A User-Centered Perspective on Engaging with Digital Health Data

Dissertation

Daniel Diethei

Supervisor & First Reviewer: Prof. Dr.-Ing. Johannes Schöning
Second Reviewer: Prof. Jonna Häkkila, PhD

Faculty 3: Mathematics and Computer Science
University of Bremen

May 2022
Date of Colloquium: April 19, 2022
Ich widme diese Dissertation meiner Oma Marieluise.
Acknowledgements

First and foremost, I cannot thank Johannes enough. For your never-ending support, kindness, motivating messages, and for your true motivation to make me a better researcher and person. For giving me the time to think when I needed it. You made the office a place we all enjoyed spending time at and were never short of a joke. Thanks for all the rounds around the Unisee, where we extensively discussed Bremen-Bavaria politics and football :)

Without a great group of people, I could have never finished this dissertation. Ashley, so many times you helped me when I had trouble finding clear thoughts. You and Jonna were wonderful hosts in Rovaniemi. Tarja, for co-producing one of my favorite pieces of research. Thanks also to my LSC DiPH friends and colleagues, in particular Christopher, Tina, Rehana, Benjamin, Julian, Laura, for letting me learn about your amazing research fields and having a lot of fun along the way. The LSC allowed me to broaden my horizon in digital health, the field I’m most passionate about.

Ankit, I learned a lot from you (you are a genius), but more importantly you became a good friend. Jasmin, you brought so much energy to the group. You inspired my work and always found time to help not only me but everyone around you. Thomas, for exciting discussions and being an awesome landlord. Other members of the HCI group, Evropi, Gian-Luca, Dasha, Nadine, Caro, Philipp, Petra, it was a pleasure working with you guys. Also, I would like to thank Tobi from the University of Würzburg for first making me passionate about HCI. Last, I thank my parents and my sister for their endless support and trust.
Abstract

The number of patients with chronic conditions, the costs for modern treatments, and life expectancy have been rising. At the same time, physician shortages are anticipated. These developments put a burden on current health systems. Digital health technologies can make health care systems more efficient, more personalized, and contribute to reaching underserved populations. Essential for the success of digital health technologies is large-scale and rigorous digital health data that facilitates health promotion, prevention, early diagnosis, and management of diseases. Digital health data empowers individuals to make better-informed decisions about their health. However, current health technologies often fail to engage users to generate, share, and understand health data.

In this thesis, from the perspective of Human-Computer Interaction, we explore users’ needs when interacting with digital health data. We introduce the relevance of digital health data, describe our contributions from four papers, and discuss the implications of our findings for HCI and digital health.

We present the Digital Health Data Engagement Model (DHD-EM) and practical implications in the form of gulfs and bridges. Our model comprises the four stages lapse, generate, share, and understand. In the lapse stage, we identify reasons for a lapse of traditional health care and a shift towards digital health. This potentially happens when the health needs of patients are not fulfilled by health providers and patients consult online health communities for informational and emotional support. Our qualitative analysis of such communities showed when and how sub-communities for specific diseases emerge. In the generate stage, we explore physical and mental needs during the generation of health data. In three studies,
a survey, a qualitative field study, and a usability study, we investigated the generation of medical images from the user perspective. The results suggest that carefully considering user preferences, e.g., in relation to sensitive body parts, and adhering to design principles paves the way for easy-to-use and trustworthy applications. In the share stage, we investigate motivations to share health data. Common barriers to health data sharing are a lack of motivation and technical difficulties. From a citizen science perspective, we show that, in times of crises, collective motives are prevalent and present design implications for fostering participation. Lastly, in the understand stage, we describe how individuals make sense of health data. In online health communities, sense-making processes are mainly facilitated in long threads about specific diseases. In digital health apps, disease-related background information increases the trustworthiness in the diagnosis provided by the app.

Based on our studies, we identify gulfs in users’ experience when engaging with their health data. We map each of the gulfs to one stage of the DHD-EM. To overcome those gulfs, we provide bridges with concrete guidance to improve the design of technologies for emerging digital health areas, such as mobile health, wearables, and online health communities.

Our findings increase the impact of digital health technologies by allowing for a more nuanced understanding of the specific stages of users’ engagement with digital health data. We foster the agency of an empowered patient who wants to understand their health and participate in decision-making. Adhering to this user-centered perspective, we argue that the proposed model and practical implications improve users’ motivation and ability to share digital health data.
Zusammenfassung


Table of Contents

1 Introduction & Motivation ........................................ 4
   1.1 Background .................................................. 7
   1.2 Related Work ............................................... 9
   1.3 Contributions ............................................... 12
   1.4 Structure of this Thesis ................................. 17

2 Generating Medical Eye Images and Trust in Automated Diagnosis .... 19
   2.1 Introduction & Motivation ................................ 21
   2.2 Related Work ............................................... 23
   2.3 EyeGuide .................................................... 25
   2.4 Trustworthiness Study ..................................... 32
   2.5 Discussion .................................................. 37
   2.6 Conclusion .................................................. 40

3 Sharing Medical Photos ............................................. 41
   3.1 Introduction .................................................. 43
   3.2 Related Work ............................................... 46
   3.3 Method ....................................................... 50
   3.4 Findings ...................................................... 54
   3.5 Discussion .................................................... 62
   3.6 Conclusion .................................................... 67
# Table of Contents

## 4 Sharing Digital Health Data for Health Surveillance  
4.1 Introduction & Motivation ............................................ 70  
4.2 Background & Related Work ......................................... 72  
4.3 Research Questions ..................................................... 78  
4.4 Method ................................................................. 79  
4.5 Results ................................................................. 83  
4.6 Discussion ............................................................. 93  
4.7 Conclusion ............................................................. 100  

## 5 Understanding Health Information in Online Health Forums  
5.1 Introduction & Motivation ............................................ 104  
5.2 Related Work .......................................................... 106  
5.3 Method ................................................................. 109  
5.4 Results ................................................................. 111  
5.5 Discussion ............................................................. 119  
5.6 Conclusion ............................................................. 121  

## 6 General Discussion ................................................... 122  

## 7 Conclusion ............................................................. 126  

List of Figures ............................................................... 129  

List of Tables ............................................................... 131  

Appendix A  Corona-Datenspende-App: Interview Guideline  
Appendix B  Coding Tree: Corona-Datenspende-App
This thesis builds upon four papers that I have (co-)authored during my PhD studies. These are full papers and I describe my contribution in them below. Publications I, II, and III are published in peer-reviewed conferences, and publication IV is currently prepared for submission. Lastly, I list other publications that I have (co-)authored during my PhD studies, which are not included in the thesis.

**Included in Thesis**


I supervised the development and wrote parts of the EyeGuide code. I planned the user studies together with Lisa Dannenberg and Mohammad Fawad Jawaid Malik. I am the lead author of the paper.

II. Daniel Diethei, Ashley Colley, Matilda Kalving, Tarja Salmela, Jonna Häkkilä, and Johannes Schöning. “Medical Selfies: Emotional Impacts and Practical Challenges”. In: *22nd International Conference on Human-Computer Interaction with Mobile Devices and Services*. 8. New York, NY, USA: Association for Computing Machinery,
I co-designed the study together with Tarja Salmela. I conducted the study together with Matilda Kalving. I am the lead author of the paper.


I collected the data for the review analysis and prepared the data analysis pipeline. All authors analysed the review and interview data. Moreover, I contributed to all sections of the paper. All authors were involved in the writing of the paper.


This manuscript is currently under review at the Tenth IEEE International Conference on Healthcare Informatics. I collected all forum data and wrote the analysis tools. The analysis was done together with Ashley Colley and Julian Wienert. I am the lead author of the paper.

Not included in Thesis


This manuscript is currently under review at Nature Human Behaviour.


Introduction & Motivation

Medical breakthroughs of the future will be increasingly defined by our ability to collect, share and understand health-relevant data in vast quantities.

European Patients’ Forum

Digital Health is the use of information and communications technologies to improve human health, health care services, and wellness for individuals and across populations [133]. Digital health can empower consumers to make better-informed decisions about their health and provide new options for facilitating prevention, early diagnosis of life-threatening diseases, and management of one’s health and well-being outside of traditional health care settings [267]. For instance, individuals who use fitness trackers improve their participation
in physical activity [29], social media can positively impact health behavior [141], and digital health interventions lead to better medication adherence [197].

Essential to the success of digital health is data. Adopting the famous analogy of "data is the new oil" [191], Perakis and Coravos [195] suggest "health care data is the new blood" and argue that they should be handled with the same caution afforded to physical medical specimens. According to the European Patients' Forum initiative DataSavesLifes, medical breakthroughs of the future will be increasingly defined by our ability to collect, share, and understand health-relevant data in vast quantities [74].

In this thesis we refer to Digital Health Data as health data collected and processed on or with the help of electronic devices. Health Data are information related to health conditions, reproductive outcomes, causes of death, and quality of life in relation to the individual or the population [230]. Currently, the majority of studies focus on organizational sharing of patient health data in electronic health records systems for exchange between physicians, health care providers, and research institutions [2, 81, 80]. Health data sharing benefits the quality of care, reduces the amount of medical errors, and reduces costs [2, 81, 80, 165]. In other words, patients’ hesitation or inability to share may result in medical errors and undesired health outcomes [2]. Prior research suggested that privacy concerns [2, 192, 9], the health specific condition [15], understanding health benefits [11, 72, 192, 125], patient-physician relationship [2, 257], types of health information [72, 281], types of recipients (with whom data should be shared) [72, 290], and preferences regarding consent [72, 279] constitute sharing intentions and behaviour. However, we know less about the determinants of sharing in personal health technologies (as opposed to organizational settings). Accordingly, in this thesis, we adopt the concept of an empowered (also known as activated [265] or digitally engaged [156]) patient and explore their engagement with health data. This type of patient wants to understand better their health [22] and increasingly takes part in decision-making, ultimately reaching a more equal doctor-patient relationship [22]. In Human-Computer Interaction (HCI), user empowerment is operationalized by putting users at the center of the design process and by having their voice heard [226]. Human-Computer Interaction is concerned with the design, evaluation, and implementation of interactive computing systems.
for human use and with the study of significant phenomena surrounding them [102]. Even though HCI has a long tradition of health research [236, 147], the perceptible impact of HCI on experiences of users in digital health has been limited [22]. This lack of impact might be due to a missing focus on the broader cultural and technological factors that shape people’s experiences of health technologies in a complex and adaptive system [22].

Therefore, in this work, we use methods and concepts incorporating these broader factors, such as the lived experience and participatory design, to assess and design for users’ experience when generating, sharing, and understanding health data. We present the Digital Health Data Engagement Model (DHD-EM) that consists of four stages, *lapse*, *generate*, *share*, and *understand* (see Figure 1.1). The *lapse* stage refers to the shortcomings of conventional health systems and the resulting turn towards digital health. The stage *generate* describes why and how digital health data is produced. The *share* stage relates to the dissemination of data to peers and medical experts. The stage *understand* highlights deriving deeper health insights from engaging with one’s own and others’ health data.

Inspired by Don Norman’s *Seven stages of action* [175], we identify *gulfs* and suggest *bridges* (see Table 1.1) when users interact with digital health data. The *gulfs* refer to issues

![Fig. 1.1 The four stages of the Digital Health Data Engagement Model (DHD-EM). Individuals turn to digital health when traditional health care systems lapse. They generate health data, share it with others, and derive a deeper understanding when integrating their data with other peoples’ data.](image-url)
in current technology that prevent users from achieving the goals of the respective stage in the DHD-EM, e.g., from generating data. The bridges are implications and recommendations of our work to overcome the corresponding gulfs. The overall goal of this thesis is consequently to support HCI researchers and designers to develop more user-centered digital health technology in the future.

## 1.1 Background

As the number of patients with chronic conditions, the costs of providing modern treatments, and life expectancy is rising [166], physician shortages are anticipated [251]. To that end, digital technologies promise to make health care systems more cost-effective, efficient, and make medicine more personalized for patients, especially in countries that have significant

<table>
<thead>
<tr>
<th>Stage</th>
<th>Gulf</th>
<th>Bridge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lapse</td>
<td>A lack of access to expertise and social support in traditional care settings for users.</td>
<td>Accessible informational and emotional support in online health forums.</td>
</tr>
<tr>
<td>Generate</td>
<td>Physical and mental barriers of health technologies prevent data generation because the user is not at the center of design and therefore experiences a lack of ergonomics, usability, and user experience.</td>
<td>Identification and consideration of inherent user needs revealed through participatory and user-centered design methods.</td>
</tr>
<tr>
<td>Share</td>
<td>Users drop out of sharing health data due to a lack of long-term motivation.</td>
<td>Communication of societal relevance of data sharing, addressing collective motives.</td>
</tr>
<tr>
<td>Understand</td>
<td>Users prefer familiar doctors to make sense of their digital health data. When doctors can not help any more, users turn to online health forums.</td>
<td>Communicate the ability of digital health technology to provide specific information about the health problem at hand. If a physician is involved, visualize the patient-physician relationship.</td>
</tr>
</tbody>
</table>

Table 1.1 The gulfs and bridges of engaging with digital health data across the four stages of the DHD-EM. The table is based on the selection of use-cases from our papers. Following this inductive approach, we can examine some of the gulfs and bridges thoroughly but are aware that our observations are not exhaustive in terms of other aspects of interacting with digital health data.
1.1 Background

Gaps in their health care delivery [267, 293]. For example, digital mental health interventions have shown to be effective in low and middle-income countries [83]. Users interact with digital health technologies to better manage and track their health and wellness-related activities [267]. With mobile phones, wearable devices, and social networks, digital health provides innovative ways to monitor our health and gives us greater access to information. These advancements converge people, information, and technology to improve health care and health outcomes [267].

We understand Digital Health Technologies as tools to capture a variety of parameters directly relevant to individuals’ health. While we are aware that not all digital health technologies depend on data, e.g., medication adherence interventions, in this thesis, we focus on technologies surrounding health data. Advances in hardware and software have resulted in the proliferation of devices that allow individuals to produce ever-larger data streams across the lifespan, throughout health and illness, and in a geospatial context. In 2016, one-third of the global population used a mobile app or a fitness tracking device to track their health [85, 21]. Applications designed to collect, store, and analyze these personal health data are increasingly being used for self-tracking.

Nearly all of the technologies that handle personal health data are outside the mainstream of traditional health care or public health research [48]. Companies that develop health technologies range from small start-ups to global tech giants, e.g., Facebook, Google, and Apple. There is a growing willingness for individuals to share their health data with others [21]. This trend fueled the Quantified Self movement, where individuals share insights gained from their self-tracking activities [291]. Wearable digital devices and sensing technologies enable users to record data about their everyday activities and to obtain feedback in the form of numbers, graphs, and illustrations [5]. Examples of self-tracking devices include Fitbit Surge, Garmin Smartwatches, Pebble Watch, Apple Watch, and various smartphone applications such as MyFitnessPal, Fitocracy, FitStar, and Nudge. Intended to motivate users by encouraging a healthy lifestyle through daily monitoring, devices and apps record a wide range of biometric data, health indicators, and vital signs, including calories consumed,
distances walked, and hours slept. Measurements typically come from inbuilt sensors for automated data collection as well as self-reporting [5].

1.2 Related Work

In contrast to the well-being applications discussed in the previous section, digital health also encompasses health conditions such as chronic conditions. To navigate their health status, people with health conditions share their data with peers who have similar issues. In online groups such as PatientsLikeMe, Crohnology, webmd.com, and r/AskDocs members discuss shared health concerns and provide informational and emotional support. Motivations of individuals to share their medical history in an online health community are often not only to get clinical expertise but also to hear other patients’ personal experiences [108]. Other motivations for posting in the forum include not wanting to see a doctor, asking for immediate discussion before a scheduled doctor’s appointment, or a second opinion following a doctor’s diagnosis [108]. The trend for sharing extends to opening up personal health data to see what insights others might see in them, as exemplified by the Open Humans Project [178]. Some studies suggest that in addition to an egoistic motivation to learn about one’s health condition, collective motives, such as contributing to science, motivate people to share data [201, 53].

In our work, we want to learn when existing health systems cannot fulfill user needs. In particular, we aim to learn about users’ motivations who turn from traditional health systems to digital health technologies.

Previous work has modeled digital health tools that track physical activity data. In 2010, Li, Dey, and Forlizzi [148] coined the term personal informatics as a class of applications that help people collect and reflect on personal information. Further, the authors introduced a five-stage model of using personal informatics systems: 1) preparation, i.e., starting to collect data and thinking about the type of information to collect and how to collect it; 2) collection, i.e., gathering data; 3) integration, i.e., preparing, combining, and transforming data to reflect upon; 4) reflection, or examining and exploring data; and 5) action, or taking one’s newfound understanding of themselves to inform changes. Later, this model was
extended to *lived informatics*, i.e., considering challenges in users’ everyday lives, such as devices breaking or changing phones [71], and incorporating users’ individual preferences, such as goal-driven and documentary tracking or social rather than personal tracking [220]. Reflection has traditionally been a concept of discussion, e.g., it has been divided into maintenance and discovery [149]. Bentvelzen, Niess, and Woźniak [19] present a model that describes conditions and barriers of reflection on personal data. In contrast to previous work on personal and lived informatics, we want to investigate not only tracking activities, which typically span over longer periods, but also health data collected or shared only once or irregularly. Moreover, our scope of health data goes beyond physical activity data and includes data related to medical conditions, such as medical images. In other words, while existing research mainly addresses health promotion, our model additionally considers managing existing conditions, e.g., through engaging in discussions with other patients who have a similar condition.

Besides self-inspection and reflection, personal health data is shared with medical professionals or analyzed on personal devices. For instance, medical photos taken with mobile devices can be remotely analysed to detect skin lesions [32, 301]. Through teledermatology, a sub-discipline of telemedicine, users can monitor moles over time [174]. Such apps may include self, automated and medical practitioner assessment of images, e.g., Miiskin [167], Molescope [170], and Skinvision [238]. Several works have discussed emotional, legal, and privacy implications of sharing health data, e.g., medical photos from users’ mobile devices [206, 41]. From the health care provider perspective, problems arise when medical practitioners receive images in emails or instant messages and when patient-taken images are included as part of medical records [32, 301]. As many types of health data are sensitive, sending data digitally requires patients to build trust in the technology. Considering that applications through which patients distribute sensitive medical photos and potentially receive a life-changing diagnosis, building trust needs to be considered for design. Interactions between humans and digital devices can be personal and intimate experiences [84, 100, 270]. Compensating the lack of physical presence of a doctor, the visual, aural, and haptic aspects of digital devices become more important to support the emotional impacts of their use.
In our work, we want to contribute a more nuanced understanding of the physical and emotional challenges related to medical photos. In particular, we want to learn which body parts individuals want to share images of and which emotions arise during acquisition.

Having generated and shared such digital health data, it is vital to learn how personal health data is understood and contribute to insights about one’s own and other people’s health. The use of internet self-diagnosis has been noted as a source of conflict between patients and medical professionals, reducing satisfaction with medical professionals when they are later consulted [219, 237, 158]. Prior studies about searching online health information have shown that information is of variable quality [144, 282]. Online discussion forums may be a useful platform through which people can share their medical history and health data, ask health-related questions, and receive answers of acceptable quality [50]. Besides providing information or suggestions on diagnosis and treatment, Online Health Communities (OHCs) have also been shown to provide social support among patients with cancer [232], diabetes [109], rare diseases [142], infertility [280], and HIV/AIDS [169]. Previous studies described OHC members’ information-seeking practices [55], as well as ways they access and appraise information [97, 109], and construct new knowledge together [159]. Based on analysis of a diabetes forum, Mamykina et al. [159] describe the value of the forum as exposing individuals to the richness and multiplicity of different perspectives, which help participants to construct their personal views. While OHCs provide a platform for people to exchange health advice, health data can also be understood in more extensive and more structured data sets. When many individual health data sets are aggregated, they can be beneficial on a population level beyond the person that has shared them. One source for such data sets is data donations which have been explored in the domain of citizen science [26]. In public health emergencies, citizen science projects contribute to identifying outbreak hotspots. In the ongoing COVID-19 pandemic, many countries have released apps to donate physiological data to detect fever and eventually disease outbreaks [213]. In Germany, the Robert Koch Institute released the Corona-Datenspende-App which provides citizens with the opportunity to share their health data from fitness trackers and smartwatches to better understand the spread of COVID-19 [213]. Our goal in the context of understanding
health data is twofold. First, we aim to investigate factors that predict the informational and emotional exchange of health information and data in OHCs. Second, we want to learn about the motivations of people who share health data to contribute to mitigating crises.

In summary, many open challenges and questions surround the generation, sharing, and understanding of digital health data. We have seen that digital health has empowered patients to take responsibility for their health status. To serve these user needs and exploit the potential of digital health, designers should support users in their engagement with health data in personal health apps and online communities. The next sections describe our contributions to fostering such engagement.

1.3 Contributions

Individuals engage with health data at multiple stages during the use of health technologies. A prerequisite for turning to digital health is often a lapse of existing health care systems, e.g., when doctors cannot help with rare or chronic conditions. What typically follows is the generation of own health data, for instance taking photos of skin lesions. Next, users might decide to share their data, e.g., with family or doctors for consultation. Ideally, they will reach a higher level of understanding of their health status, which reinforces users to generate more health data. In the following, I will outline the Digital Health Data Engagement Model (DHD-EM) presented in this thesis and map the contributions to the stages of the model (Figure 1.2).

1.3.1 Lapse of Traditional Health Care Systems

A lapse of traditional health care systems occurs when the health needs of patients cannot be met in conventional care settings. We address this issue from the angle of online health forum participants who typically discuss their health status when doctors cannot help with their medical condition. The underlying problem is the lack of access to expertise and social support in the real world, which makes users turn to digital health, for example, when they join Online Health Communities (OHCs).
1.3 Contributions

Online Health Communities provide a knowledge exchange platform for a wide range of diseases and health conditions beyond in-person doctor visits. Informational and emotional support helps forum participants orient around health issues. So far, little is known about the relationship between the level of participation and participants’ contributions in OHCs. We conducted a thematic analysis of 456 posts in 56 threads from the Dermatology sub-forum of an OHC. In our initial exploration, we identified thread length as a potential criterion for community involvement. We separated short (fewer than 26 posts) and long (26 or more posts) threads for a more thorough analysis. This cutoff of 26 posts was based on an analysis of the distribution of thread lengths in an OHC. A more detailed analysis of a specific thread indicated group building elements emerged after approximately 26 posts. In the long threads, two reasons why participants looked for advice beyond that from medical professionals became apparent. First, doctors were not able to help. Participants reported...
being prescribed many different creams, tablets, management regimes, and even undergoing repeated painful and costly operations – often with little or no success. Second, participants reported that doctors were often unfamiliar with their conditions, possibly because the conditions addressed in the threads were rare. In these situations, patients turn towards health forums or health apps and consequently take ownership of their medical status to become *activated patients* [265]. From our dataset, we noted that patients did not receive enough medical support in the real world and found solutions to their health problems exclusively in the community. We also observed a lack of emotional support in the real world. Some posters reported relief having found other people with similar issues. The real world social support circles, such as family and friends, could not fully satisfy the participants’ health-specific needs. This gap was mitigated through psycho-social support exchanged in the forums.

In summary, we show that digital health technologies can alleviate some of the limitations of traditional health care systems, e.g., a lack of access to specialized knowledge about rare and chronic diseases. Sub-communities that particularly emerge in long threads of online health forums enable participants to provide psycho-social and informational support. We suggest that designers should consider the length of discussions so the requested health resources can be accessed efficiently.

### 1.3.2 Physical and Mental Barriers During the Generation of Health Data

When digital health applications require users to *generate* health data and do not put the user at the center of the design, individuals face physical and mental barriers. As a result, ergonomics, usability, and user experience problems arise. We conducted three studies to explore issues surrounding health data generation from the user’s perspective: (1) an online survey (n=100), (2) a qualitative field study (n=12), and (3) a usability study (n=28).

In a survey with 100 participants on Mechanical Turk to examine perceptions surrounding sending medical photos of intimate body parts, respondents expressed that they mostly prefer to see a doctor in person. However, respondents stated that selfie images were the
preferred analysis mode for other body parts and less severe skin conditions, such as rashes. In a qualitative field study with twelve participants, we discovered that medical photos of the shoulder blades and buttocks were challenging to take. Moreover, photographing intimate body regions provoke uncertainty, even when done for medical purposes. When such images are taken, the aesthetic dimension remains important, highlighting the need to accommodate individuals’ awareness of their body in the design of health apps. In the last study, we recruited 28 participants for a study to test whether the user experience of taking medical eye photos benefits from an interactive tutorial based on the principles of multimedia learning. The tutorial consisted of multi-modal, i.e., visual and aural, instructions. Indeed, we discovered that participants were faster in acquiring the eye images compared to auditory-only guidance. Given the recent development of AI-based diagnosis approaches, we argue that there are many potential application areas for guided eye imaging. The implications presented in our work can be integrated as part of the existing eye-disease screening and diagnosis tools. Overall, barriers to digital health data generation can be overcome by carefully considering user preferences, e.g., the emotional challenges of photographing sensitive body parts. Moreover, adhering to the principles of multimedia learning supports users’ physical abilities to acquire medical photos.

1.3.3 Motivations to Share Health Data

After health data is generated, many digital health technologies allow or even require users to share the collected data. Common barriers are a lack of motivation or problems while using technology. We addressed both aspects from a citizen science perspective. We analyzed the German "Corona-Datenspende-App" (CDA) as a timely research application area. The app allows users to donate their physiological data using wearables to mitigate the effects of the Corona pandemic by possibly detecting local fever outbreaks associated with COVID-19. We analyzed 464 app reviews and conducted ten in-depth interviews with individuals who had experience using this application. We thus could draw a picture of the users’ lived experience of participating in this citizen science project, illustrating the cycle of engagement of the citizen scientists and people’s motivations to contribute such personal data. Our contribution
in this work is fourfold: First, we reveal that collective motives prevailed over egoistic motives, driving the citizens to contribute to the CDA. Second, the users’ emotional response included both positive and negative emotions, and we demonstrated which aspects of using the CDA resulted in particular emotions. We also show the repercussions this had on the overall use of the app and continuation of contributing to the project. Third, we uncovered that citizen scientists of this particular project demonstrated a strong sense of persistence, i.e., a large proportion continued to use the app despite various issues "for the greater good". Finally, we present implications for the design and presentation of future citizen science projects, highlighting the importance of communication, transparency, and responsibility. We conclude that especially in times of crises, collective motives are prevalent and prompt prosocial behavior, e.g., peer support. Data privacy concerns and technical problems are regarded as less severe, emphasizing the role of developers to take extra responsibility.

1.3.4 Individuals’ Understanding of Health Data

Engagement with one’s own and others’ health data often leads to understanding medical conditions. In our study of online health communities (see also section 1.3.1), we observed that sub-communities around specific health conditions emerge in long threads in Online Health Communities. Many participants will share their own experiences, adding to the knowledge pool of the community and hereby support their self-reflection. This experience exchange often manifests as a relief that there are others with the same condition and creates a sense of community among the participants. The main values of the long thread discussions were therefore building collective knowledge and community among the thread participants. In our data, we identified that a strong motive in the long threads was the exchange of informational and emotional support. Participants encouraged the community to contribute their experiences to detect patterns, identify successful strategies, and spread knowledge. Participants shared their experiences primarily through detailed descriptions of their medical history. When a successful strategy was shared, other participants followed the advice. There was an active exchange on the strategies, often resulting in participants demanding updates on others’ proposed ideas, e.g., following up to ask whether they had tried certain medications.
Consequently, progress reports of successful therapies were also shared. On the other hand, short threads usually revolve around solving an individual’s health issue through diagnosis suggestions and medical advice. Here, knowledge construction processes could be facilitated by grouping many threads into categories so that related information is easy to find.

When health data is used for the diagnosis of diseases, individuals have a preference for familiar doctors to diagnose their medical selfies (see section 1.3.2). Therefore, we suggest that digital health solutions support the patient-doctor relationship by displaying information about the diagnosing doctor, e.g., through photos and a description text. Another option to increasing the trustworthiness of health apps is to provide specific information about the disease that is the subject of analysis. A particularly sensitive topic is the communication of diagnoses provided by AI algorithms. In a user study with 40 participants, we evaluated design alternatives to improve the perceived trustworthiness of health apps. Our results indicate that providing users with disease-specific information increases the trustworthiness and positively impacts the trust dimensions benevolence and integrity. To sum up, understanding health data in OHCs is facilitated mainly in long threads. We provide recommendations to support the informational and emotional needs of forum participants. We suggest delivering disease-related information for digital health apps to ensure trustworthiness and increase the likelihood of users maintaining the data generation.

1.4 Structure of this Thesis

So far, we have introduced the topics of digital health, digital health data, and summarised relevant work in the field of HCI and digital health. The contributions from each of the papers of this thesis are mapped to the DHD-EM in Figure 1.2. The consolidated contributions, in the form of gulfs and bridges, are listed in Table 1.1 and explained in section 1.3. The following sections consist of four papers (chapters 2, 3, 4, and 5). In chapter 2, we describe the generation of medical eye photos on mobile devices and discuss trust in an automated diagnosis. Chapter 3 presents the emotional and physical aspects of sharing medical photos. Chapter 4 presents an analysis of citizen scientists that share health data for COVID-19
health surveillance. Chapter 5 discusses the findings from the previous chapters about the Digital Health Data Engagement Model and relates them to prior work. Lastly, in chapter 6, we summarise our findings and give an outlook to future work.
Generating Medical Eye Images and Trust in Automated Diagnosis

Machines will not replace physicians, but physicians using AI will soon replace those not using it.

Eric Topol

This paper contributes to the stages generate and understand (Figure 2.1) of the Digital Health Data Engagement Model (DHD-EM). As an emerging application area of generating such health data we chose medical eye images that allow for the diagnosis of preventable diseases. We describe the physical complexities during the acquisition where users need to be supported in positioning the device. We further explore how to ensure that users understand
and trust the technology-mediated diagnosis after the acquisition, when the images have been analysed on the phone or the diagnosis is delivered by the physician.

We address design considerations that impact the usability and trustworthiness dimensions of mHealth applications. We present the design, development and evaluation of EyeGuide, a mobile app that assists users in taking medical eye images using only their smartphone camera. In a usability study (n=28) we observed that users of an interactive tutorial captured images faster compared to audible tone based guidance.

In a second usability study (n=40) we discovered that providing disease-specific background information was the most effective factor to increase trustworthiness in the AI based diagnosis. Application areas of EyeGuide are the AI based disease detection and telemedicine examinations.

Fig. 2.1 The DHD-EM and the contributions of the current chapter *Generating Medical Eye Images and Trust in Automated Diagnosis* (highlighted in blue).
2.1 Introduction & Motivation

According to the World Health Organization (WHO), 65.2 million people have cataract, a cloudiness in the lens of the eye, leading to increasingly blurred vision [295]. Of the visually impaired and blind, 90% live in developing countries and 80% of visual impairments are preventable [294]. In many countries, the poorest households are more likely to have access to mobile phones than to toilets or clean water [292]. With this background, mobile health (mHealth) technologies that can support the early detection of causes of preventable blindness should be explored with high priority. Retinal screening programs for common eye diseases can provide early detection of chronic eye diseases, but come at a high cost, e.g., a stationary slit lamp camera for traditional ophthalmology costs about $15,000 [225]. One step to increase the affordability of retinal screening is to use a shared community smartphone fitted with a lens adaptor, e.g., Peek Retina ($240) or D-Eye ($435). Such solutions can be applied by clinicians and nurses in areas with limited health care facilities and have been shown to perform well in disease detection [225].

Whilst the increasing availability of smartphones has resulted in a growing market for mHealth applications, e.g., in the areas of respiratory diseases [139] and dermatology [179], there is little evidence of positive clinical outcomes [264]. Studies have suggested that users require consumer-friendly devices and apps that are self-reinforcing and enjoyable to use [247]. Reporting on experiences in the UK’s National Health Service, [87] highlight ease of use of mHealth apps as a deciding factor in their adoption. A further barrier to the adoption of mHealth services is their perceived trustworthiness, e.g., in the competence of the provider, privacy or security of the service [10, 91, 126].
To address the issues mentioned above, we developed a user-friendly smartphone app for eye disease detection. In this paper, we focus specifically on the phases of image capture and the presentation of the diagnosis to the user (Figure 2.2). To account for a wide range of individual’s living circumstances, we target a solution enabling self-taking of medical eye images, and focus on the use of the smartphone’s higher quality rear camera. Whilst there is much prior work on automated diagnosis of eye disease from images, e.g., using deep learning [297, 278, 136, 189], the effectiveness of this is limited if the source images for analysis are of poor quality.

The contribution of this work is twofold as we connect the fields of usability and trustworthiness in the context of mHealth. As suggested by others [224, 67, 127], the usability of mHealth applications impacts trustworthiness, i.e. users find applications that consider usability principles to be more trustworthy. Therefore, we describe the design, development, and evaluation of the multi-modal EyeGuide app, that guides users in the capture of eye images suitable for detection of eye diseases such as cataracts. To support the task, the app includes a tutorial phase and uses auditory and visual guidance to direct the user to position...
the camera correctly. As a second contribution, we present a study exploring different approaches to increase the perceived trustworthiness of the eye disease diagnosis app.

From our evaluation, we report that the interactive tutorial had advantages in terms of usability of quality of taken images compared to the auditory guidance. A review of the taken images by an ophthalmologist indicated that the images taken with our app are suitable for diagnosing diseases in the conjunctiva and cornea. The most effective way to build trustworthiness in the app was through increasing users’ knowledge level about the disease.

2.2 Related Work

Our work in the area of intelligent user interfaces builds on three areas of prior work on (1) smartphone based mHealth for eye diseases, (2) guiding image capture, and (3) trustworthiness in mHealth. Firstly, we introduce existing mHealth applications addressing eye-related diseases. We then present works with different approaches to guide users when taking photographs, e.g., ensuring the subject is correctly located in the frame. Finally, we give a short overview of issues surrounding the perceived trustworthiness of mHealth applications.

2.2.1 Smartphone-Based Eye Disease Diagnosis

Prior work has presented smartphone apps to diagnose disease based on eye-images, e.g., for cataract detection [205, 193], to identify high cholesterol levels [8, 135], to diagnose concussions [161] and for glaucoma screening [135, 90]. Akil and Elloumi [6] present a meta paper, investigating the image quality and diagnosis performance achieved in eight prior works using smartphones equipped with additional lenses for retinal examination. Most recently deep learning has been presented as an approach to identify eye diseases [297, 278, 136, 189]. For example, Wei, Sehgal, and Kehtarnavaz [278] presented a deep learning based smartphone app to identify retinal abnormalities. The system gives simple real-time textual feedback ‘normal / disease detected’, requires the use of an additional D-Eye lens fitted to the smartphone and cannot be self-administered. Kim et al. [129] developed and
2.2 Related Work

evaluated an automated smartphone-based system for retinal disease screening using a 3D printed housing for acquiring a series of seven images stitched together to a widefield retinal montage. Munson et al. [171] developed a free smartphone app able to identify ‘white eye’ and subsequent eye disorders based on casual photographs of small children. The authors reported the app correctly diagnosed 80% of children with eye disorders. EyeGuide builds upon work by Diethei and Schöning [62] who explored an agent-based tutorial for eye image guidance. In this paper, we present a study to assess the usability of auditory guidance and an interactive tutorial to take eye images without the need of additional lenses or other hardware.

2.2.2 Guiding Image Capture

Ensuring correctly positioned and high quality images of the eye is a critical element in the performance of the following diagnosis phases, either manually by a doctor or through AI based solutions. As the target of our approach is towards the creation of a self-administered eye examination tool, using the device’s higher quality rear camera, solutions to guide the user in image capture are needed. Similar challenges have been addressed in mHealth apps for dermatology. For example, in the Skinvision app [238] the user “waves” the smartphone over the skin lesion to be investigated, and the app automatically selects when to take the image for analysis. This unguided approach could potentially be improved on with audio feedback [59], e.g., similar to that used for navigation [103]. A summary of approaches to integration non-speech sounds to visual interfaces has been presented by [3], suggesting the use of earcons [24, 99], sounds with dynamic timber, pitch and rhythm, as suitable for localization tasks. To guide users to take better portrait photos with smartphones, [152] demonstrated a solution using verbal guidance e.g., "please move to the left", to which [14] added vibration modality.

2.2.3 Trustworthiness

The importance of the user interface in perceived trustworthiness has been reported by several works, e.g., [67, 127, 224]. Kim and Moon [127] have shown that the user interface design
2.3 EyeGuide

Factors impact customer confidence. Roy, Dewit, and Aubert [224] discovered that the usability factors ease of navigation, ease of learning, perception and support were associated with trustworthiness. In the field of eHealth, Vo, Auroy, and Sarradon-Eck [272] highlighted trustworthiness, appropriateness, personalization, and accessibility as the main weaknesses in current mHealth apps. Fruhling and Lee [82] developed a usability model for the consumer’s perception of trust in eHealth services, noting that websites with a higher usability were likely to be viewed as more trustworthy. Going beyond usability, others have highlighted the role of visual aesthetics in the design of mHealth apps in their perceived credibility [188]. However, mHealth app designers do not have free hands, the requirements for apps to meet regulatory requirements can impact the design, and consequently affect the perceived trustworthiness the app [43].

2.2.4 Summary of Prior Work

From the related work we learn that, whilst much research effort has been spent addressing automated eye-disease diagnosis from images, little research has addressed the activities preceding or following this in an operational mHealth solution. With a focus on these phases we, (1) demonstrate and evaluate two solutions using audio signals and an interactive tutorial to guide users to take optimal medical eye images (2) present design alternatives aiming to improve the perceived trustworthiness of mHealth smartphone apps. In particular our approach follows Akter et al.’s [7] dimensions of trustworthiness; ability, benevolence, integrity and predictability.

2.3 EyeGuide

The EyeGuide app guides the user to take high quality images of their own eye, using the rear camera of a smartphone. These images are suitable for analysis by a clinician or through an automated AI process. As in this configuration the smartphone’s screen is facing away from the user, the user must be guided to position the camera to fully capture the eye.
2.3 EyeGuide

2.3.1 Concept

To understand the issues related to clinical eye photography on a smartphone, we first instructed four individuals to acquire photos of their iris with their smartphone in a prestudy. The participants used the native camera application with both the front and rear camera. An analysis of the resulting images revealed that lighting and device positioning seemed to be the predominant causes of blurry images or images without a fully visible iris. The front camera did not provide sufficient image quality for our purpose of detecting diseases due to the lack of an autofocus functionality and the lower image resolution. Although the front cameras of the latest high-end smartphones are of higher resolution than those used in the test, considering our target populations in developing countries, the use of older, lower-cost devices with limited front camera resolution is more representative.

Based on previous work by Diethei and Schöning [62] and the cognitive theory of multimedia learning [47], two variants were implemented:

- Interactive tutorial (referred to as IT)

![Interactive Tutorial Diagram]

- Auditory Guidance

![Auditory Guidance Diagram]

Fig. 2.3 Flow chart of the two study conditions. Participants in the interactive tutorial group first watched a tutorial demonstrating the procedure to take eye images and then followed voice instructions. In the auditory guidance group, participants only read brief instructions to take an image of the iris and were then supported through voice and beep tone guidance.
• Reading instructions followed by voice and audio tone based guidance (referred to as ‘auditory guidance’; AG)

The procedures of the two app variants are illustrated in Figure 2.3.

In the interactive tutorial (IT) case, the process of positioning the phone is first demonstrated with an on-screen animation (Figure 2.4). After viewing the complete animation, the user is guided by voice instructions to complete each of the steps in turn. When the camera is in the right position the user is instructed to ‘stop’ and, after a countdown of three seconds, the camera flash light is activated to illuminate the eye evenly and the photo is taken. In the auditory guidance (AG) case, voice guidance is used for coarse positioning, e.g., ‘to the right’, ‘up’. When the eye is recognised in the middle of the screen, the audio output consists of a sequence of short tones, which become more frequent with better camera positioning (cf. Geiger counter metaphor, [103]). As soon as the camera is in the correct position, the audio tone becomes continuous, the camera flash light is illuminated and the eye photo is automatically taken. In both IT and AG cases successful acquisition is indicated through a

![Fig. 2.4 Tutorial to take the eye photos. Left (a): Arm outstretched in front of face. Center (b): Moving the phone closer to the eye until the word “stop” is played back. Right (c): Countdown and image acquisition with flash.](image)
camera shutter sound and vibration feedback. At the end of the process, an image review screen is shown to the user with the option to retake should the iris be blurry.

### 2.3.2 Implementation

*EyeGuide* is an Android app written in Java. To track the eye position we used the openCV library with the Haar Feature-based Cascade Classifier for Object Detection [180], which is an AI-based model for image recognition tasks. To find the optimal parameters for the eye detection (detectMultiScale method, parameters scaleFactor and neighbors), we took 50 sample images, 28 with eyes and 22 without, with the rear camera of a Samsung Galaxy S7. The images without eyes showed other parts of the face around the eyes. We then analyzed the classifications (correct, false positive, false negative) for combinations of the parameters. Finally, we chose the parameters with the highest accuracy and speed.

For the interactive tutorial (IT) case, graphics and animations were created using standard graphic design applications. To design the tutorial and the virtual agent with a focus on teaching motor skills, we adhered to the cognitive theory of multimedia learning [47], in particular the pre-training and segmenting principle. The voice instructions were recorded by author LD. To give feedback about the position of the eye in the auditory guidance (AG) case, the camera view was divided into five virtual rectangles; one in the center of the frame and four around the center. If an eye was recognised in one of the outer rectangles, speech guidance was given using the Android Text-To-Speech engine, e.g., “to the right”. The maximum frequency of the speech guidance was two seconds. When the eye was detected in the central rectangle, the guidance mechanism was changed to a sequence of short tones, which become more frequent with better camera positioning.

In both IT and AG cases, when the camera was optimally positioned, indicated either by ‘stop’ (IT) or a continuous tone (AG), a full-resolution photo was automatically saved to the phone. In the capture phase, guidance was based only on position, as the available computing power and lower quality camera stream did not support analysing the sharpness of the image in real time.
2.3.3 Evaluation

To evaluate the user experience of EyeGuide and its perceived diagnosis trustworthiness we conducted a user study. The study was conducted as a lab study and evaluated the two alternative designs, IT and AG (Figure 2.3). We recruited 28 participants and assigned them to two groups (IT and AG) of 14. The IT group had a mean age of 23, while the AG group had an mean age of 22. Both groups consisted of six female and eight male participants. Five participants of the IT group and seven participants of the AG group wore glasses to correct their vision.

The study sessions were video recorded and all data collected during the study was stored anonymously. The participants were first introduced to the target of the study and its process and were then asked to sign a consent form. Afterwards, participants completed a demographic questionnaire and the TA-EG questionnaire [123], which measures technical affinity. Participants were encouraged to think-aloud during the test. The smartphone running the EyeGuide app was handed to the participants and they proceeded to follow the instructions given by the app. At the end of the test, participants responded to open questions about the experience of using EyeGuide.

2.3.4 Results

We firstly compared the number of attempts and time needed to take an acceptable eye image in both interactive tutorial (IT) and auditory guidance (AG) configurations. We conducted independent samples t-tests with $\alpha$ set to .05. Participants in the IT group ($M_{\text{tries}} = 1.8$, $SD = .47$) and in the AG group ($M_{\text{tries}} = 1.3$, $SD = 1.19$) did not differ in tries to capture the first eye, $t(26) = 1.46$, $p > .05$). However, the duration needed to take photos of both left and right eyes was significantly lower, $t(26) = -2.17$, $p < .05$, in the IT group ($M_{\text{duration}} = 151s$, $SD = 115.3s$) than in the AG group ($M_{\text{duration}} = 247s$, $SD = 118s$).

Author LD rated the sharpness of the eye images by comparing them to a set of pre-defined reference images (Figure 2.5). The number of images rated with each quality level are shown in Figure 2.6. A majority of images was either completely sharp or sharp in both
Fig. 2.5 To rate the image sharpness in the EyeGuide study, we used a series of images as a reference. From left to right: completely sharp (1), sharp (2), slightly out of focus (3), blurry (4), completely blurry (5).

Groups (IT: 17, AG: 15) and only one image was completely blurry in each of the groups. No significant difference between the two groups was apparent, $t(26) = .437$, $p > .05$, with mean ratings for the IT group $M_{imageQuality} = 3.75$, $SD = .87$ and the AG group $M_{imageQuality} = 3.6$, $SD = .86$. Due to a technical error in the interactive tutorial condition, the image capture was triggered at a larger distance compared to the auditory guidance condition. This might have improved the image quality since a lower device-eye distance can cause blurry photos.

In the interactive tutorial (IT) group, the on-screen agent instructed the participant to watch the following steps before putting them into practice. This instruction was viewed and confirmed by every participant. However, as the tutorial started, six out of 14 participants tried to follow the instructions by turning the smartphone around while the first step of tutorial was shown. Two of the six realized that it was an instructive section of the tutorial when they were asked to press the ‘next’ button, and turned the smartphone back around to complete watching the tutorial. The other four appeared to be confused on how to press

<table>
<thead>
<tr>
<th>Interactive Tutorial</th>
<th>11</th>
<th>6</th>
<th>5</th>
<th>5</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditory Guidance</td>
<td>6</td>
<td>9</td>
<td>10</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Fig. 2.6 Distribution of image sharpness ratings. Each participant ($n = 28$) took two images, one for each eye. There were no significant differences in image quality ($p > .05$) between the two groups (IT and AG).
the on-screen ‘next’ button while physically following the tutorial steps. Eight of the 14
participants in the IT group had at least one occasion where the word ‘stop’ was not played
back, mostly due to incorrect positioning of the device in front of the face. The participants
then had to repeat the steps from the beginning. However, while the smartphone was being
moved back to the start position, sometimes one of the eyes was recognized, which triggered
the ‘stop’ instruction and resulting in participants capturing an image of their whole face.
Other problems noted in the IT group included participants thinking they should capture both
eyes at once, or not understanding that the photo would be taken automatically. Four out of
five participants wearing glasses removed their glasses to take the photo even though they
were not asked to do so. Participants were instructed to start taking each photo with a fully
outstretched arm, however, after going through the process once, most participants did not
fully outstretch when taking subsequent photos.

In the auditory guided (AG) group, participants first read the instructions before taking
any photos. After turning the smartphone around so they faced the rear camera, six of the 14
participants felt unsure about the correct distance from the eye to hold the smartphone. Four
of the participants were not sure if they were required to take the photo themselves or if the
photo would be taken automatically. In addition, eight participants were not able to interpret
the meaning of the beeping sound. Two of the participants said that they do not know what
to do when getting the voice instructions (to the left/right, up/down). Six participants turned
the smartphone at least once to recheck the written instructions. In the AG group, four of the
seven participants wearing glasses took their glasses off before taking the photos.

To summarise, the time to take images in the interactive tutorial (IT) version of EyeGuide
was lower than in the audio guided (AG) version. The sequence of steps shown in the tutorial
was clear to most study participants. However, we observed some usability issues, e.g.,
incorrectly carrying out physical instructions when participants had been instructed only to
watch in the pre-training step.
2.4 Trustworthiness Study

Having evaluated different approaches for the tutorial and image capture phases, we also addressed the diagnosis feedback phase of smartphone-based eye disease detection. To gain insight into issues of trustworthiness in our targeted automated eye disease diagnosis app, we conducted a second experimental study.

2.4.1 Study Design

As the eye diseases we target with EyeGuide are rare in healthy Western populations, for the purpose of our study we decided to simulate the detection of the more common ailment of color blindness. We did not conduct any actual diagnoses and chose color blindness as an arbitrary common condition that is related to vision. As a study probe, we developed a separate Android app that included four alternative design approaches (Figure 2.7) to building trust. Each of the alternatives is mapped to one or two dimensions of the trustworthiness

![Fig. 2.7 The four design alternatives of the trustworthiness study.](image)
model. The users’ interaction with the app was guided by an animated avatar, that represented the virtual ophthalmologist (Figure 2.8).

The four alternative designs included in our application were:

1. **No trust-building (baseline)**. This design represents the standard procedure for the assessment and diagnosis of color-blindness using the Ishihara test [113].

2. **Assessment-diagnosis mismatch**. This probe design examined if users trust the diagnosis even if there is a mismatch between the assessment and the diagnosis. Participants were asked to provide a speech sample to detect whether they have color blindness, even though there is no diagnostic relationship between these two modalities. This intervention is a negation to ‘Predictability’ from the trustworthiness model [7]. Predictability refers to the degree to which an application is expected to behave reliably by abiding to standard practices.

3. **Disclaimer**. With this design we aimed to find out whether informing the user that their data will not be used for wrong purposes and only with good intentions builds trust in the application. This targets the ‘Integrity’ and ‘Benevolence’ dimensions of the trustworthiness model. Benevolence refers to the good intentions of the application towards the user, whereas integrity refers to moral and ethical principles which an app should conform to. Therefore, to affirm both of these dimensions in the app, we display a disclaimer as well as terms and conditions at app start up.

4. **Disease information**. This design addressed the ‘Ability’ dimension, which according to [7] is the most significant in the trustworthiness model. Ability refers to skills and competencies of the app that encourage the user to accept it, e.g., the application contains the desired knowledge about the task it performs. Addressing this, our design provided information in text form about color blindness to the user before actually starting the assessment.
2.4 Trustworthiness Study

2.4.2 Evaluation

To compare the trustworthiness of our four design variants we carried out a user study. The study followed a between-subjects design with the independent variable trust-building feature (no trust-building, assessment-diagnosis mismatch, disclaimer, disease information) and the dependent variable trustworthiness. Participants were consecutively assigned to one of the four groups. Based on Akter, D’Ambra, and Ray [7]’s trustworthiness model we developed the following hypotheses:

- **H1**: Giving background information about the condition results in the highest trustworthiness score.
- **H2**: The lowest trustworthiness is observed when there is a mismatch between assessment and diagnosis.
- **H3**: The highest benevolence score is achieved through providing a disclaimer as well as terms and conditions.
- **H4**: Integrity is highest through displaying a disclaimer and terms and conditions.
- **H5**: The ability score is highest when giving background information about the disease.

At the start of the session, we informed participants about the study both verbally and in writing. Participants then completed a consent form and a demographic questionnaire.
2.4 Trustworthiness Study

After that, participants were given a smartphone containing the study app, which they opened and followed the instructions given by the app. At the end of the study session participants filled out a trustworthiness questionnaire and were given an opportunity to provide open feedback on the test process. During the study, the experimenter took notes. The trustworthiness questionnaire developed for the study was a version of that developed by [7], slightly modified for mHealth applications. The questionnaire consisted of multi-item scales with favorable psychometric properties. Each item in the questionnaire was measured in a structured arrangement on a 7-point Likert-type scale, ranging from 1 (strongly disagree) to 7 (strongly agree). There are 4 items for each of the dimensions of trustworthiness, and a final 4 questions directly investigating trust.

We recruited 40 participants ($M_{Age} = 26.2, SD = 6.0$), 18 of which were female and 22 male. Most of the participants were students of the University of Bremen.

2.4.3 Results

We first report findings on the overall trustworthiness score followed by the results referring to its sub-dimensions. To test the effect of the independent variable design alternative on the dependent variable trustworthiness, we conducted a one-way ANOVA with the significance level $\alpha = .05$. For further analysis of the differences between groups, we used planned contrasts. When referring to individual participants, we indicate the design alternative and the participant number within the condition, e.g., P1.2 is participant 2 in the no trust-building condition (no trust-building = 1, assessment-diagnosis mismatch = 2, disclaimer = 3, disease information = 4).

As expected, we observed differences in trustworthiness between the design alternatives (Figure 2.9), $F(3,36) = 3.84$, $p < .05$, partial $\eta^2 = .24$. When provided with disease information ($M = 5.33, SD = 1.26$), participants reported higher trustworthiness compared to the other design alternatives, $t(36) = 2.68$, $p < .05$. Also, a mismatch between assessment and diagnosis led to the lowest trustworthiness, $t(36) = -2.86$, $p < .05$.

While the integrity dimension score was affected by the design alternative, $F(3,36) = 5.23$, $p < .05$, partial $\eta^2 = .30$, displaying a disclaimer did not lead to the highest integrity
2.4 Trustworthiness Study

No trust-building (baseline)
Assessment-diagnosis mismatch
Disclaimer
Disease Information

Mean Trustworthiness

Fig. 2.9 Trustworthiness scores between design alternatives. Error bars represent standard deviation.

score, $t(36) = .79, p > .05$. The descriptive data indicate that providing disease information ($M = 4.63, SD = 1.29$) might increase integrity.

Similarly, the design alternative had an effect on the benevolence score, $F(3,36) = 3.08, p < .05$, partial $\eta^2 = .20$. However, contrary to our expectation, the disclaimer did not cause the highest benevolence, $t(36) = .54, p > .05$. Again, there are descriptive indications that disease information ($M = 5.55, SD = 1.49$) is more effective in increasing benevolence.

Lastly, there was a difference in ability, targeted through disease information, between the design alternatives $F(3,36) = 6.42, p < .001$, partial $\eta^2 = .35$. Consistent with our hypothesis, information about color-blindness led to higher ability, $t(36) = 3.18, p < .05$.

In the no trust-building group, seven out of ten participants complimented the app with statements like “nicely done” (P 2.10) and “fast” (2.9). When questioned whether they would trust the app or not only two of the participants said yes. Participant 2.2 said she might have trusted it if she didn’t know already that she is not color blind. Three out of ten participants in the assessment-diagnosis mismatch group were confused when presented with the results. Six participants questioned the diagnosis of color blindness based on speech samples. When shown the disclaimer and terms and conditions, seven out of ten participants read both screens. One individual skipped the terms and conditions text. All of the participants performed the Ishihara test correctly. One of the participants mentioned that they liked
2.5 Discussion

the app, especially the dialogue with the virtual ophthalmologist (P 3.6). Participant 3.5 commented that the app seems to have good intentions. On questioning them about their trust in the app, only participant 3.5 said that she trusted it, whilst participant 3.6 said she might trust it, but was a little hesitant.

We observed that nine out of ten participants in the disease information group read both the information screens, i.e. “what is color blindness?” and “what causes color blindness?”. Participant 4.2 skipped the “what is color blindness?” screen. Six participants agreed that the app contains knowledge about the disease. Three participants rated the app with comments such as “user-friendly” (P 4.4) and “nice user interface” (4.10). Five out of ten participants said that they trusted the app. Participant 4.2 said she would trust the app if it was used in cooperation with a doctor.

2.5 Discussion

Through two studies we firstly examined the usability of EyeGuide, a mHealth app to take medical eye photos and, secondly, identified design considerations to ensure trustworthiness of mHealth apps in general. While the evaluation of EyeGuide addressed the image capture phase of an automated eye disease diagnosis app, the trustworthiness study informed on design approaches to increase trustworthiness in the diagnosis feedback phase.

2.5.1 EyeGuide

As some participants did not realise that the photo would be taken automatically and that they should only capture one eye at a time, not all tutorial instructions were understood correctly. Similar confusions about the meaning of the voice instructions (to the left/right, up/down) may have been due to a misunderstanding if the object to be moved was the smartphone or their head. On some occasions, participants repeated steps or activated functions by mistake because they were holding the smartphone in an incorrect position. Participants in both groups were confused when they were required to confirm the taken image.
Participants that wore glasses had similar problems with both concepts. All the photos taken by participants wearing glasses were of poor quality. Particularly, all those rated as completely out of focus were taken by glasses wearers and other images captured by glasses wearers were either blurry or slightly out of focus. To remedy this, participants removed their glasses - in hindsight this should have been instructed by the app.

A possible reason for the interactive tutorial group being faster in taking the eye images than the auditory guided tutorial group is that the auditory guided tutorial group took longer to find the correct device position. In contrast, the interactive tutorial group just had to follow the given instructions. Additionally, auditory guided participants often turned the smartphone around to check the instructions again, which added additional time. The mean time difference between the IT and AG group to complete the two images was on average about 100 seconds. In a real-world scenario, this has probably implications on the frustration levels of users, potentially leading to drop-outs in the AG concept.

In the interactive tutorial, some participants had trouble distinguishing between the sections of the tutorial where they were supposed to only watch the instructions and the sections where they had to actively carry out the instructions, e.g., actually turn around the phone. While the agent instructed the participants to first watch the steps before following them together, this was apparently not salient enough. We suggest a multi-modal indication of whether to watch only or carry out the instructions, e.g., a voice and text instruction.

In the auditory guidance group, many users did not recognize the Geiger counter metaphor, i.e. a beep tone with shorter intervals as they approached the target position.

Overall, we believe that the interactive tutorial concept had a better user experience. While some errors were detrimental to task completion and image quality, the concept of providing instructions that have to be carried out afterwards has shown to be effective in the context of taking eye images.

2.5.2 Trustworthiness

Our results indicate that providing users with disease-specific information increases the trustworthiness and potentially also positively impacts benevolence and integrity. This is
2.5 Discussion

consistent with Akter, D’Ambra, and Ray [7] who regarded the ability dimension, i.e. the perceived skills and competences of an app that encourage users to use it, as the most significant. Therefore, we confirm our hypothesis H1. On the other hand, trustworthiness is lowest when the assessment and the diagnosis modalities do not match, we can confirm H2. The disclaimer did not lead to the highest benevolence and integrity dimension, meaning we reject our hypotheses H3 and H4. There are descriptive trends that suggest that disease information is the most effective factor to increase these two trustworthiness dimensions. Furthermore, participants reported the highest trustworthiness when shown disease information; we therefore confirm H5.

2.5.3 Diagnostic Limitations

Whilst we propose that EyeGuide requires less expert knowledge and provides better accessibility compared to existing smartphone adaptors for retinal imaging, it is important to note that the spectrum of detectable diseases between the two approaches differs. Through fundus images taken using smartphone adaptors, a range of diseases such as diabetic retinopathy, glaucoma, and age-related macular degeneration can be diagnosed [217]; with these images, algorithmic approaches can be use to predict cardiovascular risk factors [196]. Smartphone cameras without additional lens adaptors do not allow for retina inspection and can only identify symptoms of diseases that are visible on the outside of the eye (e.g., conjunctiva or cornea). However, as some of the diseases visible with the naked eye, e.g., cataracts, are preventable [294], there is an obvious need for better screening, education and intervention in eye care. We argue that easy access to EyeGuide through a smartphone and the possibility of detecting cataracts, which are responsible for 33% of global blindness, with smartphone images justifies further research on this topic.

2.5.4 Contribution and Future Work

Given the recent development of AI-based diagnosis approaches, we argue that there are a large amount of potential application areas for EyeGuide to be integrated as part of existing
eye-disease screening and diagnosis tools. The interactive tutorial, based on principles of multimedia learning, was successful in preparing most users to capture high quality images. Minor usability issues were identified in the user study and will be addressed in the following design iterations of EyeGuide.

We are one of the first to explore solutions for guiding image capture for eye images. While similar approaches in the domain of dermatology [238] include features to automatically trigger image capture, they do not provide any support for device positioning. Other authors [152] described the use of verbal guidance for taking portrait photos, however, our interactive tutorial (IT) approach combines a tutorial and verbal guidance as an integrated solution.

Based on our findings, we suggest that the designers of mHealth applications should include disease-specific information to ensure trustworthiness. While we focused on color blindness in our trustworthiness study, our findings are generalizable to other mHealth fields beyond eye-related diseases. As future work we plan to extend our study to encompass a broader sample, e.g., the elderly, and to explore the use of our EyeGuide approach as part of iris imaging for biometric authentication [260, 204, 203].

2.6 Conclusion

In this paper, we presented two studies exploring the usability and trustworthiness of eye-related mHealth applications. An interactive and multi-modal tutorial that demonstrates correct device positioning was successful in reducing acquisition time. An easy to use, self-administered app to support screening and diagnosis of eye-related conditions can increase eye care access, especially in developing countries. Providing disease-specific background information was shown to be the most effective intervention to increase trustworthiness in the diagnosis. Our findings are relevant to the designers and developers of mHealth applications in the area of ophthalmology and other medical fields.
In the previous chapter we presented findings addressing the *generate* and *understand* stages of the DHD-EM. We chose the field of medical eye imaging which is relevant due to the large number of preventable diseases that can be detected through simple screening measures. This chapter contributes more findings in the *generate* and *understand* stages (Figure 3.1) but takes a step back from the specific use case of medical eye images to the more generic *medical selfie*. This allows us to formulate contributions that are important in the teledermatology field. Further, we explain the significance of the relationship to familiar doctors to (virtually) deliver a trusted and understandable diagnosis. We discuss medical images taken with mobile phones by patients, i.e. medical selfies, that allow
screening, monitoring, and diagnosis of skin lesions. While mobile teledermatology can provide good diagnostic accuracy for skin tumours, there is little research about emotional and physical aspects when taking medical selfies of body parts. We conducted a survey with 100 participants and a qualitative study with twelve participants, in which they took images of eight body parts including intimate areas. Participants had difficulties taking medical selfies of their shoulder blades and buttocks. For the genitals, they prefer to visit a doctor rather than sending images. Taking the images triggered privacy concerns, memories of past experiences with body parts and raised awareness of the bodily medical state. We present recommendations for the design of mobile apps to address the usability and emotional impacts of taking medical selfies.
3.1 Introduction

The incidence of skin cancer has reached epidemic proportions in white populations and the trend is still rising [134]. Currently, between 2 and 3 million non-melanoma skin cancers and 132,000 melanoma skin cancers occur globally each year. One in every three cancers diagnosed is a skin cancer and, according to the Skin Cancer Foundation, one in every five Americans will develop skin cancer in their lifetime [185]. Early detection and treatment are essential in reducing mortality. While the technical equipment commonly used in this context has previously comprised expensive stereomicroscopes and digital dermoscopy systems, teledermatology has shown to be more cost effective [266, 151, 146] and at least as accurate in diagnosis [162, 273] compared to face-to-face consultations. Mobile teledermatology has good diagnostic accuracy for skin tumours [134, 137, 130]. Due to the low-cost infrastructure it is a convenient tool, especially in resource-limited settings. Examples of the successful implementation of teledermatology services are reported from the Netherlands with more than 130,000 consults between 2006 and 2015 [259] and from the Brazilian state of Catarina with more than 83,000 consults from 2014 to 2018 [274].

While the technology has evolved rapidly [146], the emotional impacts of sharing medical photos between patients and doctors have not been considered in literature to the same extent. Sending sensitive medical images through an abstract communication channel that is not as tangible as face-to-face communication requires patients to build a high level of trust in the technology. This has implications for the design of applications and systems through which patients collect and distribute sensitive medical photos, and potentially receive back life-
3.1 Introduction

changing diagnosis. In this context, we address the concept of affective atmospheres [157], whereby the combination of technology, spaces and actors contribute to the overall affective experience of an individual using a system. When dealing with health and technology use, people can be vulnerable and are exposed to strong feelings.

When there is no physical presence of a doctor, sensory engagement with the patient, an essential part of developing trust, is lacking and has to be compensated for. Interactions between humans and digital devices can be personal and intimate experiences [84, 100, 270]. In the medical context, the visual, aural and haptic aspects of digital devices become increasingly important to support the emotional impacts of their use [157]. Recent work exploring the use of digital resources for HIV, has identified that such considerations for

Fig. 3.2 Examples of study participants’ skin photos. Note: In the study protocol, the researchers did not have any access to participants’ photos. The above photos were specifically volunteered by participants for publication.
providing emotional support are currently missing from online medical services [235]. We propose that in situations where digital imaging for the purposes of disease detection takes place, a particular atmosphere is created that is dependent on, for example, the body parts being digitally documented by an individual for disease detection purposes; the current level of knowledge of the medical state of the particular body part being photographed for diagnosis; the existing relation with one’s body; the presence and form of a camera used to photograph one’s body; the presence or absence of other humans in the space; the familiarity or foreignness of the space and; the ways in which the images are to be forwarded to medical experts. Moreover, we propose a possibility that the usage of cameras of personal smartphones to take medical photographs of one’s own (private) body parts can create negative emotional states such as insecurity and discomfiture. This has a strong connection with the digital documenting of specific body parts that are considered as more sensitive, vulnerable or private to be taken pictures of.

The aim of our work is to explore the area of medical selfies, i.e. images taken of one’s self to record the state of a medical condition, in regards to their emotional impacts and practical challenges. We address the following research questions

• **RQ 1**: What are the emotional impacts associated with the process of taking medical selfies and how do they differ between body parts?

• **RQ 2**: What are the physical and practical challenges people face when taking medical selfies and how can mHealth apps be designed to overcome them?

The joint observation of the emotional together with the physical dimensions enables us to derive design guidelines specific to the intimate nature of medical selfies.

In this paper, we report our findings from an online survey (n=100) and an in-depth user study (n=12) on how people experience taking medical images of their body. Survey respondents expressed that they mostly prefer to not send medical photos of intimate body parts but instead rather see a doctor. However, for other body parts and less severe skin conditions, such as rashes, respondents stated that selfie images were the preferred analysis mode. Participants in the user study found it physically challenging to take photos of their
buttocks and shoulder blades and expressed most emotional discomfort with genital and buttock photos. Inspecting their bodies as part of the study procedure raised their awareness of skin features, e.g., dryness and moles. Self-examination made participants check on body parts they would not normally look at in their daily life. Participants were concerned about their appearance in medical selfies, even though they were aware of the sole medical purpose. We provide design guidelines (Table 3.6) that address practical challenges, e.g., correct device positioning and image acquisition, and emotional dimensions such as aesthetics and data privacy.

3.2 Related Work

As relevant prior work, we firstly summarise the current state of knowledge on selfies in general, focusing on emotional and practical aspects. We then briefly reflect on the trend towards online medical care, so called telemedicine, as one of the main drivers for medical selfies. We then detail related work looking at the medical selfie from a variety of angles. Finally, we highlight the contribution of our work to the current state of knowledge.

3.2.1 The Selfie Phenomenon

Nowadays, taking selfies, i.e. photographs of oneself, and sending them to friends, or posting them more publicly, is a common practice across much of society [25]. There is a large body of prior work on selfies, exploring a broad variety of the motivations and effects of the phenomenon, e.g. [231, 16, 154]. Sung et al. [250] identified four motivations for posting selfies on social media, attention seeking, communication, archiving, and entertainment. Higher levels of selfie activity have been reported as an indicator of higher body satisfaction and narcissism [49, 250]. However, high investment in selfie images correlates with decreased body satisfaction [153, 49]. Higher investment in the images can take the form of elaborate staging (e.g., environment and make up), time spent selecting photos for publication or effort in photo manipulation [163]. Cohen et al. note that over half of the participants in their study reported manipulating their selfies through image editing "sometimes" to "very often" [49].
Surprisingly, we were unable to find prior works specifically exploring the ergonomics of selfie taking. However, Arif et al. [12] have reported on the usage of selfie-sticks, noting their general unergonomic design and recommending improvements such as a wider diameter grip. Medical papers have reported occurrences of ‘selfie elbow’, caused by the excesses of holding the smartphone at arms length [46]. More seriously, distraction during selfie taking has caused injury or even death [289]. Selfie images may not present a fully accurate representation, being affected by lens distortions [275] and mirroring effects [31]. Issues of mirroring effects have been explored by Bruno et al. [31], who report a left cheek bias for ‘standard’ selfies.

3.2.2 Online Self-Diagnosis and Telemedicine

As an alternative to the traditional visit to a medical clinic, patients may seek self diagnosis and care using freely available online information sources i.e. ‘Googling symptoms’ [155]. Similarly many online discussion forums exist where patients may seek community diagnosis. For example Reddit hosts a number of specific ailment related forums including r/DermatologyQuestions [208], where photos and descriptions of skin-related issues are posted, discussed and in many cases diagnoses by others users are made. Brady et al. [28] highlight that such forums enable the building of trust in particular advisors, through observation of their postings over time. Others have reported that the use of internet self diagnosis can reduce satisfaction with medical professionals, when they are later consulted [219, 237, 158]. Patients’ inability to manage internet sourced information and its potential inaccuracy have been highlighted as the main reasons for patient-clinician conflict [237].

Professional clinician use of telemedicine has seen rapid growth in recent years, achieving a US market size of $26.5 billion in 2018, growing to a predicted $130 billion by 2025 [277]. The domain of telemedicine is rather broad, including e.g., mobile health platforms, real-time interactive services, store and forward services, and remote patient monitoring and addressing e.g., dermatology, oncology, psychiatric conditions and post surgery follow up. However, technology adoption challenges amongst medical staff and patients, particularly elderly patients, present barriers to the success of telemedicine services. [229]. Suggested approaches
to address the identified issues include staff training and alternating telemedicine and personal patient-to-provider interactions [229]. Teledermatology, a subset of telemedicine, is the practice of dermatology using information and communication technologies. Recently, Abbott et al. [1] summarised 16 teledermatology reviews and concluded that teledermatology is comparable to traditional in-person methods of delivering dermatological care. Most reviews state that teledermatology is similar in terms of diagnostic accuracy [146, 17, 45] and management and patient outcomes [17, 45, 145] with the exception of two studies that report lower diagnostic accuracy for teledermatology [78, 276]. Focusing on smartphone apps to address melanoma detection, Kassianos et al. [124] identified 39 such apps that included features such as education (in taking images of lesions), classification, risk assessment and monitoring change. Whilst some of the apps provided the possibility to send images to medical professionals, generally the apps were lacking clinical validation. Similar findings are echoed by Lupton and Jutel [158], who also note the routine use of disclaimers in such apps undermines trust in their diagnosis.

3.2.3 Medical Selfies

Prior work on medical selfies has addressed both their explicit function as a medical tool, particularly in the area of dermatology, as well as other indirect motivations and benefits for their use.

A concise overview of the medical selfie is provided by Burns [32], highlighting their use to document ailments with visual pathology in areas such as dermatology, rheumatology, ophthalmology and burns treatment [32]. In a meta review of skin self-examination (SSE) Yagerman et al. [298] highlight its benefits in reducing melanoma incidence and mortality. Patient demographics are noted as influencing willingness and ability to perform SSE. Older individuals may have limited visual acuity and poor flexibility, which, together with the background skin lesions that develop with age, result in low confidence in performing SSE[298]. For afflictions with transient symptoms taking selfies can provide proof or a memory to aid diagnosis and discussion with clinicians [262].
Several works have discussed legal and privacy implications of using medical selfies [32, 207, 301, 40]. Problems arise when medical practitioners provide guidance on taking selfies, when they receive images in emails or WhatsApp messages and when patient-taken images are included as part of medical records [32, 301]. Further concerns of data ownership and privacy are reported by Chao et al. [40]. On the other hand, Ray et al. [207], note that images taken by patients are not subject to the strict data protection rules of medical practitioner taken photographs.

An overview of the use of smartphones to take dermoscopic images is presented by Ashique et al. [13], who highlight consistent lighting as the main challenge in reliable assessment. Prior work has reported positively on patients’ ability to take high-quality dermoscopic images at home, using smartphones [160]. Taking selfie images of the back is a notable problem area, with solutions using two mirrors or a selfie-stick being proposed [52]. Whilst processes including the post-analysis of such images by medical practitioners have been shown to be beneficial, automated app-based evaluation have raised safety concerns [206]. In a recent work, Ngoo et al. [174] provide an overview of melanoma related apps in the app stores, reporting that the most common target of apps is to support users in monitoring their moles over time. Such apps may include self, automated, and medical practitioner assessment of taken images, e.g., Miiskin [167] Molescope [170] and Skinvision [238]. Common app user interface features include reminders to re-assess lesions and side-by-side presentation of historical and current images of the same lesion. A key challenge in lesion imaging is ensuring correct and repeatable camera positioning, this has so far been addressed though automatic shutter activation, which may be enhanced with audible positioning guidance [62].

In addition to the direct role of medical selfies in supporting a patients physical treatment, several works have investigated positive impacts of the process on patients’ mental state [258, 298, 207, 33]. Tembeck [258] discusses the use of medical selfies as a form of self expression, enabling individuals to publicly identify themselves as living with illness, and to highlight its centrality to their daily life. A common finding from several studies is the positive effect of selfie taking in encouraging patients to take ownership of their condition and treatment [298, 207, 33].
3.3 Method

3.2.4 Contributions

From the related work, we learn that selfie based solutions will play an increasingly important role in the prevention, diagnosis and treatment of skin lesions. Technical issues, e.g., image quality, have have been well addressed, and legal and privacy issues at least identified as requiring further study. Several threads related to emotional and experiential aspects of taking medical selfies have been opened by prior work, and remain largely unresearched. For example, there is little research on the presence of emotional impacts similar to those of normal selfies, ownership of one’s condition and treatment, and trust in those assessing the images, be they medical professionals, unknown individuals in an online forum or algorithms in a mobile app. As a contribution, we aim to take steps towards understanding the interconnections between these underlying experiences, and deliver knowledge enabling the creation of improved patient experiences.

3.3 Method

We collected two complementary datasets. Firstly, we collected data from an online survey (n=100), aiming to gain a broad overview of the issues surrounding the topic. Secondly, to gain deeper insights, we conducted an in-depth user study (n=12).

3.3.1 Online Survey

We developed an online survey exploring experiences and attitudes towards teledermatology, particularly aspects requiring taking photographs of one’s own body. As well as demographic data and information on the participants’ general approach to technology usage, data on feelings of body esteem were collected e.g., ‘I feel good about my body’. Gender specification was voluntary and followed the guidelines by Spiel et al. [244]. All participants reported their gender. The survey included 5-point rating scale questions, e.g., to address comfort levels of taking photos for different body parts, as well as two open questions, (1) in which
3.3 Method

situations respondents prefer to have a face-to-face examination with a physician and (2) in which situations respondents prefer to send photos to physicians.

The survey was distributed via the Amazon Mechanical Turk crowdsourcing marketplace, with the only participation criteria being that the respondents’ location was in the US. Participant compensation of 1$ was provided based on the estimated task time of 7 minutes and the average minimum wage [168]. The survey resulted in data from 100 participants ($M_{\text{Age}} = 34.2, SD = 9.1; 43\% \text{ female}, 56\% \text{ male}, 1\% \text{ non binary}$). The average completion time was 4.7 minutes ($SD = 3.6; \text{ range: 1.6 to 31.0 minutes}$). We received 111 responses, of which we excluded eleven due to missing data. The average word count for each of the two open questions in the survey was 11.23 words ($SD = 15.27$).

Respondents’ textual responses to open questions were analyzed using an open coding approach [248]. One researcher defined the codebook, two coders independently evaluated each response, and a third researcher arbitrated disagreements between the coders. Answers were coded such that each answer could produce codes in multiple categories.

3.3.2 In-Depth User Study

For deeper qualitative insights we conducted an interview based study with twelve participants. Participants were recruited through local online classified ads and flyers placed in university cafeterias. The intimate nature of the photos led to a challenging recruitment process. Due to the rather complex study procedure, two participants dropped out after the briefing session. The recruited participant sample was gender-balanced (6 male, 6 female) with half residing in Finland and half in Germany. The mean participant age was 29.66 years ($SD = 8.63, \text{ range 20-41}$) and they were all familiar with using smartphones ($M_{\text{experience}} = 7.42 \text{ years, } SD = 2.31$). Ten participants were university students, one was in a full-time working position and one was on parental leave. Participants were rewarded with a cinema ticket as compensation for their participation.

The in-depth user study process consisted of the following stages: 1) Initial briefing and equipment setup, 2) Taking medical selfies, video diary and questionnaire, 3) Follow up
3.3 Method

A key tenet of the study procedure was maintaining the participants’ privacy. Thus, taken images were at no point shown to the researchers.

At an initial meeting with the participants, the purpose of the study was explained and participants were instructed on the procedure. Informed consent, developed under guidance from the university’s ethics advisor, was signed by the participant. When required, participants were assisted in the installation a voice recording app, required for part of the study. To ensure the privacy of taken images was maintained, participants were given instructions on how to disable the automatic upload of photographs to cloud services, such as Google Photos and Apple iCloud.

Later, in their own homes, at a time and place they felt comfortable with, participants followed a set of printed instructions to capture photos of their body parts. For the task setting, participants were asked to imagine that they had been requested to take photographs by a family doctor, showing moles or other skin conditions. A list of body parts to be photographed was provided. On each body part participants were instructed to capture skin features, such as scars, moles and veins, irritations. If no salient features were present, participants were asked to capture the full body part. During the capture process, individuals were instructed to think aloud and the audio was recorded on their device through a voice recording app running in the background. While the body parts needed to be fully visible, we made it clear that it was not necessary to completely undress. If a participant felt uncomfortable taking photos of any of the body parts, they were free to skip them. We asked participants to give an explanation for their reasoning in this case. Participants were asked to retain the images taken until after the final interview. Figure 3.2 shows a collection of example photos which were volunteered by participants for publication after the completion of the study.

Immediately after participants had taken photos of all the body parts on the list, we asked them to record a selfie video in a diary manner to give feedback on the experience, guided by a set of questions (Table 3.1). Finally, participants completed an online questionnaire that covered demographics, technological affinity [123], how physically demanding it was to take the photos.
3.3 Method

How did it feel to take photos of your body for medical purposes?

While taking the photos, have you recognized and photographed features on your skin you were already concerned about before the study? If yes, please elaborate which features there were and how you felt taking photos of them.

While taking the photos, did you discover new skin features you are now concerned about? Was there anything that surprised you? If yes, how did it feel discovering and photographing these features?

Describe the experience (taking photos of your body for medical purposes) in terms of privacy and intimacy.

Have the photos changed the perception on your bodily medical state? If yes, how?

Table 3.1 Questions study participants answered in a selfie video immediately after they took the photos.

The video and audio recordings were sent via email to the researchers. The photos taken were not sent or shown to the researchers. The following day, a semi-structured interview was conducted, either at the participant’s home or at the university. In addition to preset questions (Table 3.2), individual questions were included, based on pre-analysis of the individuals’ responses to the questionnaire. The interview was audio-recorded and transcribed.

As with the survey analysis, we used an open coding approach with three researchers for the audio, video and interview transcriptions. Since both Finnish and German researchers were involved in the process, the analysis was done based on the English transcriptions. The 36 transcriptions consisted of approximately 18 000 words.

3.3.3 Data Privacy & Ethics

Since the nature of our study required participants to take highly sensitive photos, we took a lot of care in minimizing discomfort and ensuring data privacy during the study. First and foremost, we did not see any of the photos the participants took at their home and communicated this clearly at the first briefing. Furthermore, we explicitly advised to turn off cloud services to prevent the photos being uploaded to the internet. While an image quality rating would have been helpful in assessing the clinical relevance of the images, we set the focus of our study on emotional and practical dimensions. Therefore, we decided
3.4 Findings

Do you feel comfortable sharing the photos with a doctor you know? How about a doctor you do not know?
Could you imagine having an app to diagnose diseases on the photos you captured?
How do you compare the atmosphere of intimacy between taking the photos at home and being at a dermatologist screening?
If you have used a mirror: Was it useful?
Do you have any experience in sharing medical photos with friends or family?
How does it feel having the photos on your device?

Table 3.2 Common interview questions asked one day after the capture session. Additional questions were asked individually based on previous statements.

To be as least invasive as possible to make sure participants acted naturally. Moreover, for ethical reasons, we advised participants that they could skip body parts should they feel uncomfortable capturing them. Only a minority of the participants actually skipped photos, which shows that there were few concerns with the legitimacy of the study.

3.4 Findings

In this section, we report the combined findings from our online survey and in-depth user studies. We refer to individuals as survey respondents and user study participants, e.g., P2 as the second participant of the in-depth study and R36 as respondent 36 of the survey. The most common themes of the in-depth study are listed in Table 3.3. Specifically we report on prior experiences with sending medical selfies, preferences for in-person consultations, ergonomics of taking the images, emotional effects, impacts of the taking selfies and privacy concerns.

3.4.1 Prior Experience with Sending Medical Selfies

Half of the study participants reported having previously sent medical selfies. One participant (P2) mentioned sending a photo of a lesion to a friend, who advised her to see a doctor. The lesion was then diagnosed as a benign melanoma. Another participant described her regular
exchange of medical selfies with her friends and family: “[W]e are always analyzing with a
friend or one of my sisters, like, hmm, what is it this time?”, (P3).

From the survey, 88% of respondents reported having shared photographs of their body
parts for medical purposes at least once. 9% of respondents reported having done so five
times or more. A $\chi^2$ test of independence revealed no significant effect of gender on the
history of sharing medical selfies, $\chi^2 (4, N = 99) = .52, p > .05$. To explore differences in
medical selfie content, we conducted an independent samples t-test on the effects of gender
on the range of body parts respondents reported having shared (Figure 3.3). In general,
female respondents shared images of more body parts ($M_{body\_parts} = 3.71, SD = 2.99$) than
male ($M_{body\_parts} = 2.65, SD = 1.89$) respondents ($t(97) = -2.169, p < .05$). Respondents
(n=57) had the most experience in sharing photos of arms (35%), legs (34%) and hands
(32%). For particular body parts, no differences in the frequency of sharing was found
between male and female participants (all Bonferroni-Holm adjusted $p$-values > .18). Instant
messaging was the dominant tool used for image sharing (72%), followed by email (18%),
cloud services (3%) and others (7%).

<table>
<thead>
<tr>
<th>Theme</th>
<th># of mentions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Privacy</td>
<td>40</td>
</tr>
<tr>
<td>Discovering skin lesions</td>
<td>30</td>
</tr>
<tr>
<td>Memories of past events associated with body parts</td>
<td>29</td>
</tr>
<tr>
<td>Skin self examination</td>
<td>17</td>
</tr>
<tr>
<td>Relationship to doctors</td>
<td>16</td>
</tr>
<tr>
<td>Assistance used (e.g., mirror)</td>
<td>15</td>
</tr>
<tr>
<td>Prefer to see a doctor in person</td>
<td>13</td>
</tr>
<tr>
<td>Aesthetics (both positive and negative)</td>
<td>13</td>
</tr>
</tbody>
</table>

Table 3.3 In-depth user study results: Most prominent themes identified from audio, video
and interview transcriptions.
3.4 Findings

Respondents reported sharing medical selfies with a broad set of those close to them, spouse/partner (54%), family (62%) and friends (43%). Only 2% of participants reported sharing such selfies with a doctor. Advice given by the confidants based on the medical selfies was evenly divided between recommendations to see a doctor (51%) and not to see a doctor (43%). The given advice was followed by 84% of respondents.

3.4.2 In-Person Consultation vs. Sending Photos to a Doctor

Almost two thirds (63%) of the survey respondents said that conditions related to their skin were suitable for diagnosis by sending photos to their doctor. Here, saving time was a commonly mentioned justification, e.g., “sending a photo to someone is quicker. You won’t waste time driving and doctor fees” (R34). Considering intimate body parts, 41% of the respondents preferred that they would be checked by a doctor in person, whilst only a minority (5%) stated a preference for taking photos and sending them for diagnosis. A typical comment expressing preference for an in-person consultation being, “genitals, breasts, anything that wouldn’t usually be photographed or sent, photos I would be embarrassed leaking out into the internet” (R14). A similar preference ratio was reported in the study with
only two participants (17%) preferring to take photos of intimate body parts at home, rather than visiting a doctor.

For many respondents the type and location of the medical condition affected their preference for interaction mode. Respondents stated that they would prefer to send photos for less serious conditions (mentioned by 30% of respondents), but would rather see a doctor for a serious condition (mentioned by 19%). Uncertainty of a condition’s seriousness was seen as a justification to send photos (mentioned by 8%), e.g., “for a dermatological issue that is most likely nothing, but I want to double check that it’s nothing.” (R7) and “maybe a mole that looked strange but I wasn’t sure if it needed attention” (R36).

For the most commonly mentioned skin conditions (Table 3.4), the preferred analysis mode was through selfie images. Respondents managed potential feelings of embarrassment in different ways. Some did not wish to take photos of sensitive conditions or body regions (mentioned by 4%), e.g., “I would prefer to send photos if it was something not too embarrassing or uncomfortable to share.” (R44). On the other hand, others (5%) wished to avoid showing and discussing such conditions and regions with a doctor in person.

When asked “I would trust a diagnosis based on medical photos from...”, a $\chi^2$ test of independence revealed that respondents have higher trust in a diagnosis by a doctor (69.5% agree or strongly agree) compared to an AI-based diagnosis (23% agree or strongly agree), $\chi^2 (4, N = 99) = 63.67, p < .001$. Moreover, a doctor they know (89% agree or strongly agree) is preferred over a doctor they do not know (50% agree or strongly agree), $\chi^2 (4, N = 99) = 47.81, p < .001$. Our user study revealed that a majority of participants (58%) did not have any preference between sharing medical selfies with a doctor they know vs. a

<table>
<thead>
<tr>
<th>Condition</th>
<th>Prefer Doctor</th>
<th>Prefer App</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rash</td>
<td>11 (26%)</td>
<td>31 (74%)</td>
</tr>
<tr>
<td>Cut</td>
<td>1 (12.5%)</td>
<td>7 (87.5%)</td>
</tr>
<tr>
<td>Bruise/Bruising</td>
<td>0 (0%)</td>
<td>7 (100%)</td>
</tr>
<tr>
<td>Eczema</td>
<td>0 (0%)</td>
<td>4 (100%)</td>
</tr>
</tbody>
</table>

Table 3.4 Survey results: Preferences for diagnosis mode by skin condition.
doctor they do not know. Three participants (25%) preferred a doctor they knew, while two (17%) said they would rather send images to an unfamiliar doctor. One participant made the distinction based on the intimate nature of the body part: “if it’s not the private body parts, I feel it’s fine to even share with a doctor I don’t know. [...] I mean, at the first time we can meet in person, so I know his face and the personality of him and then further I share” (P7). Another participant expressing his opinion, “the unfamiliar may be the nicer one, yeah, in a funny way, I wouldn’t know who’s receiving, they just open it, and they won’t know anything else about me” (P3). The same participant continuing, “it may be nicer If I knew the person, so I won’t have the feeling that now I’ve sent these photos and now I have to meet him [afterwards]” (P3).

The attitude towards an AI-based diagnosis from an app on their mobile device differed among participants. While a majority (58%) could imagine using an app to diagnose a condition from medical selfies, two participants added the provision that a doctor should decide in the end (P4, P5). Five participants (42%) were critical towards app based automated diagnosis, preferring a traditional diagnosis from a doctor.

### 3.4.3 Ergonomics and Image Quality

A few survey respondents were conscious of potential misdiagnosis based on selfie images, commenting e.g., “[I would see a doctor for] anything that looked worse in a photo than in person, so things where I just couldn’t get a clear and direct image or one that was clear and direct to the viewer...” (P47).

In general, participants used different strategies to reach the targeted body parts with the camera. For example, to capture the soles of the feet, P7 sat with their foot on a table, held in position with their non-camera hand, while P1 and P10 needed to sit down to take the photo. The study tasks revealed challenges in taking photos of some body parts, particularly the ones on the back of the body, i.e. buttocks and shoulder blades (Table 3.5). Here, some participants employed a mirror, one participant describing standing with their back facing the mirror, placing their hand behind their back and pointing the smartphone’s front camera towards the mirror. Nine participants (75%) reported they had difficulties capturing images
3.4 Findings

<table>
<thead>
<tr>
<th>Body Parts</th>
<th>Photo taken</th>
<th>Mirror used</th>
<th>Practical issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hands</td>
<td>12 (100%)</td>
<td>0 (0%)</td>
<td>2 (17%)</td>
</tr>
<tr>
<td>Forearm or Underarm*</td>
<td>12 (100%)</td>
<td>0 (0%)</td>
<td>4 (33%)</td>
</tr>
<tr>
<td>Nostrils</td>
<td>12 (100%)</td>
<td>4 (33%)</td>
<td>4 (33%)</td>
</tr>
<tr>
<td>Belly</td>
<td>12 (100%)</td>
<td>1 (8%)</td>
<td>1 (8%)</td>
</tr>
<tr>
<td>Buttocks</td>
<td>11 (92%)</td>
<td>1 (8%)</td>
<td>6 (50%)</td>
</tr>
<tr>
<td>Genitals</td>
<td>9 (75%)</td>
<td>0 (0%)</td>
<td>3 (25%)</td>
</tr>
<tr>
<td>Soles of the feet</td>
<td>11 (92%)</td>
<td>1 (8%)</td>
<td>2 (17%)</td>
</tr>
<tr>
<td>Shoulder Blades</td>
<td>11 (92%)</td>
<td>9 (75%)</td>
<td>9 (75%)</td>
</tr>
</tbody>
</table>

Table 3.5 Study results: Body parts vs. photos taken, mirror used and practical problems (n=12).

* Due to a translation error, German participants captured their forearm, while Finnish participants captured their armpit.

of the shoulder blades, e.g., stating, “that [shoulder blades] was definitely the hardest area to capture” (P1). On the other hand, a few participants explicitly mentioned the ease of taking photos of the shoulder blades, e.g., “right shoulder, um, this feels amusingly easy now at the end” (P3). Potential reasons for the differences are the use of a mirror and the individual’s flexibility level.

Following our study protocol, the researchers did not see any of the images taken by the participants and thus were unable to make any assessment of the image quality, e.g., if the feature of interest was blurry. However, several participants described image quality issues when reviewing their own images. Three participants (25%; P1, P7, P9) had problems with camera focus while capturing their belly, shoulder blades, hands, nostrils and buttocks. Two participants (17%; P4, P8) struggled to find appropriate lighting conditions for taking the photographs, suggesting “it would be nice to provide more information while taking the photos, for example, if the light is good” (P7).
3.4 Findings

3.4.4 Emotional Impacts of Medical Selfies

As well as practical challenges in taking the images of their own bodies, participants were also affected by emotional issues. Participants in the study were willing to take photos of most body parts, the only exception being the genitals, which three participants (25%) were not comfortable capturing (Table 3.5). Interestingly, only one of the participants did not take the buttock photos. This may have been by accident, as no reasoning was given, and the the participant did photograph the genital area, which is usually the most sensitive body part. One participant took photos of their thighs instead of the soles of the feet and one participant skipped the last two photos of shoulder blades for unknown reasons.

Survey respondents were least comfortable with photographing the buttocks and genitals (Figure 3.4). Additionally photographing the belly, chest and hips appeared to raise feelings of discomfort for some respondents. To test whether the independent variable gender (male/female) had an effect on the dependent variable photography comfort level (0-4), we conducted a MANOVA. Male and female survey respondents reported similar levels of comfort for individual body parts ($F(18,80) = 1.491, p > .05$, Wilk’s $\lambda = .749$). In general, buttocks ($M_{\text{comfort}} = 1.31, SD = 1.40$) and genitals ($M_{\text{comfort}} = .90, SD = 1.18$) were the most uncomfortable body parts to be shared via photos for respondents.

![How comfortable would you feel sharing photos of...](image)

Fig. 3.4 Survey results: Mean comfort level of male (n=57) and female (n=42) respondents to share medical selfies by body part.
3.4 Findings

3.4.5 Impacts of Taking Medical Selfies

By inspecting their bodies as part of the study procedure, participants raised their awareness of their bodily health status. Four participants noticed that their skin was dry, mostly on the feet. For example, one participant discovered new moles while checking on another mole that they had not checked in years (P4). Other features participants noted were stretch marks, sun burns and peeling callus. The process of self-examination made participants check on body parts they would not normally look at in their daily life.

Individuals were reminded of past events associated with certain body lesions: “I had surgery a few years ago. [...] I had forgotten I had this scar. And as I now saw it again, I was reminded of this surgery. A bit... It was not bad the feeling, but the surgery was present again. Not so pleasant” (P2). in some cases strong feelings in relation to past events were evoked during the capture process: “I also see a spot there. I don’t know if it’s from a rash I had long time ago. [...] Indeed, something of it remained and then I remembered everything. Yes exactly. That was very tough for me back then, because I had a great birthday party, spent a lot of money and in South Africa, where I come from, [...] and then I went to the doctor [...] and spent a lot of money” (P4).

During the process of taking the photos, four participants (33%) mentioned the aesthetic dimensions of the photos. Two participants (17%) stated they should get a pedicure done after examining their feet. Armpits caused feelings of disgust and shame, as expressed by two individuals (17%): “Armpits disgust me. How can it be this hard to talk about it? It somehow annoys me, and I do not think armpits are pretty, on anybody” (P3). A second participant stating, “so I probably have to photograph there, well now it just comes to my mind that I should have shaved my armpits before this” (P8). One participant reflected on the fact that the aesthetic character of medical photos significantly differs from sending photos to other people: “Yes, it’s stupid, I know, because a dermatologist looks at it in a different way, but somehow you always want to look as good as possible” (P4). Another participant stated that they wanted to take the photos from an angle and in the lighting where their body parts look the best (P10).
3.4.6 Data Privacy Concerns

Data privacy was an important topic for most participants (Table 3.3). A major concern, mentioned by a third of participants (33%) was that photos taken with the smartphone would automatically be uploaded to the internet via cloud services embedded in the devices. Three individuals (25%) were concerned that, as they appeared in the photo gallery, they might accidentally show the pictures to other people when showing other images. To prevent possible tracking, one participant (P11) reported that they typically remove all the metadata from photos before sharing them. One participant related a prior situation, not in the context of the study, where they sent a sensitive photo, to the wrong person: “I was actually taking a picture for my mom when I had a strange mark on my breast and I accidentally sent it to my aunt’s husband and damn, I felt ashamed. Sure, I removed it right away, but since then I have been thinking every day that, damn it, if he saw it there, it would be quite embarrassing” (P3).

3.5 Discussion

In this section we discuss our findings in relation to our research questions and prior works, present design guidelines and describe limitations of our work.

Our RQ1 is related to the emotional impacts associated with the process of taking medical selfies and how they differ between body parts. Not surprisingly, emotional impacts of medical selfies became apparent especially when photographing intimate areas of the body. Both the survey respondents and study participants similarly expressed discomfort when considering, or actually capturing, images of their buttocks and genitals. Gender differences were identified in the range of body parts that respondents had captured and sent as medical selfies, with women photographing a broader set of parts, but not in the overall frequency of sending such photos. While only a small subset of our sample had prior experience with sharing medical images with a doctor, sharing medical selfies with close friends and family is a common practice, although typically done infrequently.
Taking medical photographs of different parts of our bodies without using any filters or aiming for an aesthetic photograph, differs greatly from the current mainstream selfie-culture [153, 49, 250], potentially causing affective responses that should be considered in the development of imaging applications for medical health purposes. However, interestingly, for some participants the aesthetic dimensions of selfies were also present in the medical context of our study. Participants were concerned about their appearance in medical selfies, even though they were aware of the sole medical purpose. Similar to [163] who report on staging behavior, e.g., setting up the environment to take selfies, participants in our study put efforts into making the photos look aesthetic. We derive two implications from this finding. First, this behavior could impact image quality, e.g., when participants prioritize aesthetics over accurate representation of skin lesions. However, secondly, as posting selfies encourages patients to take ownership of their condition [258, 33], the creation of aesthetic medical selfies should not be discouraged. Hence, designers and developers of mHealth apps should not reduce medical selfies to a single functional purpose of disease diagnosis, but also accept them as a means of patients’ self expression and enable users to take aesthetically pleasing medical images.

Data privacy concerns were one of the most important reasons for participants not to take photos of their genitals. To build trust in mHealth applications, data privacy should be considered a high priority. Besides the actual implementation of standards such as end-to-end encryption, these measures should be clearly communicated to the users, e.g., by continuously displaying a disconnection state from cloud storage. As proposed by the trustworthiness model by Akter et al. [7] and consistent with Lupton and Jutel [158], disclaimers in the context of diagnoses increase user trust. To prevent accidental leakage of sensitive images, medical selfie images should be stored separately from other images and require additional confirmation when sharing them.

Individuals in the survey and, to a limited extent, in the study had a preference for familiar doctors to diagnose their medical selfies. Therefore, we suggest that digital health solutions support the patient-doctor relationship by displaying information about the diagnosing doctor, e.g., through photos and a description text. This design idea supports the concept of
alternating between telemedicine and personal patient-to-doctor interactions proposed by Scott et al. [229].

Our second research question, RQ2, explores the physical and practical challenges people face when taking medical selfies and how mHealth apps can be designed to overcome them. Participants had difficulties in taking high quality medical selfies due to uncertainty in device and body positioning, poor focus and a lack of good lighting, factors which have also been observed in prior work [13]. Providing assistance in the form of model images for the user to aim to replicate or visual/auditory guidance before and during the capture process may provide solutions to mitigate some of these challenges [62]. Solutions such as the automated shutter release of the Miiskin app[167] may be effective in ensuring image quality, but without appropriate guidance may become frustrating for the user.

In terms of physical demand, the body parts on the back side of the body, the buttocks and shoulder blades, were the hardest to capture. While it could be argued that for those body parts people can ask someone else to take the photos, this is not always possible and especially difficult for the buttocks, which are among the most sensitive body parts. Potential solutions are using two mirrors [79] or selfie-sticks [12]. Furthermore, the app should enable images to be taken without the need to manually press a shutter button.

Our findings provide insights to support application developers and HCI practitioners to take into account the role of emotional and physical experiences when designing digital health technologies for skin self-examination (Table 3.6). As a result we hope this will lead to the development of digital health technologies that encourage, rather than inhibit, people to use them, and to increase their effectiveness in disease diagnosis and treatment.

### 3.5.1 Methodological Notes

In general, user study procedures such as ours, which require participants to independently follow multiple instructional steps in their own homes, face challenges of data quality and completeness. Here we discuss the challenges in our study, with the aim of providing guidance for future researchers to minimise such issues. Firstly, there was one occurrence where a participant almost accidentally sent the study photos to the researchers (instead of
### 3.5 Discussion

<table>
<thead>
<tr>
<th>Issue</th>
<th>Design Guideline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Device positioning</td>
<td>Provide model images for the user to aim to replicate.</td>
</tr>
<tr>
<td>Device positioning</td>
<td>Provide a &quot;mirror mode&quot; to guide participants to use a mirror to take photos of the back side of their body.</td>
</tr>
<tr>
<td>Image acquisition</td>
<td>Enable images to be taken without the need to manually press a shutter button.</td>
</tr>
<tr>
<td>Appearance in Medical Selfies</td>
<td>Enable users to take aesthetically pleasing medical images.</td>
</tr>
<tr>
<td>Data Privacy Concerns</td>
<td>Continuously display disconnection state from cloud storage.</td>
</tr>
<tr>
<td>Data Privacy Concerns</td>
<td>Store the medical selfie images separately from other images and require additional confirmation when sharing them.</td>
</tr>
</tbody>
</table>

Table 3.6 Guidelines for the design of medical selfie mHealth apps derived from observations in the in-depth user study.

Only the audio and video recordings). Had this happened, it would have created privacy and ethical issues that were not covered by the ethical plan for the study, and that the researchers were unprepared for. Planning how to minimise the potential for such human errors, and having a recovery plan in place would be beneficial in future studies of a similar nature. Second, three participants dropped out of the study due to problems with the study procedure; two individuals reported that they felt overwhelmed, while one did not manage to send the recordings, likely due to usability issues with using their smartphone to record audio. Our procedure could have been improved by providing an integrated approach, i.e. a study-specific application. This would also increase data privacy of the photos since cloud upload and photo gallery ‘leaks’ can be prevented by default. Additionally, the recordings could be sent to the researchers in a safe, i.e. encrypted, way instead of through e-mails.

Overall, we regard our study protocol as suitable for the purpose of collecting data about medical selfies. Especially the audio recorded thinking aloud method led to a lot of valuable in-situ insights which would not have been captured with a retrospective interview or questionnaire. However, the think aloud method led to some participants feeling that
researchers were somehow present during the photo taking and not completely immersing in the imaginary scenario of taking photos for medical purposes, e.g., commenting, “[I] could not completely forget that it’s not for my doctor, but for a study” (P1). As the data collection process was already in a rather uncontrolled setting, a less unobtrusive experimental approach would have been difficult to achieve.

The data collected in the survey and the study originated from three distinct populations. While the survey respondents were located in the USA, study participants resided in Finland and Germany. The cultural differences between these groups were not analyzed. We wanted to achieve a wide spread of study sites across the western worlds. We had to weigh up in which geographical areas we could conduct the study. Instead of studying only one western country, we chose Finland, Germany and the USA, as they have different health care systems but all share the western culture. Since the authors had access to users in Finland and Germany, we decided to conduct the study in these countries. We noticed differences in the health care systems that impacted the user study in particular. While in Germany most patients have a nominated ‘house doctor’, in Finland the relation between a patient and an individual doctor is less close. Therefore, sending photos to a specific and familiar doctor was regarded more important for the German participants. We also noted that Finnish participants had more experience with telemedicine. Our combination of rich in-depth interviews plus a larger sample survey provides both depth and breadth to our findings, which we believe supports the generalisation of our findings.

3.5.2 Limitations and Future Work

The international nature of the study, conducted in Finland and Germany, made it necessary to translate most of the collected data from the participants’ native languages to English as a working language because not all the study’s researchers spoke both German and Finnish. According to [138], every translation is an interpretation and as such is neither objective nor neutral. Hence, the translation should be considered as part of the research process. Since we conducted the analysis on translated versions of the audio, video and interview data, some of the affective clues may have become distorted from their meaning in their original language.
3.6 Conclusion

In future, we plan to conduct further studies on the topic, for example investigating the role of discussion forums, such as Reddit, in the context of SSE.

3.6 Conclusion

We collected two datasets, a survey (n=100) and an in-depth user study (n=12), exposing user perceptions of taking selfie photographs of skin conditions for diagnosis. When taking medical selfies of intimate body regions, people felt emotional discomfort and expressed data privacy concerns. For some participants, the aesthetics of the photos played an important role, even though they recognized their medical purpose. Participants faced practical problems with positioning to photograph their back and buttocks. Issues of lighting and camera focus were noted as challenges by participants. Our findings highlight the importance of addressing practical and usability issues as well as managing privacy in the user interface of e-health apps, building trust between users and the system. Methodologically, our approach of data collection, with participants recording their thinking aloud audio while capturing the medical selfies, and later video recordings provided a good trade-off between maintaining privacy and data quality.
In the previous chapter, we highlighted the physical and emotional challenges of generating and understanding digital health data using the example of medical selfies. We suggested design implications that ensure a positive user experience, e.g., to foster the aesthetics of photos, even when taken for medical purposes. In this chapter, we transition from the *generate* stage to the next stage *share* (Figure 4.1) where we explore the socio-psychological processes and motivations to share digital health data during a pandemic. Our findings indicate that
collective motives dominated among app reviews (n=464) and in in-depth interviews (n=10). We contribute implications for future virtual citizen science tools in times of crises that highlight the importance of communication, transparency, and responsibility.

4.1 Introduction & Motivation

The ongoing COVID-19 global pandemic [182, 106] has resulted in more than 29.3 million cases and 930,000 deaths as of September 15, 2020, in more than 188 countries [118]. The virus is spread primarily via nose and mouth secretions [184], and the two most common symptoms are fever and dry cough. However, symptoms can be non-specific, and some infected people can be asymptomatic [37]. In conjunction with the virus’ ranging incubation period [183], these facts facilitate the (involuntary) transmission of the virus [38].

Thus, governments, research facilities and industry have developed a variety of digital tools, which can fulfil central functions in the mitigation and control of the COVID-19 pandemic, by addressing different important aspects in the fight against it [283]. From a research point of view, technologies can be employed to control the pandemic, within five main application areas [269]: i) early detection and diagnosis of the infection, ii) monitoring the treatment, iii) contact tracing of the individuals, iv) projection of cases and mortality, and v) development of drugs and vaccines. Therefore, various mobile applications have emerged as digital tools aiming to accomplish the aforementioned goals [54].

In particular, digital tools that are currently being utilised are covering the following areas: (i) providing instant epidemic information (Johns Hopkins University Dashboard [118, 64]; Worldometer’s coronavirus update statistics [296]; WHO’s Coronavirus Disease (COVID-19) Dashboard [186]), (ii) providing health information and prevention advice (CovApp [42]; India’s GoK Direct [199], Poland’s ProteGO Safe [73]), (iii) facilitating contact tracing (Germany’s Corona-Warn-App [212]; Australia’s COVIDSafe [287]; Singapore’s TraceTogether [288]), (iv) supporting symptom checking and feedback (Italy’s Covid Community Alert, Sri Lanka’s Self Shield [286]), (v) monitoring quarantine compliance (Poland’s Home Quarantine App [94]; Karnataka’s Quarantine Watch [263]), (vi) or citizen science projects to support research (COVID Symptom Study research mobile app [285], Germany’s Corona-Data-Donation-App [214]).

The Corona-Datenspende-App (CDA, Corona-Data-Donation-App) [214] is such a citizen science tool, published by the Robert Koch Institute (RKI; the German public health institute) in early April 2020 (see section 4.4.1 for a detailed description of the app and its
4.1 Introduction & Motivation

functionalities). It provides citizens with the opportunity to share their health data from fitness trackers and smartwatches, with the aim to better record and understand the spread of COVID-19, detecting local fever outbreaks potentially associated with COVID-19 [216]. This app is a virtual citizen science (VCS) project, meaning it comprises a crowdsourced approach conducted almost entirely through virtual portals [284]. We chose the CDA as our subject of research for multiple reasons. The CDA addresses a problem of global relevance, gathered a large user base (more than 500,000 users) within a few weeks of the release, was endorsed by the government and health authorities and was one of the earliest technological efforts in collecting physiological data for the detection of COVID-19 symptoms.

Understanding the motivations behind the participation in a VCS project that takes place under such exceptional circumstances is of great significance, especially given the current high stakes. The success of such research relies on the data collected, thus a higher number of people with continuing engagement are imperative for contributing and growing the data set. In the case of the CDA in particular, it is essential for volunteers to participate in the long-term.

In this paper, we study the motivations and investigate the socio-psychological processes which led people to engage in a VCS project and share personal data during the COVID-19 pandemic. In particular, we explore the "lived experience" of a VCS project and how it affected participation. The need to study the lived experience of such projects is also pointed out in Brown et al. [30]. To that end, we researched the motivations of VCS users of the CDA.

We collected 10,202 reviews about the CDA from Apple’s App Store and Google’s Play Store (later described in detail in section 4.4.3) and coded 464 reviews to understand the motivations of the users, deriving a model about their lived experience and engagement lifecycle. Based on our findings, in a second step, we validated the model with in-depth interviews of ten app users, in which they discussed their experience using the CDA. Our findings indicate that, contrary to other VCS projects, collective motives dominated over egoistic ones. Moreover, we observed an unprecedented persistence from the app’s users to continue using the CDA, despite various deterring factors (app instability and issues, negative
media coverage etc). These motives and reasons are explored in-depth and presented in this paper, and we contribute design implications for future VCS tools in times of crises that embed user needs and aim for long-term engagement.

4.2 Background & Related Work

In this section, we contextualise our research within past efforts. First, we review prior work about citizen science in HCI and Computer Supported Cooperative Work (CSCW). Then, we describe research on the motivational processes of social participation.

4.2.1 Citizen Science in HCI

Citizen science, also known as community science, is scientific research conducted, in whole or in part, by nonprofessionals [92]. The definition of ‘citizen science’ and the terminology surrounding the concept is in constant flux [68]. As outlined by Eitzel et al. [68], citizen science encompasses a spectrum from active to passive participation. We adopt a definition in line with this broad vision. Citizen science can be described as public participation in scientific research and can lead to advancements in scientific research, as well as in an increase of the public’s understanding of science [96, 65]. Citizen science can enable people to help solve problems that are relevant for or of interest to them [128], and most of its projects are contributory ones [222]. By participating as collaborators/volunteers in a scientific research project [112], citizen scientists are more engaged and have a more active role compared to scientific research studies involving members of the public as subjects. Moreover, by incorporating members of the public, citizen science projects enable data collection on a larger scale and over a wider geographical area [120, 119, 222]. Citizen science can be employed in a variety of disciplines, such as biology [249], environmental studies [128], and mathematics [51], with applications ranging from observation documentation [187] and data collection and analysis [304], to content creation [111] and curation [150].

In 2016, Preece [198] presented a research agenda on how HCI and citizen science can mutually enrich each other. Concurrently, within the last couple of years, HCI and
CSCW scholars explored a variety of different aspects of virtual citizen science projects. Understanding how to facilitate and maintain engagement in virtual citizen science projects is one of the main questions HCI research focuses on. Eveleigh et al. [75] emphasised the need to study different groups of citizen science collaborators, namely highly committed volunteers, occasional contributors and contributors who drop out. They found that the type of motivation (i.e. extrinsic, intrinsic) was related to contribution behaviour. They determined that there is a need to design VCS interfaces that address a variety of different motivational needs of citizen scientists. Snyder [241] explored visualisations created by and for nonprofessionals in the context of the citizen science project COASST (i.e. Washington’s Coastal Observation and Seabird Survey Team). Based on their results, they derived insights on how vernacular visualisations can enrich communication and coordination in citizen science projects. Moreover, scholars have explored the integration of gamification into VCS projects to increase user activity and support motivation. For instance, Eveleigh et al. [76] found that competitive gamification elements led to various reactions from engagement to disengagement within the citizen science project. Based on their results, they derived design implications for citizen scientists with different engagement levels. Simperl et al. [234] conducted a literature review combined with an interview study to identify the value of gamification elements for VCS projects. Their findings indicate that gamification elements appeared to have little impact on the recruitment of new citizen scientists. However, such elements seemed to be a means to maintain interest of already engaged citizen scientists. These examples indicate the need to build an understanding about the motivational processes of citizen science projects across various contexts (i.e. projects), encompassing citizen scientists with different levels of experience. Our work aims to take another angle from previous work as it looks specifically at the lived experience of a VCS project on the COVID-19 pandemic. Based on our results, we strive to derive implications to design VCS systems that embed user needs in times of personal and societal crises.

Another strain of research studied the social mechanisms of VCS projects. Jackson et al. [114] described the dynamic process of volunteer engagement. Based on the example of the project Gravity Spy, their study showed that citizen scientists initially engaged with
resources by project organisers. However, their engagement shifted to community- and agent-centred resources in later stages of the project. In contrast to most studies in VCS that focus on long-term engagement, Reeves and Simperl [210] explored short-term citizen science projects. They noted that responding to discussion messages by volunteers was particularly important at the beginning of VCS projects to maintain involvement. On a similar note, Jay et al. [116] found that contribution rates to VCS projects can be significantly increased by allowing citizen scientists to contribute without registering for a designated citizen science project. Addressing the data privacy concerns in citizen science, Bowser et al. [26] report that values and norms of citizen science explicitly promote data sharing to achieve a greater good. The examples outlined above all focus on topics that do not pose an immediate threat to the individual citizen scientist or the society as a whole (e.g., Old Weather project [75], COASST [241]). Our work aims to take another angle from previous work as it looks specifically at a virtual citizen science project that addresses an immediate crises. We strive to understand the lived experience of CDA as a VCS project to derive insights about the intricacies of designing VCS tools in times of crises.

4.2.2 Motivations for Participation in Citizen Science

Multiple studies have investigated which motives are responsible for participation in citizen science. Many report intrinsic motives, such as personal interest, as the most important factor for contribution [117, 176, 221]. However, some studies suggest that collective motives, such as contributing to science, motivate citizen scientists [201, 53]. Some studies found both intrinsic and collective motives make people participate in citizen science [177, 140]. Most studies on citizen science in HCI are based on motivational models from psychology, e.g., from Batson & Ahmad [18], who identified four types of motivations for social participation towards common goals: egoism, altruism, collectivism and principlism. **Egoism** occurs when the ultimate goal is to increase one’s own welfare. **Altruism** has the goal of increasing the welfare of another individual or group of individuals. **Collectivism** has the goal of increasing the welfare of a specific group that one belongs to. **Principlism** has the goal of upholding one or more principles dear to one’s heart (e.g., justice or equality). Batson & Ahmad’s
theory was not directed towards citizen science, but explains the mechanisms of motivation in building and sustaining a community [222].

Another model has been proposed by Klanderman, who explains voluntary participation in social movements [131]. This framework includes four classes of volunteers’ motivations for participation: collective motives (the importance attributed to the project’s objectives); norm-oriented motives (expectations regarding the reactions of important others, such as friends, family or colleagues); reward motives (benefits such as gaining reputation, or making new friends); and identification (identification with the group, and following its norms). Klanderman’s framework has been extended with an intrinsic motivation dimension, operationalised as the enjoyment associated with participation in the project. This extended framework has been applied in studies of open-source software development [101] and Wikipedia editing [227]. Furthermore, reward motives have been divided into two specific motives: community reputation benefits and social interaction benefits [35, 218, 177].

Rotman et al. [222] describe a dynamic model of motivations that change during the contribution cycle of a citizen science project. Two important points in time are the initial encounter between a volunteer and scientific projects and the point at which a project ends. At the initial encounter, volunteers are driven by egoism. At the second end of a task or a project, other motivational factors, altruism and collectivism, become more important.

Citizen Science projects are conducted in different domains, and the motivations that drive participants vary.

### 4.2.3 Crisis Informatics

A large array of past research efforts in HCI and beyond explored how people utilise technologies in crisis. Crisis informatics is an interdisciplinary research area investigating ‘the interconnectedness of people, organisations, information and technology during crises’ [93]. In recent years, crisis informatics increasingly focused on the sociotechnical intricacies connected to crises [243]. Previous work in crisis informatics inter alia explored topics such as public health crisis [89], natural disasters [44, 271] and climate change [242].
In the context of the Zika virus epidemic, Gui et al. [89] qualitatively studied the perception of risk and risk communication on Reddit. Their study provided relevant starting points for successful risk communication via social media. For instance, their results highlighted the relevance of personal experiences and the resulting perception of risks and risk measures. Besides, Gui et al. found that people also thought about the challenges facing society as a whole in times of crisis. Thus, they emphasised the need to foster dialogue between the authorities and citizens in times of crisis. The importance of another social media platform in times of crisis, namely Twitter, was explored by Vieweg et al. [271]. They analysed more than 20000 tweets connected to two natural hazards (the Red River Floods and the Oklahoma grass fires). Their work provided starting points on how social media could support affected areas in current crises. More precisely, Vieweg et al. discussed how information extracted from Twitter could increase situational awareness of the public and emergency responders. Similarly, Chauhan and Hughes [44] analysed online information on the Carlton Complex Wildfire. Their results emphasised the central role local media outlets played to distribute crisis information. They also discussed the increasing difficulty for those affected by a crisis to identify helpful information in the wealth of information available.

On another note, Huang et al. [107] explored how the emotional proximity to an anthropogenic hazard, namely the Boston Marathon Bombings in 2013, affected information sharing behaviour. Their results showed that physical and emotional proximity to the crisis impacted how people sought and shared information online. Furthermore, the authors highlighted the critical role social media can play to those emotionally and physically affected by the crisis.

The aforementioned examples indicate a shift in crisis informatics from authority-centric ‘push’ culture towards citizens who are considered to be able to do life-saving work [256]. It has been suggested that future research should conduct user-centred studies for mobile crisis applications, specifically addressing the motivations of citizens. While there are usually multiple motivations involved [18, 177], the dominating factors differ between projects.

Our work is also inspired by previous work in HCI that focuses on long-term user experience (UX). A crisis itself or the consequences of a crisis are often long-lasting. In
4.2 Background & Related Work

Fig. 4.2 Screenshots of the CDA on a smartphone in German, the only available language for the app. The following descriptions are from top to bottom in each screenshot. Left: One of five screens of the app’s initial setup asking users for their gender, birth year, height and weight. Middle: Homescreen of the CDA thanking users and showing a short explanation of the app’s purpose, three graphics explaining three important measures to mitigate the transmission of the virus (1.5m physical distance, washing hands, wearing a mask), a hint that the personal pseudonym and the wearable’s registration setting can be found in the menu and the current number of donated days, here "0". Right: Menu of the CDA displaying the user’s pseudonym and offering access to the wearables registration settings, the apps’ data privacy and imprint, current information on the results (leads to the blog on the CDA website), current information about the app and an option to delete the user account. The remaining screens of the CDA, which can be accessed through the menu, do not offer (other) interactive elements or information regarding the user’s data donation.

line with that, ideally, the CDA should be used over a more extended period of time. As our study focuses on the lived experience of the CDA, there is a need to discuss previous studies on long-term UX. Karapanos et al. [122] introduced a conceptual model that focused on the temporality of user experience. Their model encompasses three forces, familiarity, functional dependency and emotional attachment. The three forces impact how people transition across three different phases of their experiences with an interactive product. The phases are orientation, incorporation and identification. Since we focus on the lived experience of the CDA, the third phase identification and the associated force emotional
attachment is of particular relevance in our inquiry. The identification phase focuses on how an interactive product is integrated into ones’ everyday life and concurrent emotional experiences. Consequently, we consider emotions as a key variable in our study. In the field of personal informatics, Epstein et al. [70], introduced the lived informatics model of personal informatics. Also, Epstein et al. [69] emphasised the importance of emotions during and after the personal informatics experience, such as feeling guilty for stopping to track. Even though using the CDA in a pandemic is not the same as a lived informatics experience described by Epstein et al. [70], there are some similarities, such as collecting and sharing personal data. Thus, in our study, we also focus on the entire lived experience of the CDA. To summarise, while citizen science projects in crises have been described in the crisis informatics literature, to the best of our knowledge, no prior work in HCI addressed emotions and motivations surrounding citizen science projects in times of crises. Hence, our work explores the lived experience of the CDA with a focus on emotions and motivations.

4.3 Research Questions

Our overarching goal is to learn more about the lived experience of the Corona-Data-Donation App. We endeavour to understand the motivations of people contributing to citizen science projects that focus on incidents posing an immediate threat to the individual citizen scientist or the society as a whole. We explore this topic by focusing on the CDA, a virtual citizen science project that aims to mitigate the impact of the COVID-19 pandemic. To that end, we strive to answer the following research questions:

- **RQ1**: What are the motivations to participate in a VCS project in times of crises?
- **RQ2**: What are the socio-psychological processes associated with the participation in a VCS project in times of crises?
4.4 Method

This study explores the lived experience of the CDA. To that end, we applied a two-step procedure. We retrieved and qualitatively analysed users’ online reviews of the CDA from the App Store and the Play Store to derive a holistic understanding about the experience with the CDA. As a second step, we conducted semi-structured interviews with ten individuals to explore themes identified in the reviews in greater detail. The reviews helped us to identify the range of different motivations present among a large group of users, however, we anticipated the reviews to be of a rather strong opinion. The additional analysis of the interviews allowed us to ask differentiated questions on specific motivations that contributed to a more nuanced picture of the users’ reasoning for participation. Additionally, as one of our research goals targets the socio-psychological processes of citizen scientists, the complementary analysis gave us two perspectives on social interaction among users. First, reviewers can interact with other reviewers and the developers in the app stores. Second, we asked our interview partners about the exchanges they had with family and friends about the CDA. Our understanding of the lived experience of the CDA is based on a qualitative analysis of these two data collections.

4.4.1 Corona-Datenspende-App

The CDA [214] (Corona-Datenspende-App) app was released on April 7th 2020 by the RKI which has worked together with the German government to coordinate Germany’s response against COVID-19. The general aim of the CDA app is to monitor the spread of COVID-19 and analyse whether measures to contain the novel coronavirus pandemic are working [34]. The app gathers vital signs from volunteers wearing smartwatches or fitness trackers - including heart rate, daily steps and sleep habits - to analyse whether they are symptomatic of the flu-like illness. They can indicate fever, e.g., when the resting heart rate goes up and the daily steps go down. However, these physiological attributes do not necessarily mean that the user has developed a fever; at the same time, multiple causes for a fever exist, so the data has to be supplemented by other sources, for example COVID-19 case
4.4 Method

numbers. So far, there is limited evidence that passive physiological data alone support the
detection of COVID-19 [200]. Preliminary results of the CDA are reported in a blog and in
an interactive fever curve on the CDA website with the aim to help the health authorities and
the general public to assess the prevalence of infections down to state level, in combination
with other data inputs [211].

As of September 15 2020, 526,727 people downloaded the app and registered a wearable
device [215]. During the CDA’s initial setup process, users are asked to provide their
information on postal code, gender, birth year, height and weight (see Figure 4.2 (left))
and to register their wearable. After completing these initial steps, the physiological data
measured by the wearable is automatically transmitted to the RKI from the user account
at the wearable’s provider [36]. Thus, the data donation begins (see Figure 4.2 (middle))
and no user interaction with the app is required after this. The app’s menu, however, offers
access to further information on the user’s personal pseudonym, data privacy, imprint, current
information related to the research results (leads to the blog on the CDA website) and current
information about the app (see Figure 4.2 (right)). Users can also change or disconnect their
wearable or delete their account.

The CDA was published by the RKI which is a federal institute under the authority of the
Federal Ministry of Health of Germany. The call to use the app was backed up by the German
government [77] and multiple news channels reported on the day of release about the CDA
on national television [255, 303], their online platforms [254, 302] and continued to report,
e.g., about the reached user number of half a million after one month [253]. Two weeks after
the app’s release the Chaos Computer Club (CCC) analysed the CDA and publicly criticised
parts of its data privacy practice [36]. For example, the fitness data was only pseudonymised
after being uploaded to the RKI servers. As a response, the RKI started to immediately
address these issues in the implementation to close the related safety gaps. Media coverage
had a visible impact on download rates as well as reviews provided (see Figure 4.3). This
paper focuses on the CDA, which was released before and should not be confused with
the Corona-Warn-App [212], a contact-tracing application launched by the RKI, utilising
Bluetooth and location tracking.
4.4 Method

Fig. 4.3 COVID-19 cases in Germany and downloads of the CDA. The CDA was released early in the pandemic, when cases were rising rapidly. Two small spikes can be seen after the RKI gave a press briefing and when the German public broadcasting news show Tagesschau reported about German Corona Apps and showed an image of the CDA on national television. On that day, Tagesschau has actually only reported about Corona-Warn-App but showed an image of the CDA with the name clearly visible. Sources: mdr.de [164] and rki.de [215].

4.4.2 Data Collection & Analysis - CDA App Reviews

Our original data set consisted of 10,202 reviews from the App Store and Play Store. The reviews were submitted in the period between April 7 (April 9 for the App Store), when the apps were published, and April 26, which was the last day with more than 100 reviews on both platforms. For the Play Store, we scraped reviews with the python package google_play_scraper. For the App Store we accessed the RSS feed which only contained the most recent 500 reviews. The reviews information included a unique review ID, submission date, title (App Store only), review text, vote sum, vote count, score, version, author, reply content and reply date. All reviews were imported into the MAXQDA data analysis software.

We applied open coding combined with thematic analysis similar to Blandford et al. [23]. Previous work in the field of HCI has utilised app reviews in multiple ways, including to characterise attitudes of users and functionality of apps [246, 86]. Extracting user experience information from online reviews has been, inter alia, done for software and video games [98]. Similarly, in this work, we retrieve the experience of users of the CDA based on app reviews. According to Vaast [268], it is important for researchers to access the context of the texts they are analysing. Some online texts, e.g., tweets, are very short (140 characters or less),...
thus making sense of them in and of itself is difficult. Therefore, tweets are to be understood within an ensemble [27], e.g., within surrounding tweets.

Accordingly, we analysed the data in the context of the pandemic, meaning that we were aware that the motivations of the users differed from previous encounters with similar apps that collected physiological data for no direct personal benefit.

In total, we coded 464 reviews from our dataset of 10,202 reviews. Three coders coded the data independently. We selected and coded the 464 reviews in four iterations, resulting in four samples, and updated our coding tree after each iteration until theoretical saturation was reached [23]. Sample #1 consisted of 84 reviews, which were randomly sampled from our dataset. This was followed by iterative discussions to establish the initial coding tree. For samples #2 (108 reviews), #3 (136 reviews) and #4 (136 reviews) we followed the same procedure: For each app store (App Store and Play Store), we selected seven versions of the CDA, all ratings (5) and sampled two reviews for each rating. This gives us a maximum of 140 reviews for each sample. Since not all versions had at least two reviews for each rating, e.g., version 1.0.7 in the App Store had only 22 reviews, none of which were 2 or 4-star ratings, the actual amount of reviews differs (e.g., just 108 in sample 2). Some versions had less than two reviews for a rating, e.g., version 1.0.7 in the App Store had only 22 reviews, none of which were 2 or 4-star ratings. Therefore, the actual sample sizes vary and are below 140.

### 4.4.3 Data Collection & Analysis - Interviews CDA Users

To complement our review data, we recruited 10 users of the CDA app through the Prolific platform and various Facebook groups. Participants’ age range was 20-39 ($M = 26.4, SD = 6.22$), with six female and four male participants. The majority of the participants (8/10) were students, including three PhD students, three psychology students and one medical student. Other occupations included research assistant and engineer. The participants were compensated with 8€, either through an Amazon voucher or directly through Prolific. Table 4.1 presents details about the participants.
4.5 Results

<table>
<thead>
<tr>
<th>Participant</th>
<th>Age</th>
<th>Gender</th>
<th>Profession</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>27</td>
<td>female</td>
<td>PhD student</td>
</tr>
<tr>
<td>P2</td>
<td>22</td>
<td>female</td>
<td>student (medicine)</td>
</tr>
<tr>
<td>P3</td>
<td>22</td>
<td>female</td>
<td>student (psychology)</td>
</tr>
<tr>
<td>P4</td>
<td>29</td>
<td>male</td>
<td>PhD student</td>
</tr>
<tr>
<td>P5</td>
<td>23</td>
<td>female</td>
<td>student</td>
</tr>
<tr>
<td>P6</td>
<td>39</td>
<td>male</td>
<td>PhD student</td>
</tr>
<tr>
<td>P7</td>
<td>24</td>
<td>female</td>
<td>student (psychology)</td>
</tr>
<tr>
<td>P8</td>
<td>23</td>
<td>female</td>
<td>student</td>
</tr>
<tr>
<td>P9</td>
<td>20</td>
<td>male</td>
<td>student</td>
</tr>
<tr>
<td>P10</td>
<td>35</td>
<td>male</td>
<td>engineer</td>
</tr>
</tbody>
</table>

Table 4.1 Demographics of interview participants.

All interviews were conducted via Zoom or via telephone using audio-only recording. Participants were asked for consent for recording before the interviews. The interviews lasted an average of 15 minutes (min = 08:40, max = 22:05, sum = 02:15:38). Using a semi-structured approach, we covered the topics of motivation, experience (e.g., installation issues and frequency of use) and data (e.g., data privacy, awareness of types of data shared). Most questions were open-ended, e.g., “please describe your experience from downloading the app until today”. We provide the interview guideline in Appendix A. All interview recordings were transcribed verbatim and imported into the MAXQDA data analysis software.

4.5 Results

In this section, we report on our findings from the analysis of the reviews and the interviews. First, we present the descriptive analysis of the reviews. Second, we explain the experience of the CDA citizen scientists and the three themes we identified in the reviews and interviews, i.e. motivations, emotions and persistence. We provide the coding tree in Appendix B.
4.5 Results

The experience of citizen scientists using the CDA during the COVID-19 pandemic is shown in Figure 4.4. Individual aspects of this cycle of engagement demonstrated with the CDA such as the motivation and persistence are highlighted in the following subsections. Direct quotes were translated from German.

4.5.1 Reviews

Our analysis included 464 reviews in total in equal proportions from the App Store and Play Store. The average number of words per review was 41.8 ($SD = 29.7$) in the App Store (title and review combined) and 22.0 ($SD = 17.4$) in the Play Store (review only). Most of the reviews were submitted within the first days after publication. For example, the number of daily published reviews in the Play Store peaked around 3,000 on the first day – almost a third of all collected reviews (Figure 4.5). Afterwards, the submission rate dropped rapidly.

The mean ratings, illustrating users’ quantified feedback between 1 and 5, were generally slightly higher in the App Store compared to the Play Store (Figures 4.6). Reviews were mostly rather negative or rather positive (Figure 4.7). Ratings went up after the release of the

Fig. 4.4 The experience of using the CDA, a virtual citizen science project, during the COVID-19 pandemic. People’s entry point is the initial motivation leading to the decision of using the app, i.e., participating in the VCS project. After the download and installation, users are confronted with the setup phase and exposed to positive and negative factors of the app. Consequently, positive and negative emotions arise causing a re-assessment of the motivation and re-assessment of the app. Based on users’ decisions, what follows is either a drop-out, i.e. removal of the app or cancellation of the data donation, or continued usage of the app. Particularly in the case of experienced negative emotions and the decision to continue donating data, users demonstrated persistence in remaining a citizen scientist.
4.5 Results

Fig. 4.5 Number of daily reviews submitted for the CDA between April 7 and April 26 (Google Play Store), and April 9 and April 26 (Apple App Store). More than 80% of the reviews were submitted in the first seven days after publication of the CDA.

Fig. 4.6 Mean daily ratings of reviews submitted for the CDA between April 7 and April 26 (Play Store), and April 9 and April 26 (App Store). Annotated are the version numbers. Version 1.0.4 fixed a common error with the day counter as can be seen by the increase in the scores. The App Store ratings were generally slightly higher compared to the Play Store.

version 1.0.4 on April 12 which fixed an error regarding a broken counter for the number of donated days. A slight downward trend followed afterwards.
4.5 Results

<table>
<thead>
<tr>
<th>App Store</th>
<th>1 star</th>
<th>2 star</th>
<th>3 star</th>
<th>4 star</th>
<th>5 star</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>414</td>
<td>137</td>
<td>182</td>
<td>112</td>
<td>268</td>
</tr>
<tr>
<td></td>
<td>37.2%</td>
<td>12.3%</td>
<td>16.4%</td>
<td>10.1%</td>
<td>24.1%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Play Store</th>
<th>1 star</th>
<th>2 star</th>
<th>3 star</th>
<th>4 star</th>
<th>5 star</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4636</td>
<td>1358</td>
<td>1260</td>
<td>600</td>
<td>1235</td>
</tr>
<tr>
<td></td>
<td>51%</td>
<td>14.9%</td>
<td>13.9%</td>
<td>6.6%</td>
<td>13.6%</td>
</tr>
</tbody>
</table>

Fig. 4.7 Distribution of ratings for both stores. Half of the Play Store ratings were 1-star, whereas only a third of the App Store reviews were 1-star.

4.5.2 Initial Motivation

Initial motivations for participation were collective and mainly revolved around the opportunity to serve the community in an unprecedented situation and the role of the RKI as a trusted organisation. These collective motivations were expressed through a variety of perspectives, such as an explicit desire to support the project due to the pandemic and its societal relevance. Many users expressed that they were "happy to support" the research of the RKI to better understand COVID-19 in such an unprecedented difficult situation. One common view amongst users was that the app enabled individuals to contribute to the fight of stopping the spread of the COVID-19 pandemic. The motivation of some users to serve their community overcame their concerns about sharing privacy-sensitive data. While the following quote also includes the potential personal benefits of donating data, i.e. an egoistic motive, it highlights the main motivation is still to serve the community.

*I had never thought I would share such intimate details, but everything that now helps statisticians to understand [the shared data], protects my family members at risk and friends. Eventually, also me! (App Store, 1.0.4, 5 stars)*

Another factor was the hope for the intended success of the app which could lead to lifting social distancing measures in the long run. This was valued higher by one user than the data privacy, which was considered to be a "luxury product" in this unfamiliar situation. Other reasons for the willingness to share personal data were the pseudonymisation of the data, the trust in the RKI and general concerns about the community’s health. As mentioned frequently in the interviews, even users who did not usually share their physiological data
with other services participated in the donation (P2, P7, P8, P9), emphasizing the trust in the RKI, both in terms of data privacy and hopes for mitigating the pandemic:

*In contrast to giving the data to Google, they go to the government. And I think they handle the data significantly better compared to when the data are sold. And this I considered a small opportunity to help.* (P9)

The commitment to serve for the “greater good” was also accompanied by requests to share more data than the app asked for, e.g., the fitness data collected by their devices before the app was downloaded. In addition, it was requested to manually insert information about temperature, pregnancy status, allergies, general health status or symptoms. P8 expresses readiness to share sensitive data:

*Yes, whatever they had asked for. Everything. In my head there are symptoms, like how I feel, if I have a high temperature. Maybe also normal things, such as how often do you go outside, so you can maybe detect preventively and protect [yourself] from Corona, not only [detecting] symptoms.*” (P8)

One user even reported purchasing a wearable for the sole purpose of donating data to the CDA.

Often, users projected their motivation to support the project to other users. This became apparent, when they peer-pressured others in their reviews to help the RKI and accept certain errors or to also donate their data.

*Let’s be reasonable and donate data!* (App Store, version 1.0.6, 5 stars)

*Help the RKI and participate! Accept bugs in the short-term.* (Play Store, version 1.0.6, 5 stars)

Responses from the interviews provided a deeper insight into certain motivational factors. Consistent with findings from the reviews, interview participants indicated that collective motives dominated the decision to contribute. Participants were aware that they had no direct benefit from using the app, but could potentially support the community. While the pandemic
was considered serious by all interviewees, concerns mainly focused on other community or family members who were at risk. Two interview participants, P1 and P9, downloaded the CDA while searching the Corona-Warn-App since it appeared in the search results and was also launched by RKI. This indicates that their initial motivation might have been egoistic or collective, since the Corona-Warn-App warns of contact with infected people, which benefits both the individual and society. In this case it is difficult to distinguish between egoistic and collective motives.

Overall, a typical behaviour observed in the reviews consisted of a general appreciation of the app’s purpose and, consequently, a strong initial motivation to contribute.

*The idea is great, but unfortunately no way to contribute with my Samsung Gear Sport. (Play Store, version 1.0.1, 3 stars)*

However, this was often followed by frustration due to errors or the incompatibility of many devices excluding users from donating their data.

### 4.5.3 Emotional Responses to Problems

Having charted the motivations of citizen scientists at the start of their contribution, we now present the emotional responses when being exposed to errors in the app. The experience of sharing data through the CDA was accompanied by positive and negative emotions. A common perception was the wish to help and the consequent disappointment when facing errors that prevented the contribution, mainly caused by a lack of feedback in the app and a limited range of supported devices. A recurrent theme in the reviews was a positive attitude towards supporting the fight against COVID-19. A few users expressed their gratitude for the "hard work to manage this crisis". Users were also thankful for updates that resolved issues with the app. We will explain the positive emotions in more detail in the next section.

It became obvious in the reviews that the app was released with some errors that prevented or complicated the intended use for a large proportion of the reviewers. This circumstance caused different negative emotions such as frustration, anger or sadness. In particular, many reviews addressed the displayed counter of donated days that was stuck at zero days. Users
were, therefore, uncertain and sceptical if their data donation was successful, i.e. if they were actually helping or if an error in the app’s development hindered their wish to contribute. This unawareness and lack of feedback about the state of the data donation caused frustration, and even when the day counter became active, some users were not entirely convinced about the correct functioning of the system:

*After the update, I can now finally see the days. Thanks a lot. However, information about which data are sent and what their purpose is is still missing.* (App Store, version 1.0.4, 3 stars)

Another dominant point of criticism leading to frustration was the limited range of supported fitness trackers and smartwatches. Many users were unable to contribute due to their own unsupported device and communicated their frustration about this issue understanding, especially after a month of waiting for an update. In some cases, frustrated users reported that this only became clear to them at the end of the registration process. Users were also frustrated that the CDA was only usable in combination with a fitness tracker or smartwatch, excluding many people who were not able to afford these devices.

While technical difficulties were the predominant reason for negative emotional responses towards the app, we also observed concerns about transparency. In particular, most users were left unaware about the exact kind of data being collected and how it would be used, as the RKI reported the generic types of data the app requested, e.g., daily steps and heart rate, on the CDA website but not in the app. Additionally, there was no option to see what data an individual had shared with the RKI.

Overall, users often had hoped to contribute, but were excluded due to reasons out of their control. A common expression in the reviews was "I would like to help, but" which was followed by the explained barriers. This led to sadness, frustration and anger, expressed in disappointment and threats to stop the data donation.

In response to the experienced errors, some users also expressed anger about a lack of professionalism in the app development and, based on that, doubted that their data was handled securely enough. Other negative comments related to the app drawing significant battery or appeared in the context of misconceptions about the CDA. For example, one user
was frustrated that the CDA would not alert users in case an infection was suspected based on the data they shared.

But, what annoys me enormously are the explanations on the RKI website\(^1\): "The Corona data donation app is not a Corona test. The users themselves are not informed about a possible infection." No one in their right mind will assume that a corona test is possible with the data from a "fitness tracker". It would be nice if the app would inform users about "anomalies" (heart rate too high in conjunction with short deep sleep phases). Theoretically, I can deduce that myself, but this way the app would provide a benefit for the user. (App Store, version 1.0.4, 2 stars).

### 4.5.4 Persistence

Our analysis showed that, after using the app, users re-assessed their motivations and attitudes towards the app (see Figure 4.4). Initially, the shortcomings of the app led to frustration and drop-outs for many users. However, when contributors reassessed their motivations, they reminded themselves of their initial motivation of supporting the collective, and applied strategies to overcome the barriers. This persistence became especially apparent when contributors updated their reviews - often indicated through "update" or "addition" or "update + revision date" at the end of the review - based on positive enforcement of their motivation, e.g., after an app update.

When re-assessing the app, common positive perceptions addressed the improvements app updates added, while negative perceptions were expressed through the errors experienced during the use. When re-assessing their motivations after initial use, participants positively highlighted the contribution of the CDA in the fight against the pandemic, while other users expressed doubts about the effectiveness of the data donation in regards to the pandemic mitigation at the same time.

\(^1\)The information this user refers to is found on the CDA website, which is provided by the RKI, but not on the RKI website.
We observed a strong sense of persistence, which was particularly noticeable in the context of negative emotions when facing errors. In many cases, users did not remove the CDA from their smartphone but rather explained their particular errors and requested the RKI to fix them. Users also often waited for multiple updates and resumed donating their data.

Another common approach of dealing with errors was through peer support. Users discovered workarounds and shared their solutions in the reviews or informed their fellow data donors in case a particular update fixed an issue, taking the role of community representatives. To do so, some returned to their reviews and, e.g., reported that it took a few days for them until the day counter started to work. In other cases, users suggested to check the CDA website, posted quotes from it or shared replies from the app support.

*I think the purpose of the app is great and I trust the data privacy. A clear 'Go' from me. From the RKI website²: "Currently, in the app the counter for days donated is not always up to date. Due to the high load of the servers, the data retrieval is delayed.". (App Store, version 1.0.2, 5 stars)*

Workarounds were also reported by users that initially were not able to contribute due to an unsupported device. For example, Huawei smartwatches were not directly supported in the first versions of the app. However, users explained that when connecting their smartwatch from an unsupported manufacturer to the Google Fit App they were able to donate. Since the Google Fit App can solely run on a smartphone without a connected wearable, it also allowed users not owning a fitness tracker or smartwatch wearable to connect to the CDA and still donate their data. Reviewers also recommended to always ensure a charged device since the app would draw significant battery. A strategy to overcome the lack of feedback given by the app was to check the day counter as an indicator to whether the donation worked:

*Especially in the beginning, I frequently checked the app to make sure, does it count the days, hoping that would mean the data for the day were transmitted (P6).*
Users’ behaviours also demonstrated a lot of patience, especially regarding errors with the day counter. Although users were uncertain if their data was even donated, they continued using the CDA, hoping the issue would be resolved in future updates. Some users reported that the set up phase and registration took quite some time and even multiple restarts did not lead to a drop-out. Furthermore, issues with the app were also seen in the context of the COVID-19 pandemic. The exceptional circumstances were considered and an understanding was expressed for issues during the early phase after publication. Fellow data donors were even asked to show more compassion due to the unprecedented times. This was expressed through statements like "Take it easy and drink some tea". Nonetheless, some users threatened to uninstall the CDA and stop sharing their data if their reported errors were not resolved in the near future. This was particularly often expressed when users were uncertain about their donated data being received. However, a "second chance" was often granted in combination with a set countdown of a couple of days:

Unfortunately, it turns out only at the very end [of the registration process] that it can not be connected with Samsung Health. But maybe this changes in the next few days. So, it will not be uninstalled yet. (Play store, version 1.0.1, 2 stars)

The strong commitment to the project was also illustrated by users’ patience in regards to the incompatibility and anticipation for their device to be supported soon. On that matter, users wished for a larger range of supported devices, since they hoped for a considerable increase in the number of data donations and, consequently, better chances for the project’s success.

The persistence of CDA users, observed particularly in these updated reviews and the related investment of time to get the app running properly in order to donate data, is particularly remarkable given the expectation of immediate availability in the context of digital technology, where long loading times of websites can already lead to dropouts.
4.6 Discussion

In this research, we explored the lived experience of the CDA, a virtual citizen science projects to mitigate the negative effects of a global pandemic. Our results contribute insights about motivational processes of citizen scientists in times of crises. We observed that users of the CDA showed collective motivation to donate their data to support the community. The crisis experience led to participants staying persistent using the CDA, even so they were dealing with technical problems and waiting for updates. While we observed frustration of the citizen scientist when faced with problems, common conceptions were positive sentiments towards the idea of data donation.

4.6.1 Motivations in Times of Crises

In relation to our RQ1, which refers to the motivations for participation in citizen science in times of crises, we found that collective motives were the dominant factor to contribute. This is consistent with Raddick et al. and Curtis [201, 53]. Contrary to Rotman et al. [222], who propose that motivations are initially egoistic and change at pivotal points in the interaction with VCS, we did not observe that users’ main motives were egoistic throughout the contribution cycle. However, we hypothesise that this can partly be explained by the limited interaction possibilities provided by the CDA. In other words, due to limited interactions with the VCS app and other citizen scientists, it was less likely for the motivations to change during the cycle of engagement with the project. Our results extend the findings from Batson et al [18] and Rotman et al. [222]. Based on the work from Batson et al. [18] there is usually a mix of motives present when citizens engage in communities and these motives, according to [222], change during the participation.

However, our results showcased that this does not necessarily seem to be the case for VCS projects in times of crises. We hypothesise that the initial motivation of citizen scientist in such extraordinary circumstances might be strong enough to remain stable even if confronted with negative experiences such as a lack of interaction with the scientists and frustrating issues with the VCS app.
In regards to RQ2, the socio-psychological processes in VCS projects in times of crises, we found that peer support was common when dealing with problems of the app. Through providing workarounds and relaying information from the CDA website, some users indicated their commitment to the project. Besides the communication among volunteers, in contrast to other citizen science projects, there was very little to no communication between scientists and users. This is critical for VCS [210] and for the field of crisis informatics, for which our model of citizens’ motivations towards such apps has important implications. In line with Tan et al. [256], who postulate a more citizen-centric communication in crises, we argue that serving the collective motives of participants will increase the understanding of the purpose of the project and the nature of their contributions. Some of the problems and questions raised in the reviews were answered by other users or covered in the FAQs on the CDA website. We assume that more effective communication could have reduced some of the frustration expressed in the reviews. Our results highlight the potential of supporting motivated participants to support their peers in using technologies such as the CDA. A practical implication could be to empower CDA users to support each other, which is motivated by our observation that peers adapted communication channels to communicate with each other that were not actually designed for such interactions. More precisely, citizens communicated about the intricacies of dealing with the CDA via online reviews in the Apple App Store and Google Play Store.

Factors deemed important for engagement in VCS, such as incorporating different motivational needs [75], visualisations [241], gamification [76], and responding to discussion messages [210] seemed to be secondary for engagement in the CDA. Still, more than 500,000 people downloaded the app and connected their wearable device to donate data. One explanation for the high participation could be that the CDA app was promoted by the German government through public broadcasting television and various press reports.
4.6.2 Design Implications

Based on our findings, we highlight possible ways forward for designing VCS projects in crises contexts, as well as starting points that can help guiding the design of future systems in the area of crisis informatics.

Communicating Societal Relevance

First, we address the importance of communicating societal relevance. The pandemic has already had drastic consequences on many societies worldwide [118]. Official authorities, such as the RKI, the German government and news outlets highlighted the role of the app in fighting the pandemic. We hypothesise that communicating the societal relevance of the CDA had an important effect on the rapid adoption rate of the app. From these observations we suggest that VCS projects in crises situations should be endorsed by official authorities and highlight their societal impact and communicate it clearly to the public. This resonates with findings from previous work. For instance, Gui et al [89] showed that people have a tendency of thinking beyond their individual needs in times of crisis. This effect, combined with the efforts of the authorities to communicate the societal relevance of the CDA may have supported the persistence of users.

Cultivating Trust and Transparency

Besides technical difficulties, concerns about data transparency were the main reason for users to stop using the CDA and contributing data. The high-level goal of the CDA, to reduce the spread of COVID-19 with health data from wearable devices, was clear to most users. However, the specifics of how their health data helped the RKI, e.g., to detect local fever outbreaks, remained unknown for many. In line with Huang et al. [107] we found that citizens were more likely to share data the closer they were to a crisis physically and emotionally. However, the lack of understanding how their data was useful raised concerns about the value of their participation. Consequently, we argue that in VCS projects that involve passive sharing of sensitive data, users need fine-grained information about the value
of their contribution. Thus, it is not enough to provide information on a high-level but on an individual level instead, e.g., through daily status reports.

The authority of the RKI as a federal German agency and the support from official authorities such as the German government led to a strong identification of the users with the goals of the CDA. Our results showed that ensuring data transparency and partnering with trusted institutions is essential to foster trust and long-term engagement of citizen scientists. Furthermore, contributors shared sensitive data with the RKI which they would not share with other platforms. While increased participation is essential for the success of VCS, developers acting on behalf of trusted institutions should be aware of their responsibility for the users who trust the institutions represented by the developers. Bowser et al. [26] refer to this problematic priming on openness in the context of citizen science with suggesting that volunteers may not raise privacy concerns on their own.

**Fostering Community Support**

As people are often overwhelmed by the range of information available, especially during crises [44], it is not enough to simply provide information. However, often, in crises, time and human resources to distribute important information are scarce. We found that many users were unaware of information that was available in the FAQ on the CDA website, for example that a non-functional day counter did not mean they were not donating data. Other topics frequently addressed in the reviews, e.g., reasons for excluding devices or data privacy policies, led to concerns about the integrity of the project. Identifying and clarifying such misconceptions could have prevented drop-outs. One way forward could be the launch of a designated platform to foster peer support such as a CDA forum directly integrated in the app. Another option could be the presentation of Twitter hashtags together with a VCS project to support community-based communication via social media [245, 271]. Previous work has shown that digital volunteers in crisis situations use Twitter to gather and disperse information and that such information is relevant, accurate and useful [271]. Online platforms such as Twitter and other internet forums support disaster-related citizen participation, enabling individual capacities and collective action [245]. Consistent with these observations, we
propose that such platforms could engage volunteers who act as community representatives for the scientists to mitigate some of the communication issues we observed in the current project, for example. As a positive side effect, these volunteers take ownership of the project, possibly synthesising user feedback for the developers.

**Personal Crisis Informatics**

Previous studies in crisis informatics primarily focused on analysing how different social media platforms, e.g., Reddit or Twitter, were used in times of crises, e.g. [89]. However, there is unexplored potential at the intersection of VCS and crisis informatics. For instance, the design and analysis of the CDA can inspire novel crisis informatics systems beyond social media. In other words, systems similar to the CDA could be an interesting extension of previous crisis informatics approaches.

The CDA is unlike more traditional VCS projects. It collects crisis-relevant, sensitive, personal data using personal informatics technologies which are normally used for private purposes, e.g., fitness trackers. Interestingly, we observed that citizens were willing to share not only their sensitive data (e.g., resting heart rate) requested by the CDA, but also broader metrics not requested by the system, e.g., body temperature.

Our analysis showed that the participants exhibited a high level of willingness to share sensitive data for a social purpose. Fostering such willingness appears to be a key design goal for technologies which use data at the time of crisis. Citizens communicated that they would have liked to actively log additional data such as their daily temperature to support the RKI and they did not understand why it was not possible to do so. Furthermore, citizens voiced their frustration with the CDA supporting only specific wearables and, which rendered them unable to donate their data. Future systems in the area of crisis informatics could take advantage of this readiness to help and provide a broader set of systems and platforms for people with different levels of openness towards technology and technology self-efficacy. This suggests that, in future public health crises, using existing health apps for data donations together with apps designed specifically for the crisis (similar to the CDA) can render the
best results. In addition, a simple website could be designed on which people who do not own a smartphone or wearable could donate their data through manual input.

4.6.3 Generalisability

Users considered the CDA as an opportunity to serve for the greater good and this perception was further reinforced by media coverage and the RKI’s role as a government organisation. We argue that our findings can be generalised to projects that are deployed within comparable circumstances, e.g., official authorities fighting large scale public health crises. For example, authorities could launch similar services to monitor outbreaks of dengue fever, measles or influenza. For the control of dengue fever, it is essential to reduce mosquito populations, as there are no effective vaccines [66]. The recommended management strategy is to eliminate unnecessary container habitats that collect water. Another public health hazard are measles, a highly contagious, serious disease caused by a virus [181]. Even though there is an effective vaccine, in 2018, there were more than 140,000 measles deaths globally, especially in developing countries. For both diseases, surveillance of fever curves in conjunction with manual reporting of infections might contribute to the localisation of outbreaks and help authorities in narrowing down their efforts to regions with high prevalence. Given the technical difficulties that accompanied the launch of the CDA, one approach for health authorities in future epidemics could be to officially endorse existing providers, such as the Fitbit platform. Users could opt-in to make their data available for the purpose of detecting symptoms based on their fitness tracker data, which has been shown to work to predict influenza-like illnesses [202]. With this method, an established technical infrastructure could be adopted by a trusted organisation to ensure reliable measurements in large populations.

In the context of a pandemic, we assume that the high perceived relevance of the projects addresses users’ collective motivations and makes them accept a range of adversities, which we will outline in the following. First, in contrast to many crisis informatics projects, where participants share observations, relay information or map hazards, the CDA requires users to contribute sensitive personal data, such as heart rate and activity data. Second, not only were there no direct rewards or a gratification system in place as a compensation for
using the CDA, nor did the CDA provide clear information on how exactly the data were useful in mitigating the effects of the pandemic. Lastly, participants shared data continuously with no defined end date, which is different to episodic crises, where citizen scientists often provide only a few data points, e.g., extent of damage at a certain location over a limited amount of time.

Therefore, the CDA shows that the societal relevance of the crisis might overrule common inhibitors for participation in VCS and crisis informatics projects. Our findings suggest that users in different types of crises with lower perceived relevance might have stopped engaging with the project at earlier stages already. Users potentially accept adversities should they have the feeling they can contribute to the greater good. In this case, developers could address users’ collective motivations and express their appreciation for serving the community. For example, the app could acknowledge the donation and send participants thank you’ notifications or let them share a message about their contributions on social media. However, we want to highlight that more research is needed to determine what drives citizen scientists’ participation during different public health crises.

4.6.4 Limitations

While we aimed to make a contribution towards understanding the lived experience of a virtual citizen science project during the COVID-19 pandemic, we can identify some limitations in our work. To this day, there is, to the best of our knowledge, no evidence that the approach of using physiological data from fitness trackers to predict COVID-19 hot-spots is effective. Moreover, most efforts to utilise mobile devices for the mitigation of the pandemic focus on contact tracing apps. These apps do not require to share personal data with the providers. Therefore, the relevance of our work for the current pandemic might be limited and is more significant for future VCS projects in crisis contexts that rely on personal data.

Our results are largely based on the analysis of reviews from the Apple App Store and Google Play Store. We are aware that most app store reviews are rather extreme [105], being very positive or very negative. Additionally, future work needs to determine if our interview
sample reflects the population of CDA users in terms of its demographics. However, our case-study presents an in-depth account of attitudes and behaviours in a nuanced manner, as it distinguishes between positive and negative emotional responses to the CDA, e.g., drop-outs or continued contribution. For a more generalisable understanding of the identified motivations, for instance in similar public health crises, future work could juxtapose the results of a more large-scale study with our findings.

The CDA was developed in a short time frame due to the dynamic of the pandemic and the urgent need for measures to detect infections and isolate potential carriers of the virus as fast as possible. Due to the time constraints, the app was released with several severe limitations, mainly the non-functional day counter, a limited range of supported devices and a lack of information on how the donated data was used by the RKI. These shortcomings could potentially have been prevented through a user-centred design process and systematic user testing. For future crisis VCS or crisis informatics projects that face the same constraints, we suggest that even when there is not enough time for user involvement in the initial releases, developers provide and maintain social media channels to synthesise feedback for future iterations and give users an opportunity to voice their opinion. This is consistent with the recommendation to utilise social media for communication among participants and between participants and developers. The negative reviews that were associated with these shortcomings, contributed to our observation of persistence, i.e. the continued participation even after experiencing problems. Had the CDA been released without these problems, persistence may not have been among our key findings. However, some of our other findings, such as the willingness to serve for the greater good, similarly support the prevalence of collective motives.

4.7 Conclusion

In this paper we studied the motivations of citizen scientists to contribute to a VCS project in a time of crises (COVID-19 pandemic). For doing so, we collected and analysed reviews of the Corona Data-Donation App, a mobile application released by the RKI in Germany.
This application allows users to donate their physiological data using wearables, in order to mitigate the effects of the pandemic by possibly detecting local fever outbreaks that could be associated with COVID-19. We then conducted additional in-depth interviews with individuals that had experience with using this application, and analysed both the reviews and interviews. We thus could draw a picture of the users’ lived experience of participating in this VCS project, illustrating the cycle of engagement of the citizen scientists and people’s motivations to contribute such personal data under these circumstances. Our contribution in this paper is fourfold: First, we reveal that collective motives prevailed over egoistic motives, driving the citizens to contribute to the CDA. Second, the emotional response of the users included both positive and negative emotions, and we demonstrated which aspects of using the CDA resulted in particular emotions. We also show the repercussions this had on the overall use of the app and continuation of contributing to the VCS project. Third, we uncovered that citizen scientists of this particular project demonstrated a very strong sense of persistence, i.e. a large proportion continued to use the app in spite of various issues "for the greater good". Finally, we present implications for the design and presentation of future VCS projects, highlighting the importance of communication, transparency and responsibility. We thus drew insights for what motivates people to participate in VCS projects in times of crises and how designers of such systems can hopefully benefit from our findings in the future.
Understanding Health Information in Online Health Forums

Any fool can know.
The point is to understand.

______________________________
Albert Einstein

In the previous chapter, we explored individuals’ motivations to share health data in times of crisis. We presented suggestions for designers of citizen science projects to responsibly and effectively engage users to share data to mitigate crises. In this chapter, concluding our paper chapters, we describe the *lapse* and *understand* stages (Figure 5.1). This allows us to draw a circle from first engaging with digital health technologies to learning about one’s health status, arguably the ultimate goal of interacting with digital health technologies.
We first investigate how and when sub-communities for specific diseases emerge in OHCs. Second, we outline how psycho-social processes vary between threads of different lengths. We conducted a thematic analysis of 456 posts in 56 threads from the Dermatology sub-forum of an online health community. In our initial exploration, we identified thread length as a potential criterion for community involvement and separated short (fewer than 26 posts) and long (26 or more posts) threads for analysis. While low participation threads (‘short threads’) revolved around solving an individual’s health issue through diagnosis suggestions and medical advice, participants in high participation threads (‘long threads’) built collective knowledge and a sense of community, typically discussing chronic and rare conditions that medical professionals were unfamiliar with or could not treat effectively. We suggest that OHC developers address the specific needs of participants in threads of different lengths. For long threads a summary of health advice that has been found useful by other members could
improve the accessibility of information. Short threads could be explored via a sub-forum level search that categories content based on medical terms, e.g., diagnosis and treatment, embedding the seemingly disjunct information in a larger context.


5.1 Introduction & Motivation

The online discussion forum format includes a range of features that may be of particular benefit to those seeking health information online. Discussion is known to facilitate better learning and knowledge absorption [143]. This has been identified as a benefit for discussion forums in general [223] and specifically for online discussion forums [261].

In this paper we follow the definition of "online community" from Hammond (2017), who state that online communities are constituted by people who meet together in order to address instrumental, affective goals and at times to create joint artefacts. Interaction between members is mediated by internet technology. In order to constitute community, members need to: show commitment to others; experience a sense of connection (e.g., members need to identify themselves as members); exhibit reciprocity (e.g., the rights of other members are recognised); develop observable, sustained patterns of interaction with others; and show the necessary agency to maintain and develop interaction [95].

In the domain of social platforms, Jaimes, Lalmas, and Volkovich [115] have defined engagement as the phenomena of being captivated and motivated, and suggest that engagement can be measured in terms of a single interactive session or of a more long-term relationship with the social platform across multiple interactions. Thus, social media engagement is not just about how a single interaction unfolds, but about how and why people develop a relationship with a platform or service and integrate it into their lives [115]. Beuchot and
Bullen [20] as well as Chalkiti and Sigala [39] stated that online forums can promote the collaborative building of information and the management of knowledge. In an analysis of a forum for self-injury support, it was reported that forums construct their own hierarchies. A forum with less moderation assumes a flatter hierarchy where participants rely on each other to set norms of behavior, and if individuals post in a forum in a way that is in conflict with the forum community, other participants will correct the behavior [239].

Shaul [233] suggested that online forums could serve as a socially productive learning tool. He divided online forums into three categories: Social/opinion forums, general discussion forums and subject-specific forums. A certain type of subject-specific forum that especially addresses health-related issues can be labeled as Online Health Community (OHC). Here, participants seek information related to a specific disease [56], including information that will help them to diagnose a particular health problem [300], but also to receive treatment suggestions, often without or before consulting a doctor. Additionally, such OHCs also provide personal experiences of those with a similar condition and social support to participants who seek advice regarding diagnosis, therapy, medication or further strategies. Such communities form a special part of the online support for patients seeking advice or information. Hence, Adler and Adler reported that online discussion forums can go beyond the limits of simple knowledge exploration and can be a source of support for an individual [4].

Another aspect, driven by the increasing load on medical services, is that of supporting individuals to shift at least some of their health information seeking behavior from a face-to-face consultation with a qualified medical practitioner to seeking information online, both before and after diagnosis [50]. However, this also requires that the information sourced online is of a reasonably high quality, so as not to pose a health risk. Although previous research on online health knowledge has shown that it is of varying quality [110], a study by Cole et al. [50] indicated that forums which contained the most inaccurate or controversial information also contained counterbalancing comments which had the potential to reduce the harmful consequences of poor quality information. Comments made by the original poster,
and most respondents in a thread, suggested that the more accurate information carried more influence.

Informed by the prior work, we aim to provide increased understanding of the community building processes at work in online health forums. With this new knowledge we believe that there is potential for the user interface design of such forums to be developed to nurture informative and supportive communities. We address this through the following research questions:

- **RQ1**: When do sub-communities emerge in online health communities?
- **RQ2**: How do the psycho-social processes vary between threads of different lengths?

We report on the detailed analysis of 56 threads from the Dermatology sub-forum of an online health community. In our initial exploration, we identified thread length as a potential criterion for community involvement and separated short (fewer than 26 posts) and long (26 or more posts) threads for analysis. We found that the issues addressed in the analyzed threads often filled a knowledge or support gap that could not be satisfied by in-person doctor visits. Comparing the level of participation, our results suggest that short threads were helpful for solving an individual’s health issues, while in the long threads sub-communities about specific rare and chronic diseases emerged. We propose that the designers of OHC user interfaces should aim to better support the collective sensemaking processes present in long threads, e.g., supporting finding information in a thread that was initiated many years ago and grouping related content, making community knowledge accessible even in short threads.

## 5.2 Related Work

To position our work, we review prior works addressing the use of online medical information in general. We then explore related work that reports on the use of OHCs, aiming to identify motivations for their use and types of contributions.
5.2 Related Work

5.2.1 Online Medical Information

Rather than consulting medical professionals, either in a traditional physical consultation or through a telemedicine service, ‘Googling symptoms’, i.e. seeking self-diagnosis and treatment guidance from freely available online information sources is a common alternative today [155]. Koopman et al. [132] explored patient readiness to engage in health information technology, identifying individual patient’s preferred mode of interaction as a key criteria. Despite individual patient preferences in communication modes, the use of internet self diagnosis has been noted as a source of conflict between patients and medical professionals, by reducing satisfaction with medical professionals when they are later consulted [219, 237, 158]. The main reasons for conflict being due to patients’ inability to manage internet sourced information and its potential inaccuracy [237]. Prior studies about searching online health information have shown that information is of variable quality [144, 282]. In a study on online health information-seeking behavior, Cole et al. [50] found that while a small amount of information was assessed as poor, the original questioner probably would not have been led to act inappropriately based on the information presented. This suggests that online discussion forums may be a useful platform through which people can ask health-related questions and receive answers of acceptable quality.

5.2.2 Online Health Communities

In addition to simply ‘Googling symptoms’, many online discussion forums exist where patients can actively seek community health advice, e.g., reddit’s r/AskDocs[209] and patient.info[194]. In such forums, OHCs can emerge. Besides providing information or suggestions on diagnosis and treatment, OHCs have also been shown to provide social support among patients with cancer [232], diabetes [109], rare diseases [142], infertility [280], and HIV/AIDS [169]. Previous studies described OHC members’ information seeking practices [55], as well as ways they access and appraise information [97, 109], and construct new knowledge together [159]. Based on analysis of a diabetes forum, Mamykina et al. [159]...
5.2 Related Work

describe the value of the forum as exposing individuals to the richness and multiplicity of different perspectives, which help participants to construct their own personal views.

**Motivations**

To understand the behavior of users in online health communities, it is important to know why they visit and join such platforms. Huh [108] identified themes in an online diabetes forum, suggesting that thread initiators’ motivations were often not only to get clinical expertise, but also to hear other patients’ personal experiences. Other motivations for posting in forums include not wanting to see a doctor, wanting immediate discussion before a scheduled doctor’s appointment, or a second opinion following a doctor’s diagnosis [108]. Investigating the motivations of pregnant women for seeking support online, Gui et. al [88] identified limited access to health care professionals, frustration with their health care providers, limited access to offline support and a mismatch between information from books of internet resources with their own experience as main factors. The types of support sought by pregnant women included advice, formal and informal pregnancy-related knowledge, reassurance and emotional support. The posts mainly addressed sharing experiential knowledge, passing on other health care providers’ opinions, suggesting going for professional help and action based on peers’ responses [88]. Tied to participants’ motivations are social roles which vary among OHC members. Yang et al. [299] identified 11 roles in an online cancer support community three of which are emotional support provider, informational support provider and story sharer. The authors found that the roles predict long-term participation and change over time from ones that seek resources to ones offering help.

**Types of Contributions to Discussions in OHCs**

Mamykina et al. [159] identified ten different types of contribution posts make to the discussion in a health forum, including asking a question, suggesting resources, agreement/disagreement with a previously stated position, further developing previously stated position, personal reconciliation and synthesis of previously stated perspectives. Besides informational support, OHCs provide emotional support, often defined as a response of the
community to a member’s desire to change their mental state, usually to be more optimistic, motivated and determined [173]. This support can take many forms but usually requires that a member who seeks it is integrated with the community and can capitalize on the existing social structures [104]. Building on these results, Nakikj & Mamykina [172] conceptualize socio-emotional needs as a member’s desire to change their mental state through social interaction with other members and as a result of their social integration with the group.

In summary, previous research established that patients frequently seek health information online. They consult online health communities not only for informational but also for psycho-social support. Prior work also highlighted the motivations that drive members of online health communities. In addition, past research addressed the types of contributions common in such communities. We aim to extend existing literature by juxtaposing the different mechanisms of engagement in short and long threads in OHCs and by exploring the collective knowledge construction and community building.

5.3 Method

In this study, we conducted a qualitative analysis of a sub-forum in a large English-speaking online health community. Although all the information posted to the forum is publicly available, we do not name the specific community to avoid reverse identification of participants. The forum contains sub-forums addressing many medical specialties. Threads contain text and optionally images. Responses can be posted to up to three levels, e.g., an original post can have up to two levels of subthreads.

To familiarize ourselves with our area of study, we first studied the “DermatologyQuestions” sub-forum on Reddit (so called subreddits, e.g., r/Dermatology). Reddit is a discussion website that hosts a range of non-medical and medical subreddits. Since the “DermatologyQuestions” consists of a large community, including some medical professionals, we first conducted an online interview with a community member, who was also a certified dermatologist. We then developed a thematic codebook by analyzing 150 threads (including 244 posts) from r/DermatologyQuestions. Analysis revealed that the Reddit posts often comprised only
of short questions and answers, making it difficult to identify types of contributions and community dynamics.

With our developed codebook, we shifted our focus to analyse another site that included health forums which appeared to have richer discussions. We specifically focused on the sub-forum “Dermatology”, as our initial exploration revealed that it was among the most popular sub-forums and included threads with a variety of lengths. For the main analysis of this paper, we used our developed codebook to analyze 50 short (with a total of 124 posts) and six long threads (with a total of 332 posts) from the OHC. The frequency of various thread lengths in the OHC is visualised in Figure 5.2.

5.3.1 Data Collection

Our dataset consists of textual information and images that were directly uploaded to the post or linked on image hosting websites, e.g., imgur.com. To retrieve the OHC posts, we used

Fig. 5.2 Frequency of reply counts of threads (n=1854) in the Dermatology sub-forum on the online health community (specific platform not mentioned to ensure participant privacy). The majority of threads (n=406) has no reply. Counts for threads with 11-15 (15), 16-20 (20), 21-25 (25) posts are aggregated into the respective category. Included are all threads publicly available as of January 27th, 2021. We selected the most recent 50 threads with fewer than 26 posts and the most recent six threads with 26 or more posts for our study.
the python package *beautifulsoup* to parse the source code. All data was collected between July 29 and August 27, 2020.

### 5.3.2 Analysis

We applied open coding combined with thematic analysis in line with Blandford et al. [23]. For the analysis of the data we created a jupyter notebook with the python package *ipyannotator* with some modifications for the forum data hierarchy and to display the photos together with the post. Existing software for qualitative data analysis was not considered suitable for this rather specific task. For the preparation of the codebook, author DD developed an initial codeset based on prior work on types of contributions to online health communities [159]. The codeset was then refined and extended with authors AC and JW, e.g., we adopted the code ‘asking question’ but also introduced new codes (i.e., ‘giving advice’ and ‘own experience’). Moreover, we added symptoms, such as pain, and body parts to the list of categories.

As our initial analysis revealed that some longer threads span over multiple years and seemed to be of high significance for a group of participants with similar health problems, we decided to investigate thread length as a factor. We initially identified a small number of threads that contained a large amount of posts. Through review of these threads, we noted one thread, *moisturizing lotion*, that was just long enough that group building elements were beginning to emerge. Hence, we defined the minimum length of what we refer to as long threads, based on the length of this thread. We are aware that this is a somewhat arbitrary definition. For our main study, we analyzed 50 short threads, i.e. with fewer than 26 posts, and six long threads, with 26 or more posts, in total 332 posts. We iteratively compared our coding tree with our dataset until we reached data saturation [23]. All the analyzed threads are still open and accepting new posts.

### 5.4 Results

In the following section we first describe key differences between short and long threads, then, secondly, present the qualitative content analysis, and, thirdly, explain our model of
5.4 Results

engagement in long threads. To ensure participants’ privacy, we do not mention user names when citing forum posts.

5.4.1 Comparison between Short and Long Threads

Figure 5.3 presents the five most frequent themes in the analyzed short (n=50, 124 posts) and long (n=6, 332 posts) threads. Whilst some similarities are apparent, there are clear differences in the weighting of themes between the two cases and several themes, e.g., diagnosis and see a doctor, were only identified in one thread type.

The initial posts in the analyzed short and long threads were generally similar, describing the symptoms with which the poster was suffering. However, in the case of short threads, a diagnosis was often specifically sought (present in 20% of posts), e.g., ‘What is this?’ In almost all the analyzed long threads the participants had a clear understanding of the medical condition with which they were suffering. Here, the *Cat scratches* thread was an exceptional case where there was debate on the underlying cause of the symptoms. Descriptive statistics are shown in Table 5.1 and the content of the long threads is summarized in Table 5.2.

A clear difference between short and long threads was the difference in balance between sharing one’s own experiences (12% vs. 45%) and giving advice (44% vs. 22%). In short threads, the discussion was mainly focused towards solving a health issue of one specific user. Respondents were typically not suffering from the health issue themselves, but

![Fig. 5.3 Comparison of five most frequent themes in the short (n=50, 124 posts) and the long (n=6, 332 posts) threads. While the posts in the short threads were mostly direct support through giving advice, e.g., about medications and treatments, members in the long threads presented their own experience with their medical condition most frequently. Often, posts in the short threads included diagnosis suggestions, which was not present in the long threads. One post could be assigned multiple themes, therefore the percentages do not add up to 100.](image)
5.4 Results

<table>
<thead>
<tr>
<th>Thread</th>
<th>Posts</th>
<th>Post Length (words)</th>
<th>Time Span</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cat Scratches</td>
<td>128</td>
<td>$M = 91.5$</td>
<td>5 years, 10 months</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$SD = 110.5$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pityriasis Versicolor</td>
<td>42</td>
<td>$M = 98.8$</td>
<td>5 years, 4 months</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$SD = 66.1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cannot sweat</td>
<td>35</td>
<td>$M = 131.0$</td>
<td>4 years</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$SD = 106.9$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moisturising lotion</td>
<td>26</td>
<td>$M = 50.8$</td>
<td>5 days</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$SD = 47.3$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLC sufferer</td>
<td>87</td>
<td>$M = 104.1$</td>
<td>6 years, 4 months</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$SD = 94.3$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sebaceous cysts</td>
<td>50</td>
<td>$M = 153.5$</td>
<td>11 years, 5 months</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$SD = 124.2$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.1 Descriptive statistics of long threads. For "Participants", all posts made by guests were counted as one participant.

provided advice, e.g., treatment strategies and medications, or shared diagnosis suggestions. Sometimes (14% of posts), posters were instructed to see a doctor if respondents suspected potentially harmful medical issues or were unsure how to help.

Long threads revolved around a single health issue that was shared by all the participants. Altogether, 67% of posts included descriptions of personal medical history (22%) or experiences (45%), sometimes together with advice, e.g., reporting on cysts after cycling long distance and recommending laser hair removal and hydrocortisone cream. Almost without exception, respondents were affected by the same condition, therefore there was less discrepancy in the role participants played in the discussion. As in most cases participants had already seen, often multiple, medical professionals with unsatisfactory outcomes, there was very rarely any advice given to consult a doctor.

Analysis of the length of individual posts in both thread types revealed a significant difference in the length of post for short threads ($M = 75$ words, $SD = 84$) vs. long threads ($M = 102$ words, $SD = 104$); $t(558) = -3.38, p < .001$. Approximately 7% of posts in short threads contained more than 200 words, whilst for long threads this amount was double (14%). Differences between the analyzed long threads are also notable. In the Sebaceous
5.4 Results

<table>
<thead>
<tr>
<th>Thread</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cat Scratches</td>
<td>This condition is mystery to medical professionals, who don’t believe the conditions exists. Diagnoses include fungal, parasites, connection to Lyme disease. Potential causes are ambient temperature, household mold. Treatments suggested are ointments, camphor oil, bathing in vinegar, covering the body with a paste of coconut oil and colloidal silver.</td>
</tr>
<tr>
<td>Pityriasis Versicolor</td>
<td>Most participants have identified the disease they are suffering from (Pityriasis Versicolor). Participants share treatments, e.g., creams, anti-fungal medication, medicated shampoo, coconut oil. Also dietary changes (low sugar, gluten free, dairy free and mushroom free) are suggested. Symptoms are reported as particularly affecting the mental health of the participants.</td>
</tr>
<tr>
<td>Cannot sweat</td>
<td>Practices to mitigate the painful effects of the condition of being unable to sweat. The original poster (OP) posts the diagnosis of his condition (Cholinergic Urticaria), noting that it has no permanent cure, but proposes an treatment regime (physical exercise) that has reduced the reduce the severity of the symptoms by 80%. For several of the participants the original poster has become the expert on the condition.</td>
</tr>
<tr>
<td>PLC sufferer</td>
<td>Participants in this thread suffer from the condition pityriasis lichenoides chronica (PLC), which is characterized as rare and mostly unexplored. Many have had the condition for more than 10 years. The thread discussion concludes in a treatment regime that has worked for several participants. Due to the lack of a cure, participants report mental challenges.</td>
</tr>
<tr>
<td>Sebaceous cysts</td>
<td>Participants in this thread suffer from a condition of reoccurring cysts on the bikini line. Many report having the condition for more than 10 years, undergoing multiple painful surgical procedures and accruing $1000 in medical bills. Participants express frustration with medical professionals. A strong theme is that the forum provides the feeling of not being alone with the condition.</td>
</tr>
<tr>
<td>Moisturing Lotion</td>
<td>Participants discuss allergic reactions to certain brands of moisturizing lotion, providing details on the products they are using and highlighting the ingredients that are the potential cause of the reaction. This is a relatively recent thread, and, at time of writing, 7 months have passed since the last post. Only 5 participants take part in the thread with more than half of the posts (14/26) being made by the OP. The discussion is particularly light-hearted and chatty, e.g., when discussing a possible allergy to formaldehyde the OP jokes about the shock the undertaker will get when they are embalmed.</td>
</tr>
</tbody>
</table>

Table 5.2 Content summaries of the six long threads.
5.4 Results

*cysts* thread almost 30% of the posts exceeded 200 words, whilst in the *Moisturising lotion* case the longest post was 124 words in length.

### 5.4.2 Model for Community Building in Online Health Communities

Based on our findings from exploring the novel characteristics of long threads, we formulated a model of engagement (see Figure 5.4) to explain the psycho-sociological processes of collective knowledge building and community building that we observed in such threads. The driver of an individual to participate in such threads is suffering from a rare or chronic diseases that likely affects their mental health. When doctors can not help, the mental condition may further deteriorate. Patients therefore seek help in online health communities and actively search for other people with the same symptoms they suffer from. Many will share their own experiences, adding to the knowledge pool of the community and providing a tool to support their own self reflection. This exchange of experience and psycho-social support, often simply relief that there are others with the same condition, creates a sense of community among the participants.

Reflecting on the six long threads in our dataset against the identified criteria for community building, we notice that the *Moisturising cream* thread, whilst meeting our thread length threshold, does not support any of the other criteria defined in our model of engagement. The described topic is neither rare or chronic and rather than negatively affecting mental health, the discussion in the thread is of a jovial nature. It is also notable that the length of posts in the thread is far below the average identified for the other five community building long threads (see Table 5.1).

**Drivers for Participation**

Community building threads emerge around participants’ experiences of conditions that affect them severely and for long periods of time. Some participants in our dataset reported a history of having suffered with conditions for more than 20 years, including frequent doctor changes and unsuccessful treatments.
Fig. 5.4 Model of high engagement threads in online health communities, that lead to the development of sub-communities

Analysis of the long threads revealed two reasons of why participants looked for advice beyond that from medical professionals. Firstly, doctors were not able to help. Participants reported being prescribed a multitude of different creams, tablets, management regimes, and even undergoing repeated painful and costly operations – often with little or no success. Secondly, participants reported that doctors were often unfamiliar with their conditions, possibly because the conditions addressed in the threads were typically rare.

It was also apparent from the long threads that many participants had made extensive research on their condition, and, based on this, critically evaluated any advice they received from doctors. In some cases, participants reported dissatisfaction with their doctors, considering the doctor had not done enough research or describing the doctor’s behavior as rude. Some participants proposed showing the discussion from the thread to their doctor, e.g., reporting that they would bring posts from this forum to their dermatologist so that they would understand that these are not scratch marks and many people are going through this. Another participant highlighted the frustration of doctors not considering their condition to be real, explaining that they had been banned from the dermatology department at their local hospital and doctors suggest the participant is imagining it. Further, in some cases the treatment proposed by some doctors was considered impractical, e.g., one participant reported she was instructed to stay out of the sun, which she considered bad advice as she is
5.4 Results

young and wants to go on holidays. Hence, alternative solutions were sought through own
research and exchange in the forum.

The severity and chronic nature of the symptoms led participants to frequent comments
on their mental health. Participants did not know what to do next, resulting in feelings
of hopelessness, frustration and desperation. For example participants stated that some
symptoms are so slow developing that it feels like nothing has changed until one day they
realized how gray everything feels. Another participant resonated that she gets depressed
while breast feeding her baby and that she is struggling with PLC and, having had the disease
for 8 years, she feels hopeless.

The dermatological conditions in the scope of our study affected patients’ quality of
life in multiple ways. For example, efforts to maintain high hygiene, both in the domestic
environment and the body, impacted to daily routines, e.g., steam cleaning bedding and
washing hair daily. Similarly, participants reported needing to wear clothes that covered the
parts of their skin affected by the condition. Feelings of shame were therefore associated
with exposing the condition in public. Other participants had to stop leisure activities such as
bike riding, due to cysts appearing in the groin area.

5.4.3 Activities in Online Health Community

The main value of the long thread discussions was the building of collective knowledge and
community among the thread participants. From our dataset, we identified a strong motive
in the long threads was the exchange of informational and emotional support. Participants
encouraged the community to contribute their experiences to detect patterns, identify success-
ful strategies and spread knowledge. Participants shared their experiences primarily through
lengthy detailed descriptions of their medical history. When a successful strategy was shared,
other participants followed the advice. There was active exchange on the strategies, often
resulting in participants demanding updates on others’ proposed ideas, e.g., following up to
ask whether they had tried certain medications. Consequently, progress reports of successful
therapies were also shared, e.g., reporting that their PLC is 75% gone, having eaten healthy
for seven weeks, which consequently changed the quality of their life and also encouraging others to beat the disease together.

However, feedback on the success of shared medication suggestions and treatment strategies was often missing. Participants indicated that they wanted to follow their peers’ advice, sometimes even asking follow up questions, e.g., the exact name of a medication, but rarely reported back to the forum on the success of the strategy. Similarly, when members announced that they were planning to visit a doctor, other members asked for updates on the outcome. In some cases information was forthcoming, but in others the thread dried up, even though the information was begged by multiple different participants.

Support was not limited to exchanging medical advice, participants created a sense of belonging through providing reciprocal psycho-social support. This was expressed as relief when hearing about other participants with similar experiences. While it seemed that participants’ doctors, family and friends had trouble understanding the physical and mental implications of their condition, participants found empathy in the forum, e.g., letting other participants know that there is a solution and that they are not alone. We observed that the sense of community was a key factor in managing the emotional stress associated with the participants’ conditions.

### 5.4.4 Emergence of Sub-Communities

Active participation in forum threads over long periods of time formed sub-communities around a specific health problem experienced by many participants. The birth of such a sub-community requires a critical mass of participants, evident from the extremely long time period many of the analyzed long threads were dormant, before becoming active (see Figure 5.5). Some participants stayed active in the community over several years. Interestingly, the notion of time appeared to be of limited importance to the community participants, with many participants responding to posts from more than 2 years previously. Rather, what was important was finding people with the same medical condition.

From our dataset we noted that patients were often left behind the “real world” and found solutions to their health problems exclusively in the community. This applied to emotional
support too. Some posters noted that they felt relief having found other people with similar issues. Apparently, the real world social support circles, such as family and friends, could not care for the participants’ specific social needs. This gap was mitigated through psycho-social support exchanged in the forums.

5.5 Discussion

From our analysis we find that one of the main motivators for people to search for medical information in long threads is that they are not satisfied with their doctor, either because the doctor could not help, often due to the condition being rare or unexplored, or because the doctor did not take them seriously (RQ1). Similar feelings, as motivation for patients to participate in online health communities, have been reported by Gui et al. [88]. While there were some occasions where forum participants forwarded health information they had received from medical professionals, such as the names of prescribed medications, many discussions revolved around treatment strategies developed through personal experience while suffering from a condition, sometimes over many years or even decades. Here our findings align with those of Huh et al. [108], who identified searching for other patient’s personal experiences and avoiding doctors as a common motives for consulting online health communities. While we did not analyze the social roles of participants in detail, we noted that
in one instance the initiator of the 'Cannot sweat' thread clearly transitioned from seeking information to offering advice, suggesting that motivations can change during the exchange with fellow patients. From these observations we conclude that the discussion forums were a valuable resource for both informational and emotional support and compensated for needs that had not been satisfied outside of the online health communities. While similar findings have been reported by others [104, 172, 173, 159], we found that such discussions in the long threads emerged around rare and chronic diseases when health care providers could not provide support anymore.

Our results also suggest that threads with different lengths serve different purposes (RQ2). While short threads are often directed towards an individual’s condition, longer threads, i.e. with 26 or more posts, evolved into sub-communities, where a sense of community developed among participants. In the short threads, there was usually a clear distinction of roles between those participants asking for advice and those giving advice. In the long threads, however, participants had similar motivations, i.e. discussing a condition they were all suffering from, and therefore the social roles were more similar. Referring to Beuchot & Bullen [20], who report that forums build their own hierarchies, we note that longer threads had a flatter hierarchy than short threads which had a deeper information seeker and advice giver hierarchy.

5.5.1 Implications

Participants who access long threads many years after the thread discussion started may have difficulty identifying relevant health advice. These threads sometimes span over dozens of posts across many years, making it hard for participants to find advice that has been “validated” by the community. Additionally, in long threads discussions might become divided into multiple streams, i.e. replies to posts, and discussions in parallel and detached from each other. We suggest that OHC developers should account for this and, for example, provide a summary of posts that include health advice that has been found useful by the community, e.g., successful treatment strategies.
Further, we discovered that threads with more participation led to deeper involvement of individuals with their condition and the community. Thus, our proposals to provide summaries of threads and highlight helpful posts could be extended to a sub-forum level, allowing for an intelligent search function based on medical categories, e.g., diagnosis and treatment. Threads and posts in sub-forums could be automatically grouped into categories and participants would then enter a keyword into the search bar, e.g., rash, to navigate through different threads that address the same or similar medical conditions. With this approach, even short threads would become embedded into a larger context, supporting knowledge construction for individuals seeking health advice.

5.5.2 Limitations

We acknowledge that our results are limited in their generalisability, as we collected our data from a single sub-forum focusing solely on dermatological conditions. By sampling across various communities and sub-forums (e.g., other medical domains), future work could extend the applicability of our findings. Nevertheless, we believe our current findings contribute an important first step in exploring approaches to provide better medical support via OHCs.

5.6 Conclusion

Based on the analysis of a dermatology focused OHC, we identified that the number of posts in a thread is indicative of the purpose of the thread, the level of collective sensemaking processes as well as the duration of participation. While short threads addressed health conditions of an individual and were mostly in a question and answer format, long threads, with more than 26 posts brought forward sub-communities where participants shared own experiences and provided mutual informational and emotional support. We provide suggestions for developers of OHCs to allow participants to take advantage of some community building and knowledge construction elements even in threads with low participation.
Adaptability is key to survival in the age of digital Darwinism.

Rob Gonda, Forbes Technology Council

Digital health can make future health care more cost-effective, efficient, and more personalized for patients [267, 293]. This thesis aimed to investigate how to make the experience of engaging with digital health data more positive for the user. We presented the Digital Health Data Engagement Model (DH-DEM) and described gulfS and bridges to support designers of digital health technologies surrounding health data. Our findings originate from a user-centered perspective on the challenges of emerging digital health technologies. While our work does not provide a generic model for all medical domains and
user experience dimensions, we covered specific use-cases in detail to guide the research and design of digital health technologies.

Our work supports HCI researchers and designers in three ways. First, it provides a model for personal health data technologies that helps to structure the design process along user-centered stages. Second, we identify current gulfs in engaging with health data from the user perspective based on qualitative and quantitative analysis. Third, we make suggestions on how to bridge those gulfs based on findings from usability and user experience studies. This thesis addresses two groups of professionals, researchers and designers. The DH-DEM provides a conceptual model of digital health technologies’ data processes. Our model is beneficial for HCI researchers and researchers in related disciplines involved in digital health research. More precisely, our model allows researchers to gain a more nuanced understanding of users’ engagement with their health data in digital health technologies. We believe that by applying the model, researchers can draw a more holistic picture of users’ progress through the stages of digital health technologies. When technologies are designed to generate and share health data, it remains essential to understand why users first decided to use the technology. Similarly, systems that aim to support individuals’ understanding of their health data, e.g., the analysis of medical images, should consider how users may want to share more data to improve the quality of their analysis results. Designers of digital health data technologies will benefit from our practice-oriented collection of gulfs and suggestions on how to bridge them. Besides the high-level gulfs and bridges presented in Table 1.1, the more fine-grained suggestions in the specific chapters may be helpful for particular medical domains.

In contrast to most prior work on digital health data, we did not focus on organizational data sharing [2, 81, 80, 165], but rather on personal data sharing. To do that, we took the perspective of an empowered patient and provided guidance to designers to foster an individual’s ability to participate in decision-making. We believe that putting user needs at the center of design adopts the concept of empowerment known from the digital health domain [265, 156].
We argue that the DH-DEM contributes to a conceptual model of digital health data that extends prior HCI work on tracking physical activity data. Li, Dey, and Forlizzi [148] introduced a five-stage model of using personal informatics systems (preparation, collection, integration, reflection, and action). Later, Epstein et al. [71] extended this model to include "lived informatics", while Bentvelzen, Niess, and Woźniak [19] described conditions and barriers of reflection on personal data. All three models encompass health data that are tracked by individuals. Our work broadens this scope of personal informatics in two aspects. First, we consider health data beyond physiological data, e.g., medical images. Second, our model integrates data recorded irregularly or shared in online forums, whereas personal informatics models focus on individuals’ tracking data. Therefore the DHD-EM also addresses interventions to manage medical conditions, rather than only monitoring or health promotion as most prior work does.

The obstacles to data sharing between organizations, e.g., health care providers and research institutions, partly also apply to individuals’ data sharing. Prior work in the organizational domain identified factors relevant for sharing intentions and behavior. In our work, we found that many of them, e.g., privacy concerns [2, 192, 9], understanding of health benefits [11, 72, 192, 125], and the patient-physician relationship [2, 257], were equally relevant to individuals’ sharing intentions and behavior. Since most of the previous work is not rooted in the HCI field, the implications for designers of digital health technologies have mostly remained hidden. Our work provides insights on designing for positive user experiences when sharing health data. For the digital health field, our findings imply that more research is needed to learn about the relationship between individuals’ sharing behavior and the effects on quality of care, medical errors, and costs.

We think that the DHD-EM is generic enough to be applied to other personal health domains not addressed in this thesis, for example diabetes tracking. For this use-case, it could structure the process of (re-)designing diabetes tracker experiences, starting from the lapse of conventional diabetes care settings, e.g., limited availability of one’s endocrinologist. Following the proposed model, the next stage could be to inspect how blood glucose level measurements affect the daily activities of patients (generate). These measurements
are sometimes shared and discussed in online health forums, where other members give nutritional and other lifestyle-related advice. This interaction with peers potentially leads to a deeper understanding of one’s diabetes condition and ideally positively impacts one’s health.

So far, HCI has had a relatively small impact on the success of digital health technologies [22]. This thesis makes concrete suggestions for specific medical domains and gives guidance for research and design. However, only future work can assess if other disciplines in the field, e.g., public health researchers, adopt our findings into their work. In line with other HCI research [236, 147] we believe that only interdisciplinary design and evaluation of digital health technologies will lead to effective and sustainable solutions.

We suggest that future work should validate the gulfs and bridges in real-world contexts. While we present an important first contribution towards designing for a positive user experience when users engage in digital health data, an evaluation of the DHD-EM and the gulfs and bridges was not within the scope of this thesis. Due to the practice-oriented nature, the gulfs and bridges could be formulated as a checklist for developing and evaluating digital health tools in the domains outlined in our studies, i.e., ophthalmology, teledermatology, and outbreak response.

In the future, large health data sets will, even more than today, shape health systems. While HCI is only one of many disciplines involved in digital health, it has been shown to play a critical role in users’ acceptance of new technology in other domains [190]. Designing for a positive user experience in the health domain, as outlined in this thesis, positively impacts the effectiveness [228] and potentially health outcomes [240]. We, therefore, consider our work to be essential for the adoption of future digital health data technology that alleviates some of the burdens on current health systems and improves health for the individual.
Conclusion

The secret of change is to focus all of your energy, not on fighting the old, but on building the new.  

_Socrates_

Digital health has the potential to address many of the challenges current health systems are facing. Digital health technologies can improve health system efficiency and make patients more empowered. However, the success of Digital Health depends on the availability of data. This thesis aimed to investigate the user experience of engaging with digital health data and contribute to a more positive sharing experience, ultimately making digital health technologies more effective.
In chapter 2, we investigated the usability and trustworthiness dimensions of acquiring medical images with mobile devices. An interactive tutorial for EyeGuide, a mobile app that assists users in taking eye images with a smartphone camera, has allowed users to take images more quickly than with audible tone guidance. An additional user study showed that the AI-based diagnosis trustworthiness was higher when provided with disease-specific background information than other interventions, e.g., a disclaimer that the AI diagnosis does not replace a doctor visit.

Chapter 3 explored the emotional and physical aspects of medical selfies, i.e., medical images taken with mobile devices by patients. In a survey and a qualitative user study, we identified physical challenges, e.g., reaching shoulder blades and buttocks, learned that patients preferably check health issues related to genitals through an in-person doctor visit, and that taking medical selfies raised awareness of the medical state.

In chapter 4, we investigated the motivations and socio-psychological processes of citizen scientists who shared personal health data during the COVID-19 pandemic. Our findings indicate that collective motives dominated among app reviewers and interview participants. We contribute design implications for future virtual citizen science tools in times of crisis, highlighting the importance of communication, transparency, and responsibility.

In chapter 5, we presented an analysis of an online health community. We learned that sub-communities for specific diseases emerge in more extended discussions and identify socio-psychological processes around informational and emotional support. The needs of community members vary across short and long threads and should therefore be considered in the design, e.g., through providing sub-forum level search functionality based on medical categories.

We modeled the engagement with digital health data through the four stages lapse, generate, share, and understand. Through the DHD-EM, we support researchers in conceptualizing user needs in digital health technologies. For practitioners, we identified gulfs and provided bridges based on our observations and studies of emerging digital health technologies.
Lastly, we suggest how future work can address remaining research questions, in particular adopting our work in broader medical contexts, and the role of HCI in the interdisciplinary field of digital health.
List of Figures

1.1 The four stages of engaging with digital health data .......................... 6
1.2 The contributions of this thesis and their mapping to the stages of the Digital Health Data Engagement Model. .......................... 13

2.1 The DHD-EM and the contributions of the current chapter Generating Medical Eye Images and Trust in Automated Diagnosis (highlighted in blue) .................. 20
2.2 A typical flow of a smartphone based eye disease diagnosis app. ........... 22
2.3 Flow chart of the two study conditions interactive tutorial and auditory guidance. ................................................................. 26
2.4 Screenshots of the tutorial to take the eye photos. ............................... 27
2.5 Reference images for sharpness rating. ............................................. 30
2.6 Distribution of image sharpness ratings. ............................................ 30
2.7 The four design alternatives of the trustworthiness study. ..................... 32
2.8 The virtual ophthalmologist avatar. .................................................. 34
2.9 Trustworthiness scores between design alternatives. ........................... 36

3.1 The DHD-EM and the contributions of the current chapter Sharing Medical Photos (highlighted in green) .......................... 42
3.2 Examples of study participants’ skin photos. ...................................... 44
3.3 Survey results: Respondents who shared photos by body part. ............... 56
3.4 Survey results: Mean comfort level of male (n=57) and female (n=42) respondents to share medical selfies by body part. 60

4.1 The DHD-EM and the contributions of the current chapter *Sharing Digital Health Data for Health Surveillance* (highlighted in yellow). 69

4.2 Screenshots of the Corona-Datenspende-App on a smartphone. 77

4.3 COVID-19 cases in Germany and downloads of the Corona-Datenspende-App. 81

4.4 The experience of using the Corona-Datenspende-App during the COVID-19 pandemic. 84

4.5 Number of daily reviews submitted for the Corona-Datenspende-App. 85

4.6 Mean daily ratings of reviews submitted for the Corona-Datenspende-App. 85

4.7 Distribution of ratings for the Play Store and the App Store. 86

5.1 The DHD-EM and the contributions of the current chapter *Understanding Health Information in Online Health Forums* (highlighted in red). 103

5.2 Frequency of reply counts of threads (n=1854) in the Dermatology sub-forum. 110

5.3 Comparison of five most frequent themes in the short (n=50, 124 posts) and the long (n=6, 332 posts) threads. 112

5.4 Model of high engagement threads in online health communities, that lead to the development of sub-communities. 116

5.5 Timeline of posts in each of the analyzed long threads. 119

B.1 Coding tree used for the interview and review analysis of the Corona-Datenspende-App. 162
<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>The gulfs and bridges of engaging with digital health data across the four stages of the DHD-EM.</td>
<td>7</td>
</tr>
<tr>
<td>3.1</td>
<td>Questions study participants answered in a selfie video immediately after they took the photos.</td>
<td>53</td>
</tr>
<tr>
<td>3.2</td>
<td>Common interview questions asked one day after the capture session.</td>
<td>54</td>
</tr>
<tr>
<td>3.3</td>
<td>In-depth user study results: Most prominent themes identified from audio, video and interview transcriptions.</td>
<td>55</td>
</tr>
<tr>
<td>3.4</td>
<td>Survey results: Preferences for diagnosis mode by skin condition.</td>
<td>57</td>
</tr>
<tr>
<td>3.5</td>
<td>Study results: Body parts vs. photos taken, mirror used and practical problems (n=12).</td>
<td>59</td>
</tr>
<tr>
<td>3.6</td>
<td>Guidelines for the design of medical selfie mHealth apps derived from observations in the in-depth user study.</td>
<td>65</td>
</tr>
<tr>
<td>4.1</td>
<td>Demographics of interview participants.</td>
<td>83</td>
</tr>
<tr>
<td>5.1</td>
<td>Descriptive statistics of long threads.</td>
<td>113</td>
</tr>
<tr>
<td>5.2</td>
<td>Content summaries of the six long threads.</td>
<td>114</td>
</tr>
</tbody>
</table>
References


[28] Ellen Brady, Julia Segar, and Caroline Sanders. ““You get to know the people and whether they’re talking sense or not”: Negotiating trust on health-related forums”. In: Social Science & Medicine 162 (2016), pp. 151–157.


References


[152] Chi-Chung Lo, Sz-Pin Huang, Yi Cheng Ren, and Yu-Chee Tseng. “Yes, right there!: a self-portrait application with sensor-assisted guiding for smartphones.” In: MobiSys. 2013, pp. 505–506.


[180] OpenCV. OpenCV: Face Detection using Haar Cascades. 2018. URL: https://docs.opencv.org/3.3.0/d7/d8b/tutorial%5C_py%5C_face%5C_detection.html (visited on 09/21/2018).


[272] Vananh Vo, Lola Auroy, and Aline Sarradon-Eck. “Patients’ Perceptions of mHealth Apps: Meta-Ethnographic Review of Qualitative Studies”. In: *JMIR mHealth and uHealth* 7.7 (2019), e13817.


Corona-Datenspende-App: Interview Guideline

Introduction

1. Thanks again for taking the time

2. Introduction of the interviewer

3. Introduction of the research project

4. Interview process: duration 20-30min. We are interested on your perception of the pandemic and your personal attitude towards the app. There is no right or wrong anwers, we are interested in your personal opinion/perspective.
5. May we record the interview?

6. We will ask for the following data: your age, your profession and your gender. These data will be pseudonymized and only used for research purposes.

7. You can cancel the interview at any time without any effects on your compensation.

8. Do you have questions, before we begin?

**Warm-up**

1. Please briefly describe the process from the moment you heard about the app until today

2. .. from download to (potential) use until today (or until the day the use is stopped)

3. Where and when did you install the app?

4. How did the installation go?

**Motivation**

1. For what reasons did you install the app?

2. How did you like the app (compared to your expectations)?

3. Do you expect "return services"?

4. Do you expect feedback on your state of health?

5. How do you think your donated data will help to fight the pandemic?

6. How do you estimate the danger for you personally by COVID-19?

7. How do you estimate the danger of COVID-19 for society?
Use

1. On which other platforms do you share your fitness data?

2. From which Tracker/Smartwatch do you donate data or have you donated?

3. Do you open the app regularly / have you opened the app? In which situations?

4. What kind of exchange regarding the app did you have with friends and acquaintances?

Your data

1. What kind of data do you think is being transferred?

2. What is your impression of data security?

3. How do you feel about the automatic forwarding of data in the background?

4. End of the interview

5. Is there anything else you would like to say?

6. Voucher will be sent to you by mail

7. prolific: ask for prolific ID so that we can accept you. If not available via telephone and ID, please send a message via prolific.

8. Again, thank you very much for your time
Coding Tree: Corona-Datenpatent-App
Fig. B.1 Coding tree used for the interview and review analysis of the Corona-Datenspende-App.