactivity of this measure.

C Facing data loss - How data loss affects facial expressions.

Losing important data causes stronger expressions of sadness, but overall only a small difference

Emotion is frequently expressed via our face, where many muscles form a very powerful system of emotional signaling. Vice versa, careful and systematic observation of the face allows researchers to draw conclusions about a person's emotional state. We used automatic, computerized detection of expressions of sadness, an emotion seen as an adequate response to experiences of loss. Experiencing the simulated loss of more important data should lead to stronger expressions of sadness.

Of the complete sample, 46 participants were included in the analysis of facial expressions. Expression analysis with the SHORE software resulted in one set of raw data for each participant, illustrating facial expressions of sadness during the alleged learning phase (phase 1), and the simulated error phase (phase 2). After data preparation, expressions of sadness during phase 2 were analyzed using the same algorithm as in the EDA analysis, just that this time, the positive changes in levels of sadness expression were summed up for each participant. The resulting individual values were then grouped into their respective experimental conditions, and group level analyses of the mean expression of sadness during phase 2 were computed.

Loss of more important personal data tends to lead to stronger expressions of sadness (M = 12.67, SD = 15.16) than loss of trivial data (M = 9.49, SD = 7.80).

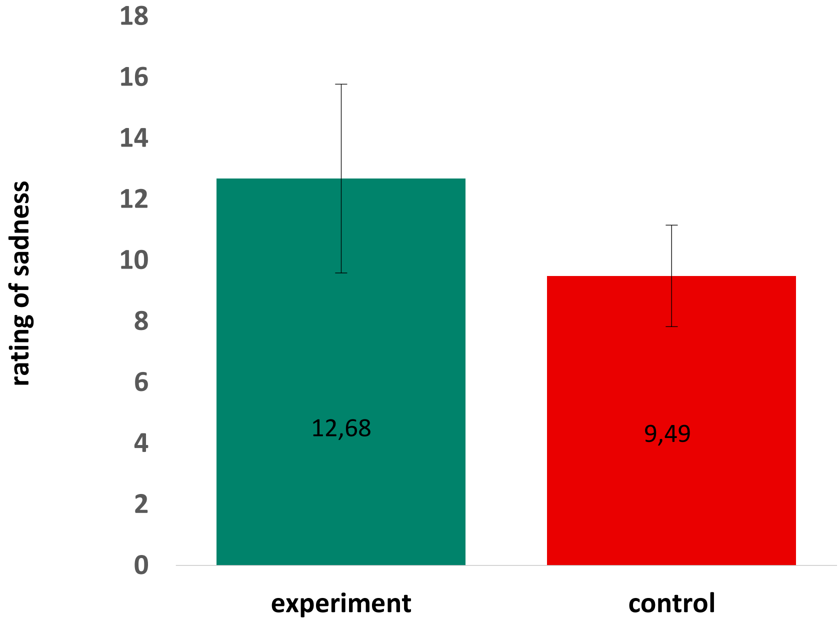


Figure 19: Positive changes in expressions of sadness, as scored by the SHORE software. Bars represent group means of positive expressive changes during the 59.5 seconds of error presentation (experiment: loss of important data; control: loss of trivial data). Whiskers represent ±1 standard error.

Statistical analyses show no statistically significant results ($t(44)$ = -0.88, p = .19, *mean difference* = -3.18, *standard error of difference* = 3.51, *effect size d* = -0.26) for the observed differences at the given sample size.

Stimulus evaluation

Contents and outer appearance of the alleged web-app had been fine-tuned based on feedback from 14 pre-test participants. However, it is unlikely that – without actually causing data loss – it will be possible to convince all participants of the threatening nature of our experimental manipulation. After all, the design of the simulated error message can only be a **tradeoff between scientific needs, ethical considerations, and technical feasibility**. In addition, it needs to have a believable appearance to convince at least some expert users (who are familiar with technical issues and e.g. error messages), while at the same time being verbose enough to convince most of the novice users (who are not familiar and rather overchallenged to elaborate the issue properly). Therefore, information needs to be easy to understand. However, information also needs to be realistic which often goes along with complexity.

Regarding the credibility of our data loss situation, not all participants perceived the simulated threat as a danger to their personal data. In post-session interviews, 90.91% of participants found the web-app's error message to be believable in itself; however, some participants described the error as not authentic for a mobile device (3 participants), thought of the error as an incident solely in our system, not affecting their smartphone (5), or already thought in the first place that this was a part of the experiment (1). When it came to the aspect of data loss, 32.73% stated that they actually feared the data was being deleted. The most common reasons given for not believing in a critical situation were: participants did not realize or did not properly elaborate the presented information (10 participants), researchers respectively a scientific experiment cannot really harm participants (10). Further reasons were objections towards technical feasibility (5), the understanding that the error was an isolated incident in our app (4), not understanding the content of the message (3), and a subjectively odd or unrealistic appearance of the error screen (2).

As a common limitation of self-report, these numbers might include cases of participants who actually did believe the data loss situation, but later reported that they were not impressed; vice versa, participants could also have not believed the situation, but then report that they felt in an experiment-conform way. In other words: people just do not tell the truth – for different reasons. Additionally, general users' judgements could be biased by their technological beliefs about mobile devices, e.g. they do not actually know if the given scenario was technically feasible or not, but rather assumed their devices' technology to be safe in general.

Experiment II: Summary/Take-away

Loss of data causes physiological reactions as expected for a negative experience.

All measures show trends in direction of theoretically grounded hypotheses

Loss of important data leads to…	**more electrodermal activity**
	lower nose tip temperatures
	facial expressions of sadness

However, across all measures, effects were small and not statistically significant in our sample

Experiment I and Experiment II: Final take-away

The three measures in study 2 show consistent trends, and are also consistent with study 1 results; this suggests that loss of personal data is of emotional relevance, although on a low level.
These results are in line with previous research on decision making which revealed that our decisions are anything but perfectly rational analyses. Frequently, our decisions are rather heuristic and serve as mental short cuts. These heuristic decisions are often efficient and useful. However, they may also lead to inaccurate conclusions:

- **We tend to ignore low-probability risks**
  ("It is not true because I cannot imagine it might happen").
- **We cannot evaluate probabilities in an accurate manner**
  ("1 in a 1 million is just as much as 1 in a billion").
- **Only if risks are imaginable and emotionally relevant, we will not ignore them**
  ("I can image it and it is important for me personally. It must be risky!").

Our study reveals that the risk of data loss seems to be perceived as a low-probability risk. Furthermore, participants have difficulties imagining that data loss might actually happen. Therefore, data loss is "not possible" and does not elicit strong emotional reactions. To assess this risks more properly and to avoid underestimation, people need a clearer understanding of what data loss means to them personally – regarding (1) the loss of emotionally relevant goods and (2) the functional principles of data loss. Both seem to be far from today's average smartphone user's knowledge.

References

Boucsein, W. (2012). *Electrodermal activity*. New York: Springer Science & Business Media.

Braithwaite, J. J., Watson, D. G., Jones, R., & Rowe, M. (2013). A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments. *Psychophysiology*, *49*, 1017-1034.

Butler, T., Pan, H., Tuescher, O., Engelien, A., Goldstein, M., Epstein, J., Weisholtz, D., Root, J.C., Protopopescu, X., Cunningham-Bussel, A.C., Chang, L., Xie, X.-H., Chen, Q., Phelps., E.A., Ledoux, J.E., Stern, E., & Silbersweig, D.A. (2007). Human fear-related motor neurocircuitry. *Neuroscience*, *150*(1), 1-7.

Dawkins, R. (1982). *The Extended Phenotype: the gene as the unit of selection.* W. H. Freeman, Oxford.

Khalfa, S., Isabelle, P., Jean-Pierre, B., & Manon, R. (2002). Event-related skin conductance responses to musical emotions in humans. *Neuroscience letters*, *328*(2), 145-149.

Leiner, D., Fahr, A., & Früh, H. (2012). EDA positive change: A simple algorithm for electrodermal activity to measure general audience arousal during media exposure. *Communication Methods and Measures*, *6*(4), 237-250.

Merla, A., & Romani, G. L. (2007, August). Thermal signatures of emotional arousal: a functional infrared imaging study. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE* (pp. 247-249). IEEE.

Merten, J. (2003). *Einführung in die Emotionspsychologie*. Stuttgart: Kohlhammer.

Muslukhov, I., Boshmaf, Y., Kuo, C., Lester, J., & Beznosov, K. (2012). Understanding users' requirements for data protection in smartphones. In *Data Engineering Workshops (ICDEW), 2012 IEEE 28th International Conference on* (pp. 228-235). IEEE.

Pew Research Center (2015). *The Demographics of Device Ownership.* Received from: http://www.pewinternet.org/2015/10/29/the-demographics-of-device-ownership/ [22.02.2017].

Scherer, K. R. (2001). Appraisal considered as a process of multi-level sequential checking. In K. R. Scherer, A. Schorr & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research* (pp. 92-120). New York and Oxford: Oxford University Press.

Scherer K. R. (2005). What are emotions? And how can they be measured? *Soc. Sci. Inf.*, 44, 693–727.

Scherer K. R. (2009). Emotions are emergent processes: they require a dynamic computational architecture. *Phil. Trans. R. Soc. B.*, 364, 3459–3474.

Schwan, S. & Hesse, F. W. (2004). Kognitionspsychologische Grundlagen. In R. Mangold, P. Vorderer & G. Bente (Hrsg.), *Lehrbuch der Medienpsychologie* (S. 73-99). Göttingen: Hogrefe.

Statista (2017a). *Number of smartphone owners worldwide from 2014 to 2020 (in billions).* Received from: https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/

Statista (2017b). *Percentage of U.S. cell owners in each age group whose cell phone has been lost or stolen in 2012.* Received from: https://www.statista.com/statistics/241365/us-cell-phone-users-whose-device-has-been-lost-or-stolen-by-age-group/

Statista Smartphone user penetration as percentage of total population in Western Europe from 2011 to 2018*https://www.statista.com/statistics/203722/smartphone-penetration-per-capita-in-western-europe-since-2000/

Appendix

Data categories and various values (existing on phone / considered important / deleted for money)

	number of participants who reported the data category to…		
data category	**… exist on their phone**	**… be important**	**… be deleted for money**
sensitive photos: myself	1	1	1
sensitive photos: partner	2	1	1
financial information	12	6	2
contacts	35	22	5
documents: personal	18	11	8
photos: general	41	18	4
photos: family, friends	37	23	9
passwords	12	6	5
emails: work	22	7	3
emails: personal	37	11	3
photos: partner	18	7	4
call history	41	8	5
scans: id-cards	7	3	2
documents: work	15	6	2
photos: myself	36	13	7
music	32	6	4
general videos	35	10	3
text messages	40	10	4
calendar	26	4	3
notes	24	3	1
apps	42	4	1

Data categories and various values (monetary value / relative amount of money / value for selling)

data category	mean of monetary value in €	amount of money distributed in %	mean monetary value if sold in €
sensitive photos: myself	30.00 €	12.00%	30 €
sensitive photos: partner	15.00 €	6.00%	30 €
financial information	13.33 €	14.46%	35 €
contacts	11.89 €	18.10%	15.2 €
documents: personal	10.56 €	10.19%	17.5 €
photos: general	10.37 €	13.37%	31.25 €
photos: family, friends	9.05 €	13.93%	15.56 €
passwords	8.75 €	18.27%	19 €
emails: work	8.18 €	5.82%	40 €
emails: personal	7.03 €	8.37%	23.33 €
photos: partner	6.11 €	7.22%	18.75 €
call history	5.85 €	8.46%	20 €
scans: id-cards	5.71 €	7.76%	10 €
documents: work	5.67 €	5.33%	15 €
photos: myself	5.00 €	6.48%	15 €
music	3.91 €	4.32%	21.25 €
general videos	3.29 €	4.86%	11.67 €
text messages	2.88 €	3.62%	11.25 €
calendar	2.69 €	2.01%	13.33 €
notes	1.71 €	5.83%	30 €
apps	1.07 €	3.57%	10 €