Master Thesis, Interaction Design

Understanding User Experience of mHealth and Sensor Technology as part of Anxiety Treatment

Mette Elsborg and Louise Gram

AALBORG UNIVERSITY
STUDENT REPORT

Department of Computer science
Selma Lagerløfs vej 300
9220 Aalborg Ø

AALBORG UNIVERSITY
STUDENT REPORT

Title:	Understanding User Experience of mHealth and Sensor Technology as part of Anxiety Treatment
Semester:	10th semester
Semester theme:	Master Thesis
Project period:	January - June 2019
ECTS:	30
Supervisor:	Anders Bruun
Project group:	hci1020f19

Abstract

Mental health disorders pose a major concern worldwide. Anxiety is one of the most common mental disorders, and though there already exists well-established treatment programs with Cognitive Behavioral Therapy (CBT) being the golden standard, there is room for improvement. Sensor technology provides the opportunity to measure anxiety as it happens, which can help patients understand when their anxiety happens and potentially what causes it. There is however a lack of research on how the use of such sensor technology is experienced by the intended user group. In this project we conducted a longitudinal field study with four participants to understand how anxiety patients experience the use of a GSR sensor for measuring their anxiety and by using the GSR data as cues in a Cued Recall Debrief method setting. Furthermore, we conducted interviews with psychiatrists to learn their thoughts on the use of such technology in treatment. The three main results indicated that the system could be used for recalling anxiety episodes, confirming episodes and disconfirming episodes. We also propose a design of a mobile application to replace the corresponding desktop application to the GSR sensor used in the field study to make it more accessible, mobile and personal.

Number printed:	1
Pages:	77
Appendix:	126

Louise Gram

Mette Elsborg

By signing this document, each member of the group confirms participation on equal terms in the process of writing the project. Thus, each member of the group is responsible for the all contents in the project.

Summary

The use of sensors for providing objective data on the users propose a new possibility for personalized treatment through mHealth technologies. mHealth technologies are becoming increasingly popular as both a way to make treatment more accessible to a wider population and to improve the effectiveness of the current treatment. Galvanic Skin Response (GSR) sensors have proven effective in measuring arousal, which can be an indication of increased anxiety. Recent research has shown that the combination of mHealth and sensors provide the possibility of helping patients suffering from anxiety, yet studies also indicate a high drop out rate and furthermore there are not sufficient research on how these technologies are used in the field by the intended users.

This Master Thesis focuses on the user experience of using sensor technology in anxiety treatment. First a longitudinal field study was conducted and afterwards a prototype of a mobile application was designed and evaluated. We also had a meeting with a group of psychiatrists at a Professional Forum, as well as interviewed two psychiatrists.

Longitudinal Field Study
The longitudinal field study was conducted with four participants, all suffering from some kind of anxiety disorder. In the study, the participants received a wearable wristband: Empatica E4 and the following system to use for one-two weeks. The participation was conducted individually. The device was to be worn every day in the study period and become a part of the users everyday life. The participants were recruited either through the psychiatric ward at Aalborg University Hospital, or through a Facebook group. The four studies all varied in length which were 2 days, 3 days, 7 days and 27 days. The four participants consisted of three women and one man in the age of: 23, 29, 22 and 24 respectively.

The participants were instructed to use the system: E4 Connect that came with the Empatica E4 wristband, to monitor GSR levels in order to compare them to their own anxiety experiences throughout the day. The GSR data served as cues to help participants remember their anxiety episodes. The cues are inspired by the method: Cued Recall Debrief (CRD) known from HCI studies for assessing user experience of technology in the field. This method has been proven to be an effective method for eliminating memory bias. The participants were instructed to write down their anxiety episodes on custom made fill-out sheets based on the GSR data. The fill-out sheets were developed by the researchers for the purpose of this field study, but based on questions from the Day Reconstruction Method, which is also a method for assessing user experience of technology in the field. The fill-out sheets were similar to those used in Cognitive Behavioral Therapy (CBT), which is an acknowledged therapy form in anxiety treatment.

The results were analysed through open-ended coding and categorized into themes and then concepts. The results of the field study indicates a potential in using the technology of GSR data as cues in CRD in combination with CBT.

The results were presented to two psychiatrist, and an interview with them about the concerns and possibilities was a contribution to this project as well. The interviews showed that the use of the system would be dependent on each individual patient as the psychiatrists saw multiple potential use cases. They also had some concerns in regards to patients who are too focused on their anxiety as they might abuse the technology. Therefore, they stated the importance of an individual assessment of patients to determine the suitability and potential use case of the system. Also the two psychiatrists saw benefits in using the technology as an extension to current CBT and also using it in different types of mental disorders.

The main findings from the field study and the meeting and interviews with psychiatrists were that the system could be used for recalling anxiety episodes, confirming episodes and disconfirming episodes.

Design of Mobile Application

In extension to the results from the field study, on the fact that the current system is too rigid, a proposal for a mobile application, using GSR data as cues, was developed. The mobile application collects GSR data and gives the user the ability to directly input anxiety episodes on the mobile phone. The result was an interactive prototype that made it possible for the user to try out the functionalities of the app. The app consisted of a GSR data overview which is synchronizable to the Empatica E4 wristband. The user is then able to use an Episode Wizard, in which the user considers the GSR spikes and whether they were anxiety episodes or not. If so, the user is able to log the episodes and save them in order to later show the data in a therapy session. The app also consists of: 1) Insights or facts on anxiety, 2) Reflection of the user's life, 3) Setting up goals and 4) Writing in a diary. The app was evaluated with one user who found it helpful and more usable than the desktop application: E4 Connect.

Table of Content

1. Introduction

Mental health disorders pose a great global health concern as they contribute to a tremendous financial as well as personal burden worldwide both in developing and developed countries. In 2010 mental health disorders were the leading causes of years lived with disability, with anxiety disorders being some of the most common disorders (Whiteford et al., 2013). Anxiety disorders have an early median onset in childhood and a peak during adolescence, and is furthermore linked to a high comorbidity with other mental health disorders, as well as an overall reduced quality of life (Bry et al., 2018).

Anxiety is a state in which a person experiences renounced nervousness or fear for a situation, either be present or in the future (Chen et al., 2017). Anxiety covers several disorders such as: General Anxiety Disorder, Phobias, Panic Disorders and Social Anxiety. The most effective treatment method for anxiety is Cognitive Behavioral Therapy (CBT) (Tolin et al. 2010). The CBT method focuses on altering negative or destructive thought patterns within the patient (Beck, 2011). CBT consists of sessions with a psychiatrist where homework is assigned to the patient to challenge them to handle their anxiety and help them keep track of episodes through a diary or fill-out sheets. Completing homework has been proven effective but does put an increased burden on patients (Kazantzis et al., 2010), and CBT has a relatively large dropout rate of up to 40%, studies show (Winslow et al., 2016).

Recently, attention has been drawn to the potential of using mobile health (mHealth) technology, such as mobile phones and wearable devices, to monitor mental health and support treatment. The accessibility and functionality of these technologies make them ideal for health interventions as they have the potential to limit the costs and geographical limitations of current treatment (Bry et. al., 2018). Furthermore, these technologies often include sensors that can collect both physiological and contextual information about the patient, which can be used to provide more personalised mental health inventions to the patient (Garcia-Ceja et al., 2018).

Many of the mHealth technologies that currently exists for monitoring anxiety, mainly mobile applications, make use of subjective data through self-reports and are furthermore often not validated by health professionals (Bry et al., 2018). Relying solely on subjective data can be problematic, as these are prone to memory bias and can be a burden to the users (Boukhechba et al., 2018b). Physiological and contextual sensors can provide objective data to support subjective recollection. However, the use of sensors in existing mHealth applications have not reached the consumers yet (Bry et al., 2018), possibly because there is a lack of research on how representative users experience the use of a system that combines both physiological, contextual and subjective data for monitoring anxiety. Boukhechba et al. (2018b) did a study where they used self-reports, GPS and Heart Rate (HR) to assess the

change in emotional state for socially anxious patients participating in a treatment program. They used the self-reports to determine whether the sensors could successfully detect the participants' mental state. Levy et al. (2016) likewise used HR data to assess a change in mental state in patients with acrophobia following two different types of treatments.

There is a potential for not only using the sensors to assess a change in mental state following a treatment program, but also to use them actively as part of the treatment. As mentioned above, CBT is the most common treatment for anxiety, where homework is an essential element. Homework can consist of fill-out sheets for tracking anxiety and emotions connected to certain situations, places and times throughout a day (Mørch and Rosenberg, 2005). Some of these fill-out sheets are similar to those used in HCI studies for assessing users' experience with technology in the field, such as using the Day Reconstruction Method (DRM) (Kahneman et al., 2004) or Cued Recall Debrief (CRD) studies. In CRD studies the user can fill out self-reports based on cues from throughout the day rather than free recall (Omodei et al., 1994). In CBT these cues could be measurements of anxiety through Galvanic Skin Response (GSR) along timestamps and potentially GPS data, to support anxiety patients in recalling episodes, as this can sometimes be difficult for the patients.

Anxiety is an emotion that activates the Autonomic Nervous System (ANS) which regulates the body's physiological response by for instance increasing sweat production and HR (Poh et al., 2010). This state is also referred to as arousal, which is one of the dimensions of the *Pleasure-Arousal-Domination* (PAD) model commonly used for describing and measuring emotions. The arousal dimension span from excited to relaxed, while pleasure spans negative to positive in terms of how pleasant an emotion is. Domination describes how dominant the emotion is, spanning high to low. These dimensions can be assessed subjectively through questionnaires, or objectively through physiological sensors (Bruun and Ahm, 2015). Physiological sensors generally measure one of the dimensions, where arousal has been proven to be effectively measured through GSR, also known as Electrodermal Activity (EDA), and HR sensors. However, GSR sensors are not able to determine how pleasant an increase in arousal is (Lang, 1995), which could determine whether the arousal was a sign of anxiety or not. GSR sensors could therefore be useful in monitoring anxiety episodes by providing cues on increases in arousal but allowing the patient to subjectively assess the pleasure through self-reports.

It is proven that arousal can be measured through sensors, but more attention needs to be placed on how anxiety patients experience the use of such sensors for monitoring their anxiety. Therefore, we conducted a field study with four anxiety patients who wore a GSR wristband sensor for a period of time to assess their experience with it. We also discussed the potential of using such a system as part of the current treatment with a group of psychiatrists. Based on the field study and the discussion with psychiatrists, we designed and developed a prototype of an application to be used with the GSR sensor.

1.1 Research Question

With this study we want to further investigate the user experience of using a GSR sensor technology in treatment of anxiety with representative participants. We want to understand the possibilities and drawbacks of this technology as well as understand what tools can be made to support it. We therefore formulated the following research question and sub-question to base the study on:

> *How do anxiety patients experience reflection of continuously measured GSR sensor data, with the purpose of understanding their anxiety and how can this be a potential part of anxiety treatment?*
>
> - *How can we design tools that supplement this?*

1.2 Timeline of Project

A timeline shown in figure 1.a shows the structure of the process that formed this project. The process started out with framing the research area, and then looking into related work. Afterwards the field study was planned and executed. During the course of the field study, a meeting with psychiatrists at their weekly Professional Forum at the psychiatric ward in Aalborg was held. Shortly after this, and after finishing the field study, the process of designing a prototype of the mobile application began. Then we conducted two interviews with psychiatrists and presented the mobile application. The mobile application was then evaluated with one previous participant from the field study.

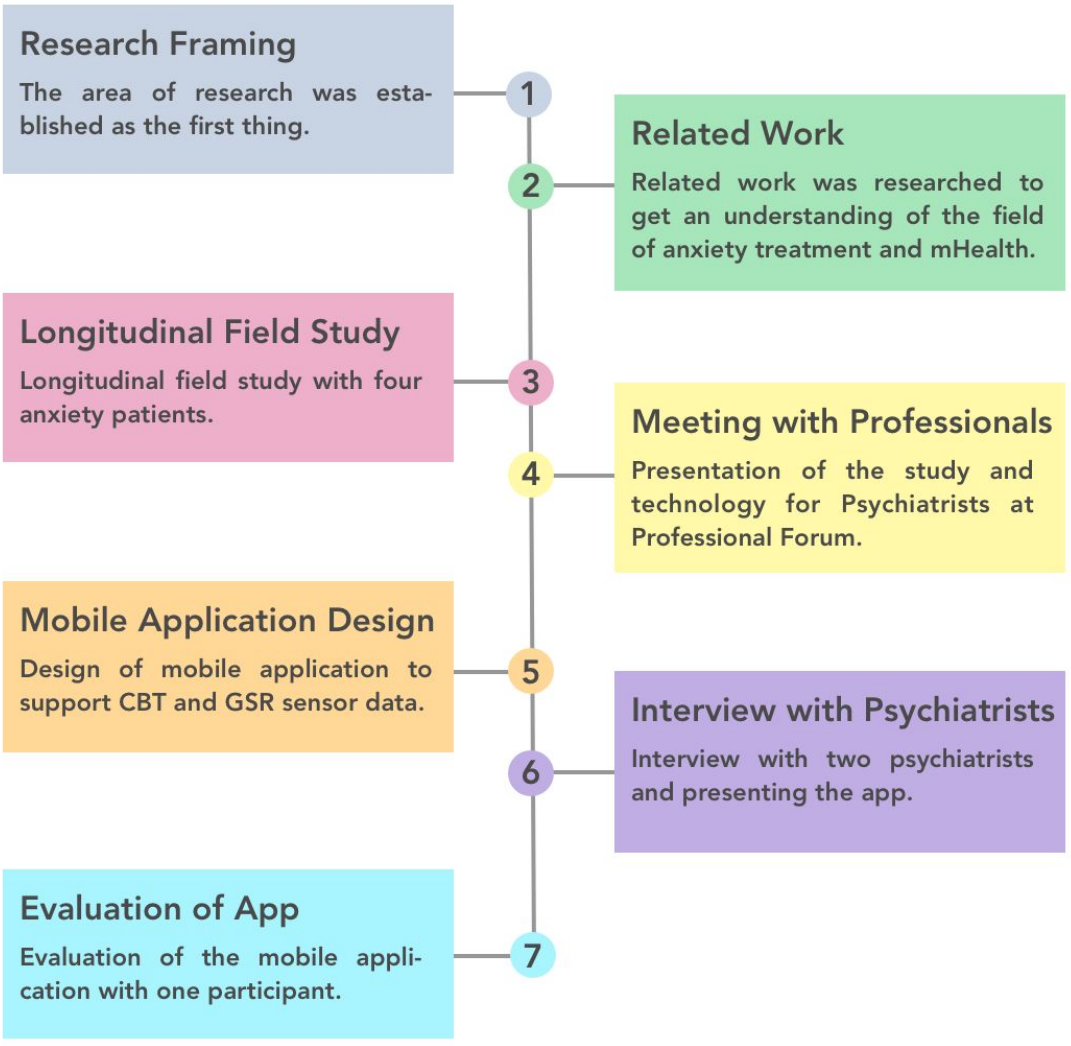


Figure 1.a Timeline of the project process

1.3 Report Structure

Chapter 1. Introduction
Firstly, there will be an introduction to the topics of mHealth, anxiety treatment and sensors. A research question and sub-question was also developed. The chapter also includes a timeline of the project process.

Chapter 2. Related Work
Following, there will be a chapter on related work in the field of anxiety treatment using sensor technology.

Chapter 3. Study Method
This chapter is where the study of user experience of sensor technology in anxiety treatment is presented. This chapter explains the methods, procedure, setting and data analysis of the longitudinal field study as well as the methods used for interviewing two psychiatrists.

Chapter 4. Field Study Results
The results from the longitudinal field study with four participants suffering from anxiety will be analysed and described in this chapter.

Chapter 5. Results of Interviews with Psychiatrists
Following the field study, this chapter describes the results of two interviews conducted with two psychiatrists.

Chapter 6. Design of Mobile Application
The design chapter focuses on the prototype development of an app that uses CBT theory to monitor and reflect on anxiety episodes for patients. In this chapter, there will also be a short evaluation of the prototype with one participant and two psychiatrists.

Chapter 7. Discussion
The discussion focuses on how current literature on mHealth, anxiety and sensors relates to the work and findings found in this project.

Chapter 8. Conclusion
In the end, a conclusion raps up the work done for this project, and answers the research question.

2. Related work

Using sensors in combination with mHealth technology to treat mental disorders is an emerging field of research. Especially in relation to anxiety treatment, the amount of literature and empirical studies are limited. A literature review conducted prior to this study yielded four empirical papers regarding anxiety with use of sensors in treatment. These four papers lay the foundation of the related work chapter, and will first briefly be described. Afterwards, the chapter will cover the following topics in relation to the four papers:

- Collaborating with Therapists and Use in Treatment
- Data Collection through Self Reports and Real-time Detection
- Cued Recall Debrief to Supplement Current CBT
- Sensors Usage
- Different Types of Study Methods

2.1 Introducing the Papers

Here is a short description of the four papers and their focus of study which is all related to anxiety and sensors. Two of the papers are by the same authors, Boukhechba, one from March 2018 and one from July 2018.

Acceptability of mHealth Augmentation of Collaborative Care: A Mixed Methods Pilot Study by Bauer et al. 2017 (from now on referred to as Bauer)
Bauer did a study to assess the feasibility and acceptability of a mobile health platform to support the Collaborate Care program, which is a patient-centered approach to delivering care for patients with depression and anxiety disorders. The participating patients as well as the caregivers used the platform during the study.

E-Virtual Reality Exposure Therapy in Acrophobia: A Pilot Study by Levy et al. 2016 (from now on referred to as Levy)
Levy did a study lasting three weeks with six participants who suffered from acrophobia. Their study consisted of two sessions of virtual reality exposure therapy each week, with the first three being remote sessions (E-VRET) and the last three, traditional sessions (P-VRET) with a therapist.

Physiological Changes over the Course of Cognitive Bias Modification for Social Anxiety by Boukhechba et al. 2018a, July (from now on referred to as Boukhechba A)
Boukhechba A recruited 20 highly socially anxious people to participate in a computerized Cognitive Bias Modification (CBM) treatment program, which is a new treatment program that does not require therapist assistance.

Predicting Social Anxiety from Global Positioning System Traces of College Students: Feasibility Study By Boukhechba et al. 2018b, March (from now on referred to as Boukhechba B)
Boukhechba B did a study with only GPS data and no physiological sensors, in a study to assess the feasibility of using mobile technology to passively examine college students' social anxiety levels. 228 college students participated in the study for two weeks.

2.2 Collaborating with Therapists and Use in Treatment

Collaborating with professionals doing development of mHealth technology for mentally ill is highly important (Bry et al., 2018). In a study conducted by Bry, out of 121 apps on the app store, very few were tested by professionals. This correlates with the four papers, in which only one, Bauer, has a close collaboration with the caregivers. Levy uses therapists in executing the study, but not in the development process. The same goes for both of Boukhechba's studies where there is no incorporation of therapists or professionals.

The four papers all use different types of therapy for treatment of anxiety. Levy uses Virtual Reality Exposure Therapy, Bauer uses Collaborative Care, Boukhechba A uses Cognitive Bias Modification and Boukhechba B does not focus on treatment but points to CBT in future work. In the study done by Boukhechba A, they recruited 20 highly socially anxious people to participate in a computerized Cognitive Bias Modification (CBM) treatment program, which is a new treatment program that does not require therapist assistance.

Levy did a study to compare traditional virtual reality exposure therapy (p-VRET) and remote virtual reality exposure therapy (e-VRET). VRET is a well-accepted treatment program for anxiety disorders and consists of exposing the patients to their anxiety triggers in a virtual world rather than in the real world. The difference between p-VRET and e-VRET is the presence of a therapist. Bauer uses a mobile application to join the caregiver and the patient. The caregivers used an online platform with an overview of all patients as well as individual pages, where they could see the data transmitted via the app. Even though high dropout rates occurred after the first four weeks, Bauer argues that there was potential in giving patients and caregivers an online platform to communicate.

2.3 Data Collection through Self Reports and Real-time Detection

Data collection can either be objectively detected through sensors or subjectively collected through self-reports. A mix of the two is Ecological Momentary Assessment (EMA). EMA is a method where the user is prompted several times a day to log their activity to avoid memory bias (Scollon et al., 2003). All four papers used EMA in one way or another.

In Boukhechba A's study, as an example of EMA, the participants used a mobile application which prompted them several times a day to fill out a questionnaire. The questionnaires were related to affect and social context, such as how negative they were feeling and who they were with. The app alongside a wrist-worn sensor also collected data such as GPS, accelerometer, HR and GSR. Boukhechba A used the self-reports as a 'ground truth' to

determine whether the physiological data, mainly HR and accelerometer, could successfully detect negative affect. Their motivation for studying the correlation between physiological changes and changes in mental states was to potentially eliminate self-report measures, as these often propose a great burden for users, and are furthermore prone to recall bias. The reasoning behind taking any type of measurement is to assess the effect of the treatment program, in this case: Computerized Cognitive Bias Modification (Boukhechba et al., 2018b).

In Bauer's study, the participants were likewise using a mobile application which gave them 3-4 notifications throughout the week to complete self-reports. The notifications consisted of regular clinical surveys, occasional satisfaction surveys and health tips. The clinical surveys were respectively Patient Health Questionnaire (PHQ-9) used to assess degree of depression severity and General Anxiety Disorder (GAD-7) used to assess the severity of anxiety. GPS data was also passively collected to assess movement. Patients reported that they found the amount of self-reports reasonable, however most of them dropped out of the study halfway through. Bauer claims the reason for the dropouts were due to a change in the platform during the study and some ambivalence and nuisense in the users' experience with the app, such as lack of personalization. Bauer also learned that it could be burdensome for the caregiver to go through the self-reports and filter out the important information, but did find the access to the information, along with alerts, useful. Alerts were based on the PHQ-9 and GAD-9 responses and GPS data, which flagged patients who seemed to be isolating themselves.

Boukhechba B started off by assessing participant's social anxiety level using the Social Interaction Anxiety Scale (SIAS) questionnaire, which consists of 20 items rated from 0-4. Throughout the study, they solely used passively collected data from a custom-made app that collected GPS location every 150 seconds. They used the initial SIAS score to see if their system could predict the participants' level of social anxiety based on their movement patterns.

Levy used both a HR sensor to passively detect the participants' mental state and had the participants fill out questionnaires before and after each session. Before starting the sessions the participants filled out four different questionnaires: Beck Depression Inventory (BDI), Spielberger State-Trait Anxiety Inventory (STAI), Attitude Towards Heights Questionnaire (ATHQ) and Acrophobia Questionnaire (AQ). Before and after each session the participants also filled out the STAI questionnaire and rated their anxiety on a visual analog scale (VAS) ranging from 0-10. They used the different types of data collection to determine the change in mental state following the E-VRET and P-VRET sessions respectively.

2.4 Cued Recall Debrief to Supplement Current CBT

In studies like the ones from Bauer and Boukhechba A, self-reports or EMA have been combined with mHealth technology to assist treatment of anxiety. The downfalls of using EMA, or Experience Sampling Method, which it is also known as, is the high participant burden (Scollon et al., 2003). Both in the Bauer and in the Boukhechba studies they experienced participant dropouts due to this. A high number of dropouts are also seen in CBT

(Winslow et al., 2016) due to homework. But homework or self-reports is an important part of CBT because it helps the user keep track of thoughts and when they occured.

Therefore, there is a need to address this issue of high burden on participants and an answer might be the Day Reconstruction Method (DRM). With the DRM participants are asked to fill out a schema or a diary at the end of the day about their experiences (Kahneman et al., 2004). The issue with the DRM is the memory bias because the user is forced to remember back on the day that has passed. In order to accommodate this, the Cued Recall Debrief (CRD) method might be more suitable. CRD is a method where participants at the end of a session is provided cues about the happenings of the session to recall certain events. The CRD was first invented by Omodei (1994) as a tool for eliminating the memory bias, by adding cameras on participants when conducting a study. The method has proven useful in situations where detailed responses of emotion is needed, and help minimize memory bias (Bruun et al., 2016).

2.5 Sensors Usage

The four papers use either physiological sensors, such as HR and GSR, or contextual sensors, either GPS or accelerometer. Boukhechba A used HR, GSR and accelerometer data to track the changes in the participants' mental state following the CBM program over three weeks. The participants did not have access to this data. The result showed preliminary evidence, that using HR and accelerometer can be used for identifying a change in mental state following a CBM intervention. Levy also used HR data to assess the different effects of the two types of treatment.

Bauer did not use any physiological sensors in their study, but obtained GPS data passively to assess movement to determine if a patient was isolating themselves. The caregivers' platform used the GPS data to flag patients who seemed to be isolating themselves. In relation to Bauer's study, Boukhechba B used GPS data to analyze the participant's movement patterns to predict if the participants suffered from social anxiety.

2.6 Different Types of Study Methods

In the four papers, the studies conducted are done long-term, in the field and with representative users. Representative users is understood as mentally ill participants appropriate for the given study. For instance, in the study done by Bauer, they did a long-term study over eight weeks with patients with anxiety and depression disorders. However, in their study, most of the participants dropped out after four weeks, probably dues to the high number of notifications provided by the app leading to high participant burden.

Doing a study in a laboratory setting limits the ecological validity of the results (Boukhechba et al., 2018a). Boukhechba B stated that there is a need for studying the different methods of assessing mental health symptoms in the field, both subjective, through self-reports, and objective, through sensors; as most studies have been conducted in a laboratory setting.

Boukhechba B study lasted for two weeks with 228 college students. Before the study, Boukhechba B assessed the participants' social anxiety symptoms in a laboratory setting using self-report instruments. They classified the participants as either low or high socially anxious. They used both low and high socially anxious participants to examine if their technology could correctly predict the participants' social anxiety level. Boukhechba A, also did a study using socially anxious college students. Their study lasted for three weeks and involved 20 highly socially anxious participants who used a mobile app in their daily lives.

Levy did a study lasting three weeks with six participants who suffered from acrophobia. Their study consisted of two sessions of virtual reality exposure therapy each week, with the first three being remote sessions (E-VRET) and the last three, traditional sessions (P-VRET) with a therapist. However, all sessions were conducted in a laboratory setting at the hospital. For the E-VRET sessions there were no direct contact between patient and therapist. The results showed no significant difference between the two therapy types indicating that E-VRET can be used to treat acrophobia.

2.7 Summary

Four papers reveal very different approaches to treatment of anxiety, even though they all rely on mHealth and sensors. Something they can agree on is using representative participants in long field studies to gain a representative result and using roughly the same sensors (HR, GSR, Accelerometer and GPS) to carry out their studies. Furthermore, the papers use both subjective measures through self-reports and objective through sensors, or a combination of both referred to as EMA. In the following chapter, the method for the field study we conducted will be described.

3. Study Method

To learn more about anxiety patients' experience with using a system for monitoring and measuring their anxiety using sensors, we aimed to conduct a longitudinal field study. Four people suffering from anxiety participated in the study. One of them used the system for 27 days while the others used it for 2, 3 and 7 days respectively. We also presented and discussed the system with a group of psychiatrists and conducted two individual interviews with two of them to learn their thoughts in terms of potentials and concerns when introducing the technology in the current treatment.

In this chapter, there will be a presentation of the system the participants used during the field study. Then, the reasoning behind the methods chosen for conducting the field study will be presented, followed by a presentation of the four participants. Afterwards, there will be a description of the setting, materials and procedure of the field study as well as the data analysis method. Finally, there will be a section describing the presentation and the individual interviews with the psychiatrists.

3.1 System Overview

To get an idea of what the system that was given to the participants looked like, the following is a quick overview and description. The technical part of the system used during the field study consists of two elements:

- The Empatica E4 Connect desktop application, seen in figure 3.a
- The Empatica E4 wearable wristband, seen in figure 3.b

The Empatica E4 wristband (Empatica, 2019) was chosen due to the possibility of measuring GSR that has been proven to measure arousal effectively (Lang, 1995), which increases when a person experiences anxiety. The wristband also measures Heart Rate (HR), skin temperature, accelerometer and Blood Volume Pulse (BVP).

To present the GSR data associated with the wristband, the desktop application E4 Connect was used. A mobile application is available as well for iOS and Android that supports the Empatica E4 wristband and is called: E4 Realtime. E4 Realtime consists of a livestream of GSR, HR, accelerometer and temperature. The app only live streams data and does not support an overview of the full duration of a GSR measurement. The app needs constant connection to the wristband in order to function. This made the app unsuited for monitoring anxiety episodes for a full day. The E4 Connect on desktop was therefore a better choice because it allows the user to get an overview of the monitored sessions. The E4 Connect for desktop is where the user can see five different graphs of all the measured data. The E4 Connect does not allow for filtering out any of these five graphs so the participants were

asked to solely focus on the top graph that shows GSR. A screenshot of the desktop application, showing the sensors: GSR (blue line), BVP (red line) and Accelerometer (purple line) can be seen in figure 3.a, the last two graphs are not seen in the screenshots but are HR and skin temperature. The wristband is seen in figure 3.b.

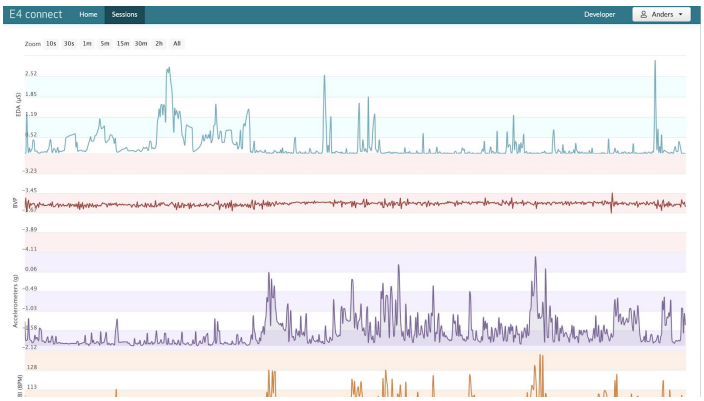


Figure 3.a Screenshot of E4 Connect desktop application

Figure 3.b Empatica E4 wristband

3.2 Methodology

We chose to do a longitudinal field study in order to understand the user experience of sensors and mHealth in the treatment of anxiety. This section will include a description of the methodology behind the field study as well as longitudinal studies and the reasoning behind choosing these methods for the purpose of our study. Afterwards the correlation between Cued Recall Debrief (CRD) and CBT will be described in terms of how it was used in the field study.

Field study

As today's technology is so embedded in the users' everyday life, field studies are becoming increasingly popular as a way to evaluate the user experience of these technologies (Preece et al., 2015). Field studies is an approach for studying users in a naturalistic setting. The downside of doing a field study instead of a laboratory study, is the lack of control due to not being able to plan tasks ahead. But what is lost, is gained as a more naturalistic view on the interaction between the user and the product. Especially in the case of initial research on a topic in HCI, field studies prove beneficial in uncovering the unexpected instead of confirming or declining what is already known (Rogers and Marshall, 2017). As the area of anxiety, sensors and mHealth is an emerging topic in HCI, the amount of research on the user experience is scarce. Though there are a few papers, as presented in chapter 2: Related Work, who studies this area in the field with representative users over time, there is a need for more research on the topic (Adams et al., 2017).

Longitudinal Research

When it comes to studying technology, the novelty effect of experiencing a new product might give an unclear result of the user experience of a product over time, because this is likely to develop as the user gets more familiar with the product. To learn more about this development, time is essential (Karapanos et al., 2009). Longitudinal studies are conducted with participants over a period of time in the participants' natural context. This is done to study how the user experience change as the novelty effect wears off and the product becomes more integrated in the user's everyday life. In longitudinal studies, the focus is on measuring changes in a set of metrics over time (Gerken and Reiterer, 2009). Gerken and Reiterer made a taxonomy for longitudinal studies to get an overview of how to conduct a longitudinal study. The taxonomy consists of set research considerations, that help the researcher scope the study at hand, yet the most important element of any study is the research question, also highlighted by Gerken and Reiterer. Our research area was the user experience of using GSR sensor data cues in combination with CBT treatment. The metrics that were measured during the study was mostly focused on usability and user experience. Some of the research considerations from Gerken and Reiterer are: Research Objective, Research Design and Data Gathering Design. These are discussed in the following.

To put our study in perspective to the taxonomy, we have worked with what Gerken and Reiterer calls: Change as a Confounding Factor, as our Research Objective. The reason behind this is that we know when we give the participants the system, a change will happen in their everyday life. We are more interested in understanding how the users everyday life change and what that means in their anxiety treatment.

The Research Design is a Prospective Panel, which allow for a wide understanding of the participant and system use, due to the length of these studies. When it comes to Prospective Panel studies, the user might start questioning his behavior, which might lead to a modified experience. This is why some length is preferable when conducting a study on user experience of a technology. (Gerken and Reiterer, 2009).

In relation to the Data Gathering Design, we used a combination of Continually and In Waves data gathering methods (Gerken and Reiterer, 2009). Continuous data gathering comes from the use of GSR sensor data and diary keeping. The In Waves data gathering comes from the interviews that were held in the beginning and end of the study, and for one participant also in the middle. There is a quality in combining these two types of data gathering techniques, as one emphasizes the other. For instance the data from the GSR was discussed during the interviews.

Cued Recall Debrief and CBT

As described in chapter 2: Related Work, there is a correlation between Cued Recall Debrief (CRD), a method in HCI for assessing user experience of technology, and Cognitive Behavioral Therapy (CBT), the golden standard of treatment for anxiety disorders. In CBT,

patients are asked to complete homework which could consist of filling out sheets about their anxiety episodes in relation to i.g. what happened, where it happened and how they were feeling. In CRD, participants are likewise instructed to fill out similar qualitative measures about their experience, but are provided quantitatively collected cues, throughout the session to support their recollection. Introducing the cues in CBT homework might therefore be beneficial.

CRD has shown potential in improving the free recall method, minimizing memory bias (Bruun et al., 2016). CRD uses cues to help participants remember emotions (Omodei et al., 1994), and in order to give the participants in our study something to relate the cues to, a custom made fill-out sheet was developed. The cues are collected as GSR measures through the E4 wristband and presented to the participants to help them complete the sheets. The fill-out sheets used for this study were based on the DRM presented by Kahneman (2004), and can be seen in Appendix 1. Current CBT homework sheets (Mørch and Rosenberg, 2005) are similar to the ones used in DRM. The fill-out sheets consists of seven questions in relation to the anxiety episode they experienced. The questions were: *What happened? Where did it happen? Were you alone or with somebody? What did you feel? What time did it begin? What time did it end?* The idea is that the GSR measures can help the participants see a timestamp of when a potential anxiety episode was experienced, to help them remember it.

3.3 Participants

While researching the existing literature on these type of studies it was made clear how previous studies often make use of non-representative users, or convenience sampling, which is problematic in terms of validating how the system will work outside a research context (Mohr et al., 2017). There are though a few studies who do use representative users as presented in chapter 2: Related Work. It can be difficult to get access to these types of participants as stated by Doherty et al. (2010), which was also evident during our recruitment of participants, which is why we ended up with four participants who took part in the study. We collaborated with the psychiatric ward in terms of getting access to patients, where two of the four participants were recruited. The other two were recruited through a Facebook group. The participants have been anonymized and the names presented in the report and Appendix are cover names. A list of all participants are seen in figure 3.c. Following the list, an individual description of each participant is found.

#	Name	Age	Occupation	Participation length	Participation length (actual use)
1	Caroline	23	High school student	2 days	2 days
2	Mathilda	29	School teacher	29 days	27 days
3	Maya	22	Medicine school student	14 days	7 days
4	Eric	24	English literature student	10 days	3 days

Figure 3.c Full list of participants

Participant 1: Caroline

Caroline is a 23-year old girl living on her own. She has no known anxiety diagnosis, but reacts strongly with hyperventilation, shiver and sweating. She is a student and has been in treatment within the psychiatry for approximately half a year. Sometimes it is difficult for her to remember when she had an anxiety episode, and sometimes she is unsure if it is anxiety or something else. She was provided with some fill-out sheets from her psychiatrist, but did not think she could use them. She is not keen on using technology, for instance she does not own a smartphone. She participated in the study for two days and then she did not feel comfortable wearing it anymore. Looking into her GSR data we saw that she wore the wristband for 4 hours the first day and 12 hours the second.

Participant 2: Mathilda

Mathilda is a 29-year old girl from Aalborg who works as a primary school teacher. She suffers from panic disorder and OCD, of which she is medicated for. She had been seeing a psychiatrist for some time, when we met her. It is easy for her to remember the big anxiety episodes but with the small ones, she does not even realise she is having one before it is over. She also fears anxiety episodes. She is familiar with writing down when she experience an anxiety episode in fill-out sheets provided by the psychiatrist and she has a regular routine in writing it in the evening time. She seems happy and very open towards the idea of monitoring her anxiety, and has no concerns regarding privacy. Her motivation for doing this study, is that she wants to help others with the same condition. Mathilda is very pro technology, and wears a smartwatch everyday herself. She uses an app to monitor her body signals and has tried multiple apps to relieve anxiety in the past. She enjoys monitoring her data, and is especially interested in her sleeping pattern. She is a math teacher which makes her happy with numbers and used to reading graphs. Mathilda was first told to wear the Empatica E4 wristband for 11 days, and then we had an interview about how it went. She was positive towards the wristband, but also stated that some questions were hard to answer after only 11 days, so we asked if she would like to extend the study period. She was willing to wear the wristband for another 18 days, making it a total of 29 days where she participated in the study. Looking at her GSR data afterwards it showed that she used the wristband for 19 to 33 hours a session and it was in use for 27 of the 29 days.

Participant 3: Maya

Maya is a 22-year old girl from Aalborg who is studying medicine at the university. She has had anxiety since she was little but has only been diagnosed for two years. She is diagnosed with panic disorder in a generalized way as they have not found the trigger yet. She has been in therapy before but not anymore as they were not seeing much progress in figuring out what triggered her anxiety and furthermore, she had learned some effective techniques to get the anxiety episodes under control when they happen. When she was in therapy she used fill-out sheets that she made herself. She now experiences small anxiety attacks a couple of times a week and has a difficulty in remembering them. In relation to technology, she uses a wide variety of apps and has also previously been writing down anxiety episodes in an app to keep track. She does not use or own a smartwatch. Maya were scheduled to use the system for two weeks but due to illness during her course of the study she therefore used the wristband for seven out of the 14 days planned. She wore the wristband from 1 hour in a day to about 15 hours in one session.

Participant 4: Eric

Eric has suffered from anxiety for four years and is diagnosed with panic disorder. He has been in CBT treatment for a limited period of three-four sessions in total. He lives in a dorm with three other students in Aalborg, while he is doing his bachelor's degree in English literature. He is reluctant to undergo treatment for anxiety especially with medicine, but wants to participate in the study to help others. The anxiety episodes he experiences were at the most powerful two years ago. Today he gets them very rarely, and claims he only experience them once-twice a month, though he also states that he believe his baseline anxiety is higher than the average person. When an anxiety episode approaches, he remembers them in detail, both time and place. He does not have difficulty in remembering episodes, and does not write them down when they occur. Eric does not own a smartwatch but uses a smartphone with a variety of apps, none for managing his anxiety though. Eric wore the wristband for seven days, but due to system issues, only half of the days only three days, were captured. In the three days he wore it from 8 hours to 25 hours in a session.

3.4 Setting and Materials

In this section the setting and materials used for the field study will be presented. First, there will be a description of the interviews, including the setting and the interview guide. Then a presentation of the home kit which the participants received at the initial interview.

Interviewguide

The interviews were semi-structured as this would allow the researchers to ask follow-up questions and have a fluid conversation. The questions for the initial interview where based on getting to know the patient and his or her therapy form. The questions from the initial interview is seen in Appendix 2.a. In regards to the concluding interview the questions were inspired by the User Experience Questions (UEQ) from Hassenzahl (2001) and a paper on

technology acceptance in healthcare (Gao et al., 2015). To see the concluding interview guide, see Appendix 2.b. For the third interview that was held with one of the participants, the interview guide was based on her answers from the first concluding interview. To see this interview guide see Appendix 2.c.

The UEQ are questions based on the user experience of a product or technology. The questionnaire is made up of anchoring conditions such as *good* and *bad.* The UEQ are usually listed in a questionnaire format, where the user positions themselves on a five-point scale where 1 is bad and 5 is good (Hassenzahl, 2001). For this field study the conditions where made into questions which the participants could answer. The questions are used to get an understanding of a product's perceived user experience. An example of one of the UEQ in the concluding interview was: *Do you think the wristband is exciting or boring?* Where the user will answer the one or the other and the interviewer would ask a follow up question to elaborate.

Gao et al. (2015) proposed an integrated technology acceptance model specifically aimed at wearable devices in healthcare. The model is based on unified theory of acceptance and use of technology 2 (UTAUT2), protection motivation theory (PMT), and privacy calculus theory. The model consists of 10 factors that influence users intention to adopt wearable technology for healthcare. Under each factor there are three statements, the user is asked to evaluate on a scale. These statements were changed into questions for the concluding interview in our study. An example from Gao is, under the factor: *Performance Expectancy: "I find the device useful in my everyday life",* which was formulated into "Do you think the wristband was useful in your everyday life?", and then the participants were asked to elaborate.

Home kit and preparation

The home kit consisted of: a wristband (Empatica E4), desktop application (E4 Connect), binder with instructions and fill-out sheets based on the Day Reconstruction Method. The kit is seen in figure 3.d. In preparation of the interviews, we asked participants to bring a laptop, where we could help them install the software during the interview. Two home kits where created for this field study to be able to conduct two studies in parallel.

Figure 3.d Home kit consisting of: Empatica E4 wristband, desktop application E4 Connect and a binder with instruction and fill-out sheets

3.5 Procedure

All four participants followed the same procedure, which consisted of:

1. Recruitment
2. Initial interview
3. Wearing wristband at home for one-two weeks
4. Concluding interview

Recruitment

Participants were either recruited through the psychiatric ward in Aalborg or through a Facebook group for recruiting participants. The contact with the psychiatric ward was established through a chief psychiatrist at Aalborg University Hospital. The participants recruited from the psychiatric ward were asked by their individual psychiatrist if they wanted to participate. The participants received information on the study prior to meeting, which can be seen in Appendix 3. Then a meeting was set up with the researchers and the participant.

Interviews

Interviews were conducted as semi-structured and by the two researchers. One researcher being the the primary interviewer, responsible for following the interview guide, and the other researcher being the secondary interviewer, responsible for: taking notes, recording the interview and ask follow up questions. All interviews were recorded and transcribed, to see

the transcripts see Appendix 4.a-d. In the beginning of the interview the participants were asked to fill out a consent form, seen in Appendix 5. For the initial interview there were also time to introduce the system and trying the wristband for five minutes before looking at the data together: patient and researchers. This was done to make sure the participant was on board with using the system and did not have any immediate questions to the use. If there were questions later in the study the participant was instructed in looking in the binder for the manual or calling the researchers.

Location

The interviews were conducted at the place the participant would find most convenient or safe. With the two participants recruited through the psychiatric ward, the interviews were conducted on the location of the psychiatry. For the other participants, they were interviewed either at home or on university property.

Duration

The length of the participation was: 2 days, 10 days, 14 days and 29 days for each participant. The actual use of the system was 2 days, 3 days, 7 days and 27 days. The researchers had two Empatica E4 wristbands available therefore only two studies could be conducted simultaneously.

3.6 Data Analysis

Qualitative data in the form of interview transcripts make up the core of this analysis, which was analysed inductively. Inductive analysis or open ended coding is beneficial when the field of research is emerging because no theory has been formed yet. Inductive analysis and open ended coding derives from grounded theory. Grounded theory's most important asset is its inductive nature, which means the knowledge that comes out of inductive coding does not come from predefined concepts, but rather concepts deriving directly from within the data. Grounded theory contrasts with experimental research, where a hypothesis is tested and either accepted or declined. With grounded theory, the process of gaining knowledge is based on being able to organize data into theory of the unknown (Lazar et al., 2010). According to Lazar, doing grounded theory has four stages: Open coding, development of concepts, grouping concepts into categories and formation of a theory. The procedure of the data analysis for this study is described here:

- Transcribe interviews
- Read transcripts through
- Open coding
- Finding themes and patterns
- Categorizing themes into concepts

The procedure followed the one Lazar (2010) suggests. Lazar suggests that when doing open coding the material is read through before coding, as it gives the researcher the opportunity to get an overview of the qualitative data. This was done in this study, as the researchers conducted the interview and transcribed the material, and read the transcripts through before coding. After having read the data, the open coding began. The coding was conducted individually, where both researchers coded all transcripts by marking interesting quotes in the text and creating themes and patterns in a separate document. The interesting quotes and findings were in the initial interviews focused on the participant and their anxiety treatment as well as their relation to technology. The findings from the concluding interviews focused on the perceived value of the use and the experience with the technology. After the individual coding, the researchers got together and categorized the themes into concepts.

Since the interviews conducted happened over a long period of time (the first was held on January, and the last in March) the researchers had to start the analysis process while not having all data available. This meant that when it was time to go through the third and fourth participant, preconceptions had influenced how the data from the third and fourth participants was perceived. This is an inevitable struggle when conducting qualitative data analysis, but the same procedure was followed in analysing the data, thus trying to keep an open mind towards the data was upheld.

Another part of the data analysis was looking at the data from the wristband and relating it to the notes made by the participants, to see if there was a match between the fill-out sheets and the E4 Connect data. This was done to further support the claims of the participants of how the system had worked for them. In the following chapter, the result of the field study will be presented by describing the concepts and relate them to quotes and screenshots from the participants, but first a description of the involvement of psychiatrists.

3.7 Involving Psychiatrists

The importance of incorporating clinical professionals when studying technology and mental health is highlighted in both Garcia-Ceja et al. (2018) and Doherty et al. (2010), as clinical professionals might have other reasonable and genuine concerns regarding the introduction of technology in treatment in terms of potential pitfalls than HCI professionals. In this project the involvement of clinical professionals consisted of a meeting with a Professional Forum at the psychiatric ward, and two individual interviews conducted with psychiatrists. In the following two sections the method used for the Professional Forum meeting and the individual interviews will be described.

Meeting at Professional Forum

A meeting was held at the Professional Forum at the psychiatric ward at Aalborg University Hospital. The meeting consisted of a review of the project and focused on the technology and the field study results so far. Firstly, we presented the motivation behind the project, the system we used and how we imagined it could be used as part of the current treatment. We

presented some of the results we had already received from the first two participants in the field study. Afterwards, we asked the psychiatrists to comment, discuss and ask questions about the system and the potential use in treatment. The presentation and discussion was not audio recorded, but notes were taken throughout. The meeting yielded many insights that could be taken into consideration, when designing for anxiety patients.

Individual Interviews with Psychiatrists

Two interviews were conducted with two psychiatrists who were Caroline and Mathilda's psychiatrists respectively, and have been given the names of Camilla (Caroline's psychiatrist) and Anna (Mathilda's psychiatrist) which are not their real names. The purpose of the interviews were to learn how the psychiatrists viewed the technology in terms of potentials and concerns and gain knowledge on current CBT treatment forms. The interviews were conducted separately in the psychiatrist's office and present were the participant and three interviewers, one primary, one secondary as well as one who took notes. The interviews were semi-structured and an interview guide was developed based on the results of the field study and the meeting at the Professional Forum, seen in Appendix 6. The interviews were transcribed and analysed the same way as the interviews from the field study, as presented in section 3.7: Data Analysis. The transcripts are presented in Appendix 7.a and 7.b. The results of the meeting at the Professional Forum and the two interviews are presented in chapter 5: Results of Interview with Psychiatrists.

3.8 Summary

This chapter described the study method of how this project was conducted. The empiric data that lays the foundation of this study is based on: 1) Longitudinal field study with four anxiety patients, 2) Meeting at Professional Forum and 3) Interview with two psychiatrists. The study method follows field study and longitudinal methodology as well as a CRD and CBT approach. The data from the field study and interviews with psychiatrists was inductively analysed using open coding. The results of this study is presented in the following chapters: Chapter 4: Field Study Results and chapter 5: Results of Interview with Psychiatrists.

4. Field Study Results

In this chapter, the results from the field study will be presented. The results consists of concepts that emerged during the analysis and will be presented along quotes from the participants. During the analysis of the interviews, 27 themes emerged, both in relation to the physical design of the wristband and the overall use of the system, a list with all themes can be found in Appendix 8. The emerging themes are sorted into the following six concepts:

- Confirmation of Episodes
- Understanding of Oneself and Patterns
- Usability and User Experience
- Use in treatment
- Concerns and Drawbacks
- Design Considerations

These six concepts will be described in the following and corresponding quotes from the participants will underline the findings.

4.1 Confirmation of Episodes

During the study it was learned that the E4 wristband could assist in getting confirmation of anxiety episodes. Besides getting confirmation of episodes, the data can also assist in remembering them as this can sometimes be difficult.

For instance, Mathilda was able to get confirmation from the data on how she had experienced her anxiety and found that all the data fit with how she had felt it. She used it by first writing her episodes down and then checking the data afterwards, to see how they matched her experience. She said it was nice to get the confirmation from the data that something was actually happening in the body and not only in her mind, but she also explained how it was a bit terrifying at the same time as the data made the episodes more real. In a quote she states:

"... (It is)Both terrifying and confirmative at the same time ... When it is written on paper in front of you it becomes more real. And when you then get such numbers, and if you like numbers, it becomes even more real, right?" - Mathilda

Another example is Caroline, who is sometimes unsure whether she has an anxiety episode or not. She found that the wristband could potentially be useful in knowing when she did experience anxiety.

"... sometimes I can be unsure whether it is it (anxiety), and so it might be a help, I think"
- Caroline

Caroline only wore the wristband for a day and a half which is why the statement is more a speculation, but during the time she wore the wristband, she experienced one anxiety episode which showed on the graph. For this one episode, Caroline said: *"In a way it was a little nice to get it confirmed"*.

Remembering episodes
To the initial interview, three of the participants said it could be difficult to remember all their anxiety episodes and details about them. After the study period, Maya said the system helped her remember some episodes:

"... there is definitely two of the times, no three of the times I believe, where I think it was great to see that one like, then you could just think, oh yes, that was then" - Maya

Caroline also said the wristband could help her remember her episodes, as she was sometimes unsure why she felt the way she felt.

4.2 Understanding of Oneself and Patterns
This concept relates to how the system can help the user get an understanding of themselves, their anxiety and discover patterns in their anxiety. Mathilda, who used it for a month, was able to learn something about herself and could start to see patterns. In relation to getting an understanding of herself and how she felt, Mathilda said:

"... it is actually pretty good for when you are about to spaz out over why you are so tired or why you don't want to do anything when you get home, and then you think back ... and then I can actually see, it's not only right there and there, it is actually also some time after or, there is a reason why I am just used up" - Mathilda

The data also helped her be aware of the physical change that happened in her body:

"It's pretty nice to know. It may be that you are a little weird but not completely weird. It is real enough that something happens in the body" - Mathilda

During the extended period, Mathilda was able to start seeing certain patterns in the data, such as a significant difference between her anxiety both during the day and night when she was working and when she had time off.

"... I think some patterns are becoming more clear, and I think that's very interesting and something to dive into ... I think now it's been a month so it's no longer just a day picture or a momentary picture, right?" - Mathilda

She knew beforehand that there were a difference in her anxiety when working and when not, but she was not aware just how big a difference there actually was. This became clear after wearing the wristband and looking at the data. She could see the patterns both during the day and the night. In figure 4.a and 4.b, two screenshots from Mathilda's data are shown. The top screenshot in figure 4.a shows a day at work, and the bottom one, figure 4.b, shows a day during her holiday, where she was off work. Though there are still some spikes when she is not at work, it is clear that there is a significant difference in the GSR data (the top blue graph) when at work and when she has time off. However, it should also be noted, that the accelerometer data (the bottom purple graph) also shows that there are more movement during the day when at work, so some of the spikes might be due to that rather than anxiety.

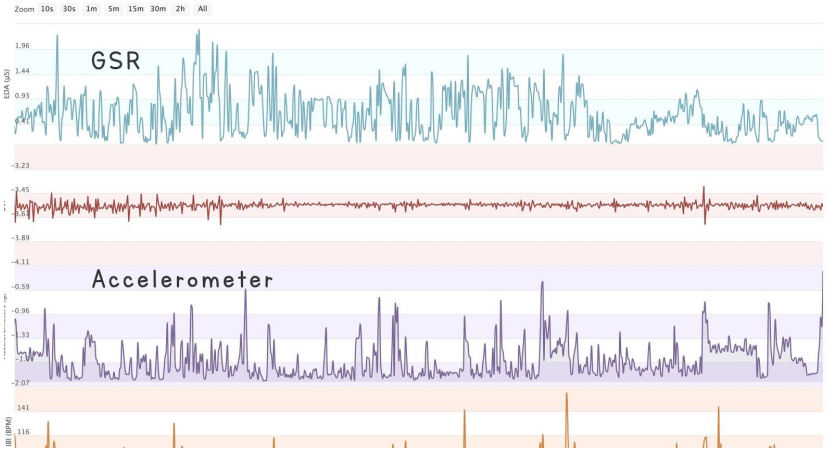


Figure 4.a Shows the GSR data, in the top blue graph, for Mathilda on a day where she is working

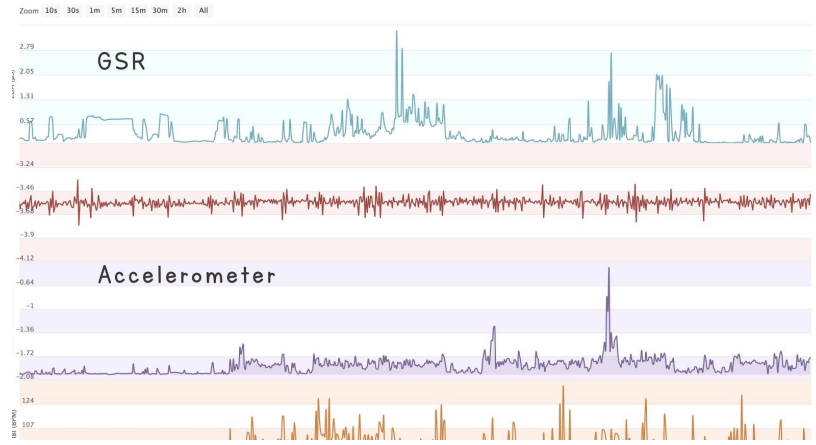


Figure 4.b Shows the GSR data, in the top blue graph for Mathilda during her holiday, where she was off from work

4.3 Usability and User Experience

The themes for this concept span: the ease of use of the system, the overall user experience and the trust in the system as a result of the use.

Ease of Use

Overall, all participants found the system easy to use. For example, Maya said: *"Well I think it was easy to use … I don't think it was complicated or anything".*

However, some participants found the graph confusing to read. For example, Eric said he would not have been able to understand it, if we had not explained it beforehand, but after being told what to look for, it was fine. Caroline also stated:

"I think it was a bit confusing ... just like, when it goes up, how much is then a lot and when is it just like normal?" - Caroline

Mathilda had also discussed the graphs with some colleagues, who had the same concerns as Caroline, they also wanted to know when the graph was 'normal'. However, Mathilda herself had no problem reading and understanding the graphs. She was able to differentiate between different types of situations and anxiety and furthermore, see patterns in the data. Maya was also able to read and understand the data.

Being able to differentiate between spikes on the graph as a result of anxiety or something else is highly relevant in terms of understanding the data. All participants were able to tell when the spikes were due to physical activity rather than anxiety. This was mostly based on the time of the spike, as they knew they were working out at that particular moment. For instance, Caroline said: *"... I knew at that time I had been working out, so I knew it was that."*

User Experience

Two of the participants found the system interesting and fun to use. For example, Maya said: *"... then you look forward a little to, you just have to look, what does it say today".*

However, Maya also stated that, for the duration of the study she found the system intriguing but if she were to use it for a longer time, she might find it boring. Mathilda however, who used the system for a full month, found it very interesting to use, but did mention that she might find it boring at some point when there was nothing new to learn about herself through the system:*"Well, I think it would be exciting as long as I am insecure if I am feeling it right."*

Neither Caroline nor Eric found the system fun to use. Caroline did not find it fun to use as she had some concerns about the constant monitoring and privacy, elaborated in section 4.5. Eric experienced a lot of technical issues which made him very frustrated towards the system. However, he said if uploading was less frustrating it might be fine to use the system.

Trust in the system

Whether the participants trusted the system was mainly based on how well the data fit with their own experiences. For three of the participants, Maya, Caroline and Mathilda, the GSR fit with how they felt, so they indicated that they trusted the system. Maya had looked at the other graphs as well and found the HR to be very high. She had tried to measure herself to compare the data and she did not agree with what the E4 wristband measured. She did however still trust the GSR data:

"Well I will say I trust that GSR at least, not the pulse, the temperature I also think sounds very realistic" - Maya

Maya experienced one time during the period that the graph did not spike even though she had felt an anxiety episode. For that episode, she trusted herself more than the system and said that even though the graph did not spike she knew she had had an episode, the GSR might just not have been affected at that time. This experience did not change the fact that she still overall felt she trusted the GSR data.

Mathilda had a lot of trust in the system. All the data fit, as she said: *"annoyingly well"*. The reason she said it was annoying was that the data had confirmed some concerns the psychiatrist had regarding Mathilda, which Mathilda had tried to convince the psychiatrist otherwise. To this, Mathilda said:

"... and then I think it is pretty interesting that I try really hard to convince myself that this work is so healthy for me ... and the psychiatrist tries really hard to convince me that it is not, and it is a little annoying that I have now borrowed a device that does not speak my case" - Mathilda

An example of a night and day in Mathilda's data is shown in figure 4.c, the measurements are from 23 o'clock until 18 o'clock the following day. Timestamps and notes have been added to the screenshot, and is shown as a blue overlay. During this day, Mathilda wrote down four episodes. The first episode was during the night which she noted as a giant nightmare, and the following three were all at work: A conversation with a parent, a conflict with a student and a meeting with the boss. There are spikes on the GSR graph for all the episodes she wrote down. The graph also shows some spikes that she did not write in the fill-out sheets, but these could potentially be due to something other than anxiety, or could be episodes she did not realise she was having. This fits well with how she said she was trying to convince the psychiatrist that work was good for her, but the graph showed how it was actually not as good as she thought.

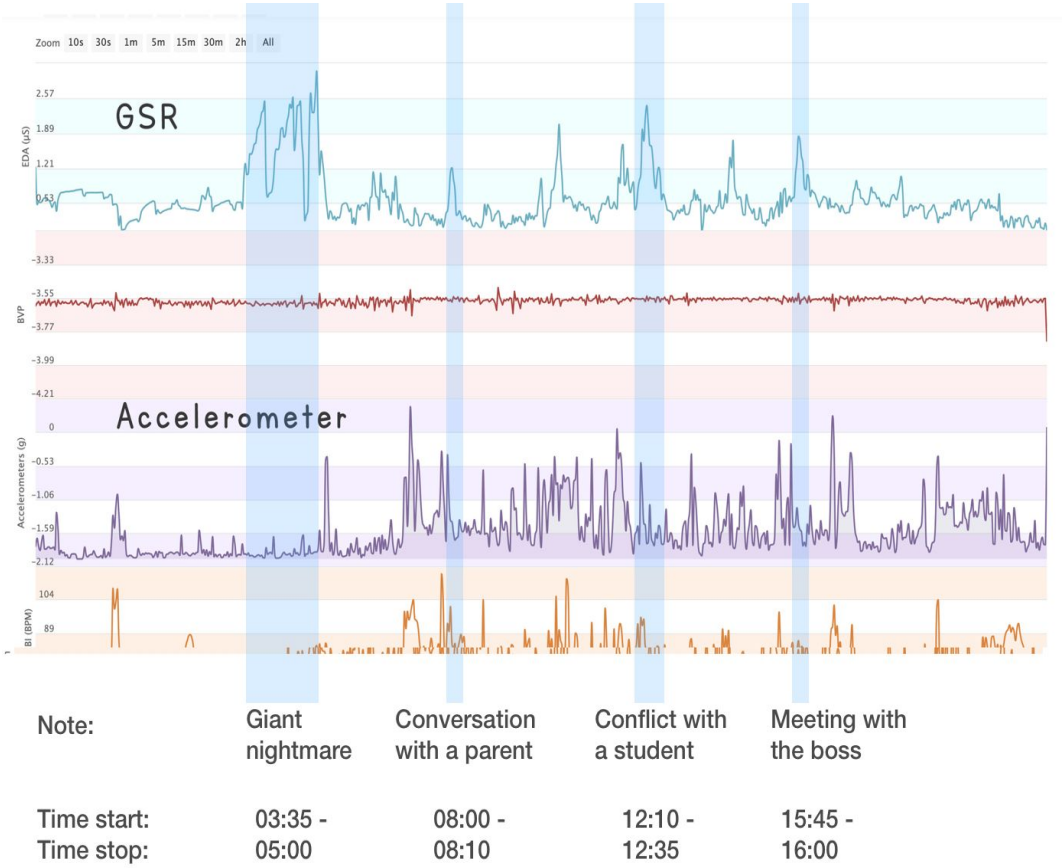


Figure 4.c Mathilda's day from 23:00 to 18:00 the next day

Eric did not have a definite anxiety episode during the study period but had felt he had an increased level of anxiety at some times that week, however, this did not show on the graph. He could see that the GSR spiked when he was working out, so he trusted that it measured his sweat level. He said:

"... I can see it measured sweat so I trust that it did that, but if it is because I don't sweat when I have anxiety episodes or I don't sweat until later. I don't know. In any case, it did not measure it." - Eric

4.4 Use in Treatment

One of the participants, Mathilda, used the system as part of her treatment by bringing the data to her psychiatrist and discussing it during one of her sessions. This was interesting as the data showed something about some of the things they had already discussed:

"... some of it I think we had talked about, (psychiatrist) and I, was becoming better, maybe some of it, which was completely gone, that maybe it isn't, it has become better, but maybe it is actually still unconsciously in some contexts without us really completely (knowing it)"

- Mathilda

As described in the previous section, the system had also confirmed some concerns of the psychiatrist in relation to the patient's work effecting her anxiety, which they had discussed. In regard to this Mathilda said:

"... we have talked about that unfortunately I can no longer find a counterargument against that I'm right, and she thought that was great" - Mathilda

Use of fill-out sheets

The participants were given fill-out sheets to complete during the study. This was to see whether the GSR data would be helpful in filling out CBT homework. Mathilda was the only one to fill out the sheets throughout the study duration. She liked to write it down and then get confirmation from the data:

"But it is the funniest when you can write it down on paper and then get the confirmation, I think, I believe if I should give a piece of advice to others I think it would be to have it written down by yourself so you can see, well I definitely think that there is a difference in what I believe is just normal human behaviour or yes, and what is anxiety" - Mathilda

Maya used only one of the sheets but then continued writing anxiety episodes down on her phone instead, as she found it more convenient. She did it the other way around as Mathilda by looking at the data first and then writing it down.

"I have connected it first and then I have looked at it and then I have filled out afterwards, but I have not filled that many out in that one (the fill-out sheets), because it became a little strange to write it there, so I wrote it in my phone instead ... I thought that was a bit easier" - Maya

Maya did fill out two days on the sheets we provided her, and we looked at the graph and her notes to see how they fit together. A graph for a day she noted three anxiety episodes is shown in figure 4.d, shown from 17.45-21.30 o'clock. She wrote down an anxiety episode for each of the two spikes marked with a blue overlay on the graph, and wrote that all throughout the evening she had to control her anxiety. The graph clearly shows the two spikes, and during the evening (approximately from the second spike and the rest of the graph), the graph likewise shows spikes. The small spikes throughout the evening clearly shows compared to when she is not having anxiety, which can be seen in figure 4.e. The red lines shows when she pressed the button on the wristband.

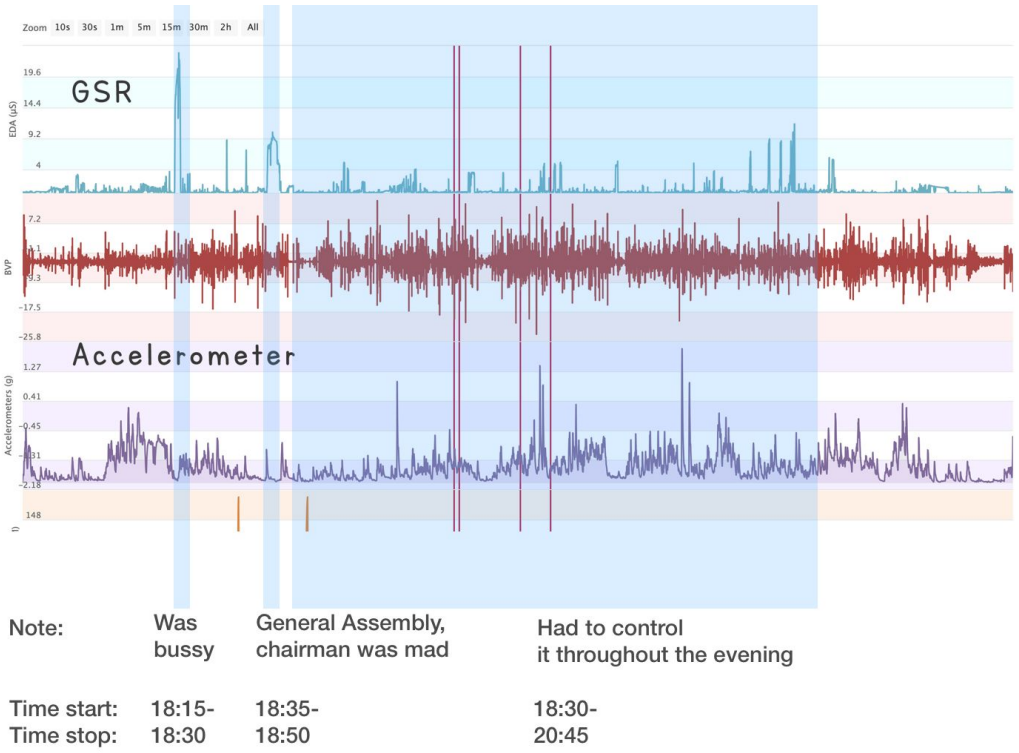


Figure 4.d Maya having three anxiety episodes in a day. Data from 17:45 - 21:30.

In figure 4.e, another day from Maya's data is shown. She noted an anxiety episode while she was biking, which shows clearly on the graph. The screenshot shows her day from approximately 10.30-17.30 o'clock.

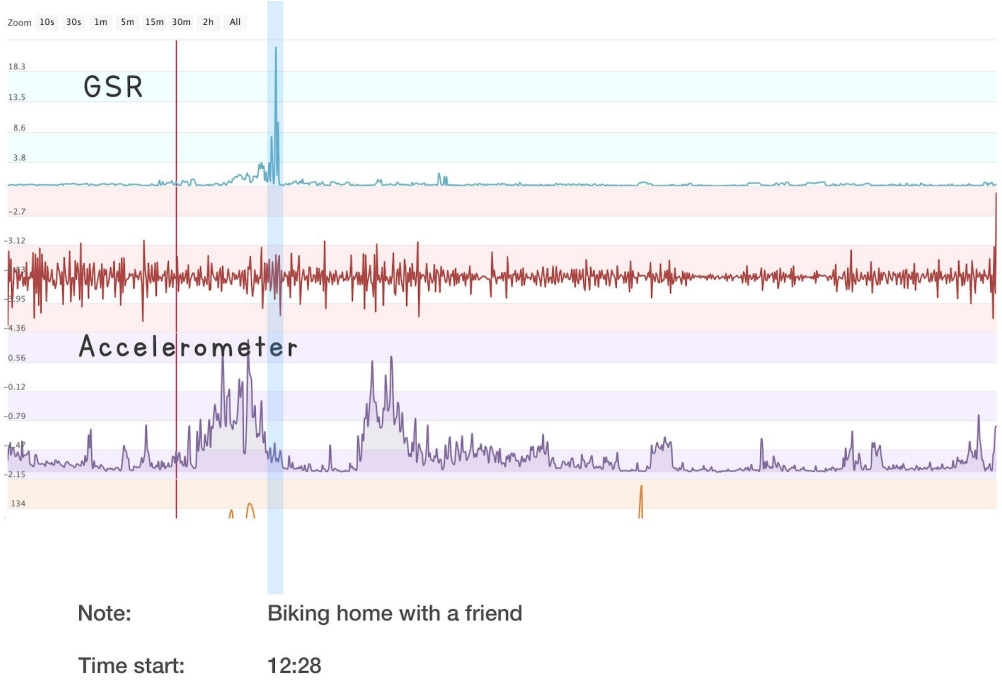


Figure 4.e Maya having anxiety while biking. Data from 10:30-17:30

Caroline was uncomfortable writing anxiety episodes down on the fill-out sheets, and Eric did not think he had the need to write it down as he had not had a definite anxiety episode. Eric had also said to the initial interview that he never wrote anything down in regards to his anxiety.

Use in periods

Two participants, Mathilda and Maya, indicated that there were certain periods of time where it would make more sense to use the system than others. Both certain periods in a therapy program and in life in general. Maya said she would more likely use such a system in vulnerable periods of time, because when things were going well she did not want the system to tell her otherwise:

"... I think maybe I would say in those periods where I feel more vulnerable I think I would use it and then those where I don't, except for during the exams period, then I think I would say it might be better to not do it." - Maya

Maya did not want to use the system in the exam period as she already knew she would be right on the edge all the time during that period and did not want anymore stress added. Maya also said the system might be useful as part of the therapy program she had been in a few years ago. Mathilda said a system like this might be useful in the beginning of a therapy program:

"... If this had been in the beginning, I think it would have made much more sense. At that time there were almost constantly (episodes). Now we are down to where it can fit on one sheet of A4 paper" - Mathilda

Mathilda also stated how the system could be useful in periods where it could be difficult to differentiate between her emotions in relation to what was happening in her life:

"... I could imagine that if you were in a really stressed period and you were not able to separate the things, like separate: when is it because you are really stressed at work, or when is it because the excessive thoughts just take over or when do you have too much in the calendar, or whatever it might be. Like if you could not separate the things, I think it might make good sense because, a lot of the things I have written down has something to do with sleeping. So right now I think my main problem should be found in my sleeping pattern. And maybe you would figure out it was mostly when you were with people or mostly when you were alone or yes, mostly at work, or mostly in your spare time." - Mathilda

The screenshot in figure 4.f shows a night during the period from approximately 22.30 until 9.00 the next morning from Mathilda's data. Mathilda wrote in the fill out sheets that she was not able to sleep due to her thoughts from between 4 and 7 during the night. This is shown in

the screenshot where the blue overlay is and the GSR spikes are clearly seen during this time period.

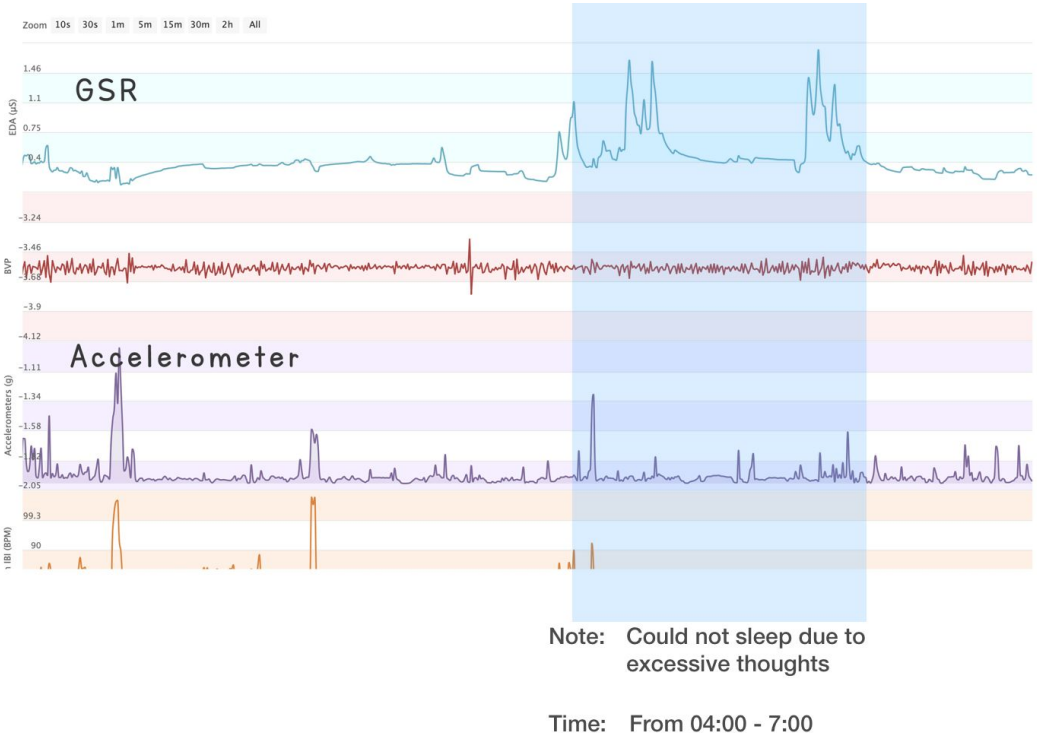


Figure 4.f Mathilda's night from 22:30 - 9:00.

4.5 Concerns and Drawbacks

There were a few concerns regarding the use of the system during the study. These were related to the constant monitoring and privacy.

Monitoring and Privacy

Caroline was the first to use the wristband and after a day and a half she felt she was being monitored too much by the wristband which she found too uncomfortable to continue:

"Well I have only worn it for a day and a half because I felt like I was being monitored and was very aware that it took measurements constantly, and I thought that was uncomfortable"
- Caroline

Concerns were also related to privacy in terms of who has access to the users' data. Two participants had some concerns regarding this. Caroline did not like that anyone could see her stuff, and also stated that she might feel differently towards the whole system if she knew, she was the only one who had access to it:

"I don't know if it had been different if all this was just me who could see everything, then I might have felt differently towards it" - Caroline

Mathilda was not concerned if others could see her data because they only made sense to her and someone who were trying to interpret them. However, when asked if she would mind her psychiatrist seeing her data, she said:

"Then I think I would like to have some kind of feature that could show that now someone else has been following along ... I would like to know who feels entertained by my data, if I knew it, it would not bother me." - Mathilda

The other two participants, Maya and Eric, were not concerned about neither privacy nor monitoring.

4.6 Design Considerations

There were several comments regarding the physical design of the E4 wristband, mainly related to the size and visibility of it as well as the two GSR sensors on the bottom of the wristband being irritating for some participants and left marks after a day of wearing it. Because of the size of the wristband it was uncomfortable to wear for some participants. For instance Maya explained how it felt like the wristband locked the bones in the wrist when she did sideways movements. All four participants found the wristband big and clumsy.

"... it is a litte big, I think, that is probably what I think is most, was most annoying" - Maya

"It's a little annoying when it is so big and clumsy" - Eric

The problem with the size was not only related to the uncomfortability of wearing it, but also that it drew unwanted attention to itself. For example, Caroline said that people kept asking what it was, which was uncomfortable to answer:

"... I think it is hard to hide it, so I was asked a lot what it was and I don't think that was very awesome to have to answer either" - Caroline

Several design considerations emerged during the interviews, both in terms of the physical design of the wristband and the future use of the system. In terms of physical design all participants would prefer a smaller device that was less visible, or include a clock in the wristband so it looked more like a regular watch. Furthermore, two of the participants mentioned that they would prefer the wristband to be waterproof.

In relation to the future use there were some thoughts regarding computer vs mobile and alerts.

Computer vs. Mobile

When asked if they wanted the system on a computer or a mobile application, two of the participants, Eric and Mathilda, said it should be available on both, however, Eric said he would prefer to use the mobile because of its mobility. Maya was certain that she preferred it in a mobile app so it would become a part of her day rather than sitting down with her computer and filling it out:

"... when you are sitting and waiting and just have two minutes then it is not your computer you pull out, then it is your phone you just like, so it does not become something like 'now I have to do this', so it becomes a part of the day." - Maya

We asked her whether she had any concerns in terms of strangers following along if she used the app on the bus for example, and she said it did not bother her at all.

Personal Consideration

Another important feature that Mathilda mentioned, was the fact that the system should support customizability. There might be difference in understanding the graph depending on you preconceptions. At one point she stated:

" ... If you somehow could make it personal, I imagine that there is a great difference between people when it comes to what numbers come out, but the fluctuations are probably there anyway." - Mathilda

Alerts (Just-in-time interventions)

We asked the participants whether they would like the system to alert them if it had detected an increase in GSR. Alerts could both be to warn the user or to help alleviate before the episode happened. Both Maya and Mathilda were a bit sceptical towards alerts. For instance, Mathilda said:

"... if you were warned ...it would be smart, but that would require you knowing what you were warned about, ... and I definitely think, I don't always know what triggers it (anxiety) right here and now" - Mathilda

Mathilda had a suggestion for how it could be done rather than warning. She suggested that the system could make a countdown to when the episode would be over, as she knew they could never last for more than 15 minutes, which could feel like a long time when experiencing anxiety:

"... if you could do it somehow personal,... if you had a watch or device that knew your rhythm, if it could say: "Okay, now it has started", and then you know, ... you can maximum be very anxious for 15 minutes, then it goes down, ... because 15 minutes can feel tremendously long, and sometimes you can have the feeling several hours pass by, so if your

watch could tell you: "Yes, yes, we can see you're anxious, but in 10 minutes then,... then it starts to go down"... that could really be smart, I would like that" - Mathilda

In relation to alerts, Maya said it depended on what type of episode it was:

"Well I believe, there should be a limit then, because if it (the anxiety) just suddenly says 'poom' (shows it increases quickly), then I don't think it should alert because then I think it will just make the stress worse ... then you are so much in yourself and in your, that there is not anything else that should come and go 'pling pling pling' ... but if it is slowly increasing then it might say something in the beginning but then it should stop to not make bad worse"
- Maya

4.7 Summary

The results of the field study indicates a potential for using this type of technology for keeping track of anxiety episodes. It can especially be helpful in remembering episodes and getting these confirmed. Furthermore, using the technology for a longer period of time might help the user get an understanding of themselves and their anxiety, and as a part of this also help uncover patterns. There are however also some concerns and design recommendations that needs to be addressed for the technology to be fully accepted by the users. These include concerns in relation to constant monitoring and privacy. To make the system more convenient for the users it might be beneficial to also have the data available from a mobile application. Furthermore, a redesign of the wristband will be meaningful as all participants found it large and clumsy.

As the system might be effective as a part of the current treatment form for anxiety it is relevant to discuss the potential and the concerns with psychiatrists as well. Therefore, we conducted two interviews with two different psychiatrists. The results from these interviews will be presented in the following chapter.

5. Results of Interviews with Psychiatrists

As stated by both Garcia-Ceja (2018) and Doherty (2010), it is crucial to also involve clinical professionals when studying and designing these type of systems. Therefore, we presented the project at a meeting called Professional Forum with a group of psychiatrists to discuss potentials and concerns. Furthermore, we conducted two interviews with two psychiatrists to get an understanding of their view on the use of the technology. The psychiatrists were Mathilda and Caroline's psychiatrists, and have been given the names of Camilla (Caroline's psychiatrist) and Anna (Mathilda's psychiatrist). During the analysis of the interviews, 22 themes emerged which could be grouped into four concepts. These will be presented in the following section along with corresponding quotes from the participating psychiatrists. The quotes will all be from the individual interviews, as the Professional Forum discussion was not recorded and transcribed.

The four concepts are:

- Individual Assessment of Patients
- Potential of Multiple Use Cases
- Current Treatment
- Concerns towards Technology Abuse

5.1 Individual Assessment of Patients

A recurring theme during both the interviews and the Professional Forum discussion were how the patients differed from each other making it difficult to point to one particular user group or use case for the technology. Both diagnose and personality traits would affect how suitable the patient would be for finding the technology useful. Both psychiatrists in the individual interviews stated that they would assess each patient individually to determine who could gain something from using the technology. Some of the traits mentioned, that might affect how suitable the patient would be, were: independence, rumination and perfectionism. For instance Camilla said:

"... there is also again something individual about patients you might consider ... like how independent you are in your thinking and in general." - Camilla

Both psychiatrists said that patients who ruminate a lot might not be suitable for the technology as they have a hard time differentiating between what is relevant, so the increased focus on arousal might have a negative effect. For instance, Camilla said:

"... I think maybe that (rumination) can also be connected to feeling monitored, that you think like someone is following along all the time and everything is relevant and you lose the

focus in it, what is the focus, what should I pay attention to or what will there be payed attention to when we look at it." - Camilla

Anna also said:

"I think those with panic disorder, that would be great. Because these patients with generalised anxiety have a lot of worries about everything" - Anna

Camilla also mentioned patients with panic disorders as a potential user group as the anxiety attacks seem to happen out of nowhere. However, she was unsure whether it was necessary to determine the trigger, which is also related to the next concept: Potential of Multiple Use Cases, which will be elaborated in the following section. Another diagnose that was brought up was Obsessive Compulsive Disorder (OCD) because of the perfectionism, Anna said:

"A lot of patients are actually a good group because they are very perfectionistic. If you press the right button they will make an effort to be good patients who completes all their homework." - Anna

Anna, however, also stated that there were also negative sides to the perfectionism that might affect the use:

"So it is just how you use it, also perfectionism in the good way, you can get a lot out of it, or you focus again on these negative sides of perfectionism which are also there. So it is a bit of an individual assessment" - Anna

Both psychiatrists agreed that it was up to them to assess how suitable the patients are, so they also found it important that they knew the patients a bit before introducing the technology. Anna was certain the technology could be helpful for a lot of patients: *"Yes, it makes sense to many. No doubt about it"*.

5.2 Potential of Multiple Use Cases

Also related to the individual assessment of patients were the different potential use cases for the technology as it might serve different purposes depending on the patient and where they were in their treatment. The use cases that were brought up were:

- Disconfirmation of episodes
- Connection to the body
- Prevention

Other considerations under this concept were the psychiatrists' role in it, when to introduce it in the current treatment program and the connection between the technology and the current

fill-out sheets used. This theme was both apparent during the Professional Forum and the individual interviews.

Disconfirmation of episodes

As it was learned during the field study, the system could help in confirming anxiety episodes, however, both psychiatrists also mentioned how the system could help disconfirm episodes. The argument for disconfirming episodes were targeted some anxiety patients who feel they are anxious constantly. For instance Anna said:

"I also think monitoring is important. Especially from the beginning so you can see how much space it takes, also on the other side some say "I feel bad"... all the time. So it can also be a way to disconfirm this" - Anna

Camilla also mentioned these patients who feel anxiety all the time: *"... these who feel anxiety constantly, also a little good to use it as potential disconfirmation"*.

Connection to the body

Camilla stated how it could be difficult for patients to establish a connection between what was physiologically happening in the body and what was happening in the mind. She said that some patients cut off completely in terms of acknowledging the body's signals:

"And then of course there is the connection to the body which is often really bad, ... like really, some can, is completely cut-off (shows cut-off from the neck and down) and doesn't want to relate to anything at all" - Camilla

She said how the technology could potentially help re-establish this connection, as understanding the body's signals was crucial: *"But to learn your, the body's signals, I think is primary for understanding it better"*. Furthermore, Camilla stated how being aware of what was happening in the body in certain situations could help understand what was happening if you were in the same situation again.

"... but it could also be you could, if you are in the situation again, where you have been aware that here something is increasing, what is it then that happens next time I am in such a situation, can I be more aware of what is happening in the body" - Camilla

Even though it could potentially be helpful to learn to connect more with the body, it once again comes down to an individual assessment of the patient as the increased focus on the body might also be negative for some patients, Camilla said.

Prevention

Anna focused a lot on how the technology potentially could help prevent anxiety from happening by making the patients determine why it was happening. Camilla was not

completely convinced that it was helpful for patients to learn what and why their anxiety was triggered, but rather learn how to handle it when it did. Anna said:

"... that we could identify these triggering factors and prevent and make a crisis plan or a plan of action … This tool (the presented system) could help with that" - Anna

In contrast, Camilla said:

"Yes, but it is probably also because we are a little like … back and forth on should we figure out where and when and what caused it, or should we actually learn to handle it when it comes and get ahead?" - Camilla

Introduction in the Current Treatment

When to introduce the system in the current treatment would once again depend on the patient and the purpose. Anna suggested a good time to introduce the system would be about ¼ way into the treatment, as it would not make sense neither in the beginning nor in the end:

"... a fourth part in the program or something like that. ¼ part of the program where you get to know the patient a little better. That would be my suggestion based on the experience I have" - Anna

Camilla said it could be interesting to introduce it at different times during treatment depending on the purpose:

"... it can be used differently in the programs, and it might also be a little interesting when in the program you used is, also maybe depending on the purpose of it." - Camilla

Prior to discussing the potential of using the technology for anxiety treatment with psychiatrists, we considered how the technology could be used before starting the actual treatment program, as there can be a long waiting time before a patient receives treatment. We presented this at the Professional Forum, where several of the psychiatrists raised concern regarding this. As described previously, the use of the system as part of treatment, requires an individual assessment of each patient, which is why they did not think it would be a good idea to introduce the technology before starting treatment. They were concerned it might make the patients more nervous and have the opposite effect.

Psychiatrist's role

We asked the psychiatrists whether they wanted to have access to the data as well. Anna said that seeing the data might make sense to her as well as the patient:

"It would also help me I think, to also maybe again connect data to symptoms. Yes that would make sense" - Anna

However, she also mentioned that it would be better if the patient brought the data and they discussed it, like they also do with fill-out sheets, instead of the psychiatrist looking at the data alone:

"You could also say that patients could bring it and we look at it together. Or else it becomes very like keeping an eye on what he is doing" - Anna

Camilla also agreed that it should be the patients who should see the data, as she was also concerned the patients might be affected by knowing their psychiatrist could see the data:

"I think it should be the patients themselves who could see it (the data), also to strengthen their agent in it, ... and they should also take responsibility in the treatment themselves and be curious, like it is really important it is them who drives it or else it falls to the ground." - Camilla

5.3 Current Treatment

To better understand how the technology would fit in, we asked the psychiatrists about the current treatment form they used and what the biggest challenges and successes with this form was. They mentioned that they used Cognitive Behavioural Therapy (CBT) but with variations depending on the needs of each patient. Most sessions are conducted in a group with other anxiety patients as the experiences they can learn from each other is key in CBT:

"Yes, successes is also in our group, where they meet each other and can mirror themselves in each other and see that it's normal and yes, learn from each others' experiences, ... it is really really good experience and successful when they experience they can do the things regardless, and even if it is uncomfortable, they can complete it" - Camilla

Another important element of the current CBT treatment is homework, which the patients are assigned to complete between sessions. Homework could for instance be to complete fill-out sheets about what has happened. This can be really helpful, but again depends of the individual patient:

"... It is probably a little different how much you use the schemas, I think there are a lot of patients who likes to have something written down, and we also see it in the groups, that they have a great advantage in getting some schemas and something, and just getting it down on paper might mean you become more aware of some things" - Camilla

In relation to writing things down, Camilla also said:

"... it could be something like continuously writing down during the week, what have you experienced, how have you tried to handle it? How did it work? Consequences in the short and long run for example, it can also sometimes be difficult to write down all the time, or it

becomes a little against the intention, then you get an enormously big focus on the anxiety maybe"- Camilla

There were also some challenges in terms of writing down, as the patient could forget to write it down and then try to remember and write everything down at once, which could be affected by their current situation:

"... I think it (writing down) is a challenge, and it will always be colored by the situation you a probably sitting in, are you sat with a really bad feeling ... then you might remember a lot of the bad episodes, so such reporting is like, it can be doubtful" - Camilla

In general, it can be difficult for the patients to remember their episodes or details about it. For instance, Anna said:

"When I try to ask they are all confused and say: "No, I can't remember it, I was sitting on the sofa" I say: "Were you sitting on the sofa, what happened before?" They can't connect it to anything at all. There is actually a decent group who says that" - Anna

Camilla also mentioned how the patients' memory was affected and it could be difficult for them to get a hold of different elements in relation to their episodes:

"I think it is often hard (to remember episodes), because they don't remember that well all the time, like their memory and concentration are affected and, it also depends on each patient what it is they can easily get a hold of ... some have a really hard time getting hold of the thought content and some have a really hard time getting hold of what is happening in the body, and the feelings are always really difficult, besides anxiety" - Camilla

She continued explaining, how they, the psychiatrists, were not always that good at telling the patients to write it down, and then when the patients came for their sessions they had to sit there and try to recall, which could be really challenging:

"And it's not certain that we are always too good at ... telling them they can do it (write down), and then everytime you come here you have to try and remember, and then you are sat here in a bit of an interrogation situation and not, then you become a bit trapped ... it can be a bit hard" - Camilla

Trying to recall the episode during the session could then once again affect which situation the patient would remember, as she said:

"and then you might think of one episode that was really meaningful for some reason or following reason, but there could have been a different situation that had been really good to remember"- Camilla

5.4 Concerns towards Technology Abuse

Neither of the psychiatrists had major concerns regarding the use of the technology. Again the individual assessment of the patient was brought up. In terms of concerns, Anna for instance said:

"'Now I have anxiety again and it is really bad for me', because I was just thinking of a patient who is not suitable. He is exactly so focused on that" - Anna

The concern of being too focused on the data and the anxiety was also voiced during the Professional Forum meeting. Some of the psychiatrists called it a double-edged sword, because though there were positive sides to being aware of your anxiety and the change in the body there were also negative sides. They said it would especially be problematic if the patient could follow along continuously as the wristband measured. A part of their concern in regards to this was based on the fact that the wristband could not differentiate between different types of arousal, and a part of their treatment was to teach the patients that not all arousal is anxiety or bad in general.

Furthermore, it was discussed at the Professional Forum how the system and the measurements might make the patients avoid certain situations, where they could see the GSR increasing. A part of CBT is to try to expose yourself to uncomfortable situations rather than avoiding them.

Camilla's concerns were mainly related to whether the use of the technology would take any time away from the current treatment.

"... the ethical might be to use the time on it ... so it becomes clear to the patient that it won't go from their treatment for example if we, you will quickly start talking about it in the sessions" - Camilla

Another concern, Camilla mentioned, was that it was important to her that she and the other psychiatrists knew about the technology, so they were able to discuss the use of it with the patients: *"... I think it is good if we are prepared for it and can talk with them (the patients) about how it could be"*. This was also related to the concerns the patients might have in terms of monitoring for instance. Camilla also mentioned how it would not destroy anything if the patients used the system and it did not work for them. For instance, as she was the psychiatrist for Caroline, who dropped out of the study due to the feeling of being monitored, she had thought about how it was actually a good experience for the patient to be able to end the study when feeling uncomfortable.

"And I think there is something in relation to saying no, I think that for patients it is ... it can be difficult in it self, even if it is really uncomfortable to wear and say no, I don't want to do

that, can be a hurdle for some, and it can also be good to then do it and be supported in doing so" - Camilla

5.5 Summary

All in all, both of the psychiatrists could see a potential for using this technology as part of their current CBT treatment program. The discussion at the Professional Forum also showed an interest in how the technology could be used. However, the use of the system would be highly dependent on each individual patient. Several potential use cases for the system were brought up including preventing of anxiety episodes, establishing a connection to the body and disconfirmation of episodes. The psychiatrist's role was also discussed and the importance of their knowledge of the system was highlighted in terms of being able to discuss the use of the system with the patient. It was agreed that it should mainly be the patient who controlled the use of the system but it might be relevant for the patients to bring the data to the sessions and discuss it with the psychiatrist.

Some concerns were also raised in regards to being overly focused on the body and the anxiety, making the system have a negative effect. This was especially a concern if the patients had access to continuous live streaming of the measurements. Another concern was the system taking too much time away from the existing treatment program.

Based on the findings from both the field study and the interviews with psychiatrists, we designed and prototyped a mobile application to be used with the wristband sensor. The design process as well as the prototype and the evaluation of it is described in the following chapter.

6. Design of Mobile Application

Several of the participants thought it was a hassle to use the desktop application, E4 Connect, in the field study. Therefore a mobile application to increase the mobility and incorporate CBT procedure was developed. The design of the mobile app is based on the results from the field study, the meeting at the Professional Forum and guidelines for designing mental health technologies. This chapter will start with a presentation of the design concept, then a description of the process followed by a description of the prototype. The app was evaluated with one of the participants from the field study and presented to two psychiatrists.

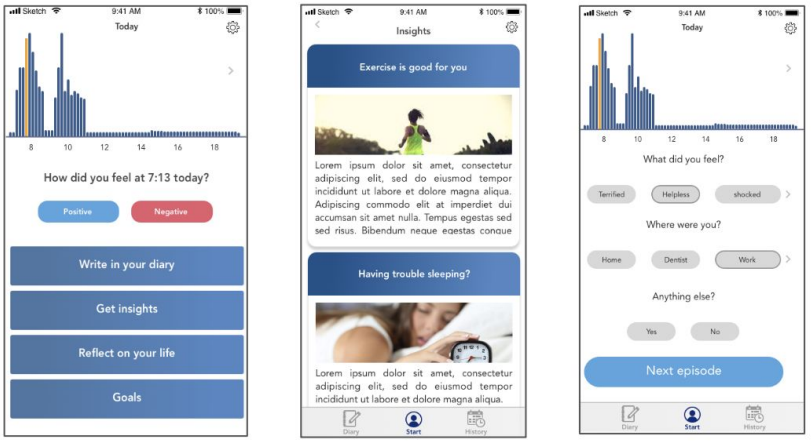


Figure 6.a Overview of mobile application

6.1 Design Concept

The concept of the app is for it to monitor and handle anxiety episodes recorded from the GSR sensor on the Empatica E4 wearable wristband. The app will be synchronized with the Empatica E4 wristband, and use GSR and accelerometer from the wristband. The reason behind developing an app for anxiety monitoring, comes from a wish from the participants in the field study, to create a more mobile solution for everyday and on-the-go use, and be able to get an overview of recorded GSR values. A snapshot of the mobile application is shown in figure 6.a. Later in this chapter a full overview of the application and all of its features will be described.

Storyboard

A storyboard was created, as seen in figure 6.b, to describe a use case scenario of using the app. The storyboard follows Sophie, who suffers from panic attacks, and shows how the mobile app is supposed to be incorporated into her everyday life. The storyboard also shows how Sophie takes the app with her into therapy and shares her findings with her psychiatrist.

STORY BOARD

Meet Sophie

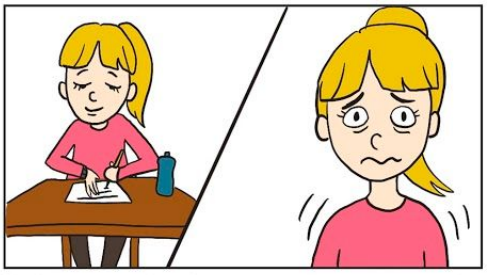
- Sophie is 23 years old and is a student

- She has been suffering from panic disorder for less than a year

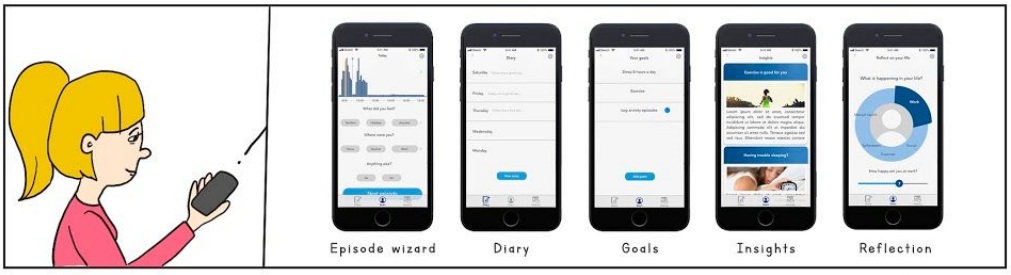
- She has been receiving therapy for 6 months

Sophie is seeing her psychiatrist once every two weeks. Today her therapist suggests trying the Empatica E4 system. She is instructed in the use.

When Sophie gets home, she creates an account on the app on her phone, and syncronises the wristband to the phone.



Sophie wears the device throughout her everyday activities. When she feels an anxiety episode coming. She clicks the button on the device to mark it.



At any given time of the day, Sophie can monitor her anxiety. She does this by looking at the EDA levels and going through five questions regarding her mental health. In the app she also finds: a Diary, Goals, Insights and Reflection.

When it is time for the next therapy session, she brings her phone and shows the data to the psychiatrist. Together they discuss whether there is a pattern in the data, that might help them understand what causes her anxiety episodes.

Figure 6.b Storyboard of mobile application in use

6.2 Process

The app was designed on the knowledge gained from the field study and interview with psychiatrists. In the development of the app, design guidelines for designing for mentally ill was taken into consideration, as well as following the CBT approach. The researchers used the information to brainstorm possible features and functionalities of the app. To see the brainstorm, see Appendix 9. The prototype development started with defining use cases, then sketching, then designing wireframes and at last an interactive prototype.

Prototyping

Prototyping is an important part of any interaction design process. Prototyping as a tool helps the collaboration between designers and help share and present ideas to stakeholders (Preece et. al., 2015). Prototyping is an important tool for reflection and testing different aspects of a design. Preece describes low-fidelity (lo-fi) prototyping and high-fidelity (hi-fi) prototyping. While lo-fi prototypes are quick and cheap to produce, they may not support the full understanding of a product, but can be quickly modified. Hi-fi prototypes are more detailed and is more expensive to modify, but may provide a more real look into a finished design. Either way, the lo-fi to hi-fi can be seen as a scale. The process in this project started with a lo-fi prototype in the shape of a brainstorm, while the endresult, a interactive prototype, is more hi-fi.

Use Cases

Use cases were done to get an understanding of what the user journey would be like. Use cases is presented by Preece et. al. (2015) as a possibility to get the users point of view in a system. The main uses cases were:

1. Gain an overview of GSR values for a given time period
2. Determine when the user experiences anxiety in that time period
3. Write down notes for the anxiety episodes they experience (more or less specific)

Other cases were also discussed such as:

4. Gaining insights of oneself
5. Writing in a diary
6. Reflection on life
7. Personal settings

Sketching

Sketching is a lo-fi prototype tool used to rapidly communicate ideas and collaborate through designing. Eva Brandt (2006) describes how this type of prototyping aids the conversation in design, and how tangible prototypes can help gain a joint understanding of a product. When discussing multiple design proposals, being able to communicate through sketching is a really great asset. A collection of sketches from the prototype is seen in figure 6.c-f. At this point in the process a lot of explorability was conducted, thus the many different ideas are seen.

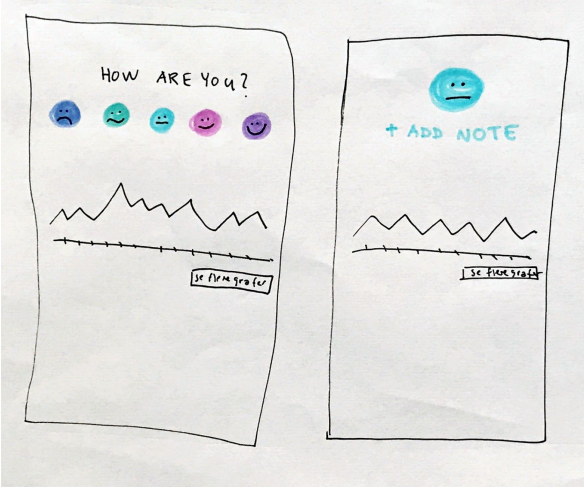


Figure 6.c GSR data as emoticons

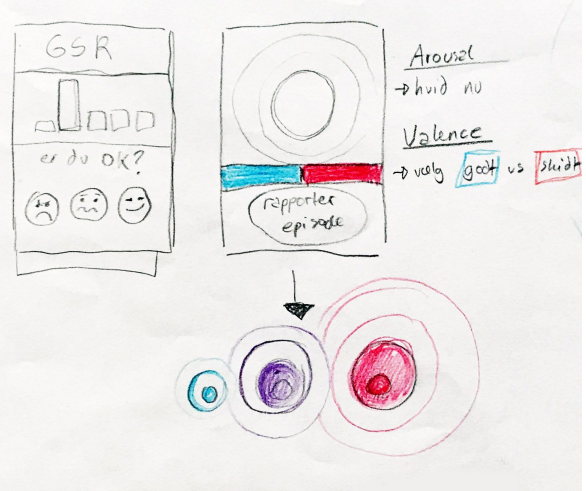


Figure 6.d Live-streaming GSR

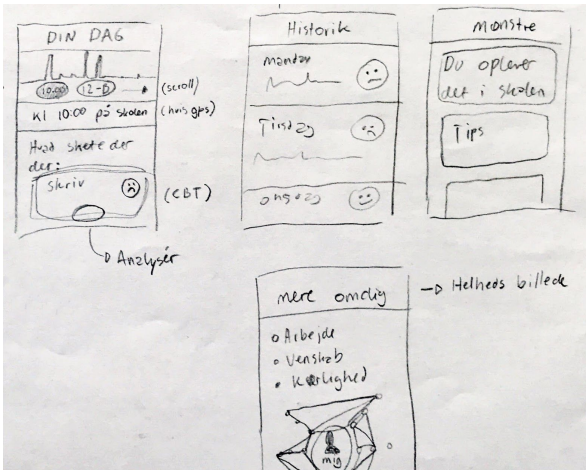


Figure 6.e GSR with time stamps

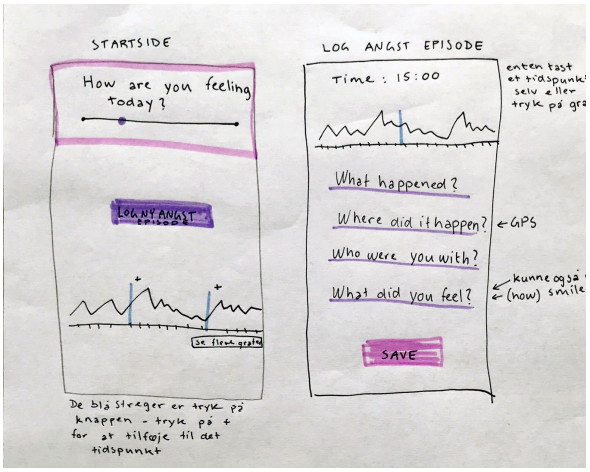


Figure 6.f Sliding gesture to indicate mood

In the sketch in figure 6.c the front screen presents the user with five emoticons used to rate how they are feeling overall. It also shows the GSR graph and the user can see other graphs as well if they press the button in the bottom right corner of the GSR graph. After the user picks an emoticon, they can write a note of how they are doing. The sketch in figure 6.d shows an idea of live streaming of GSR data, where the data is presented as glowing circles in either blue for good and red for bad in terms of anxiety. The user is able to write down an episode if they press the button at the bottom of the page.

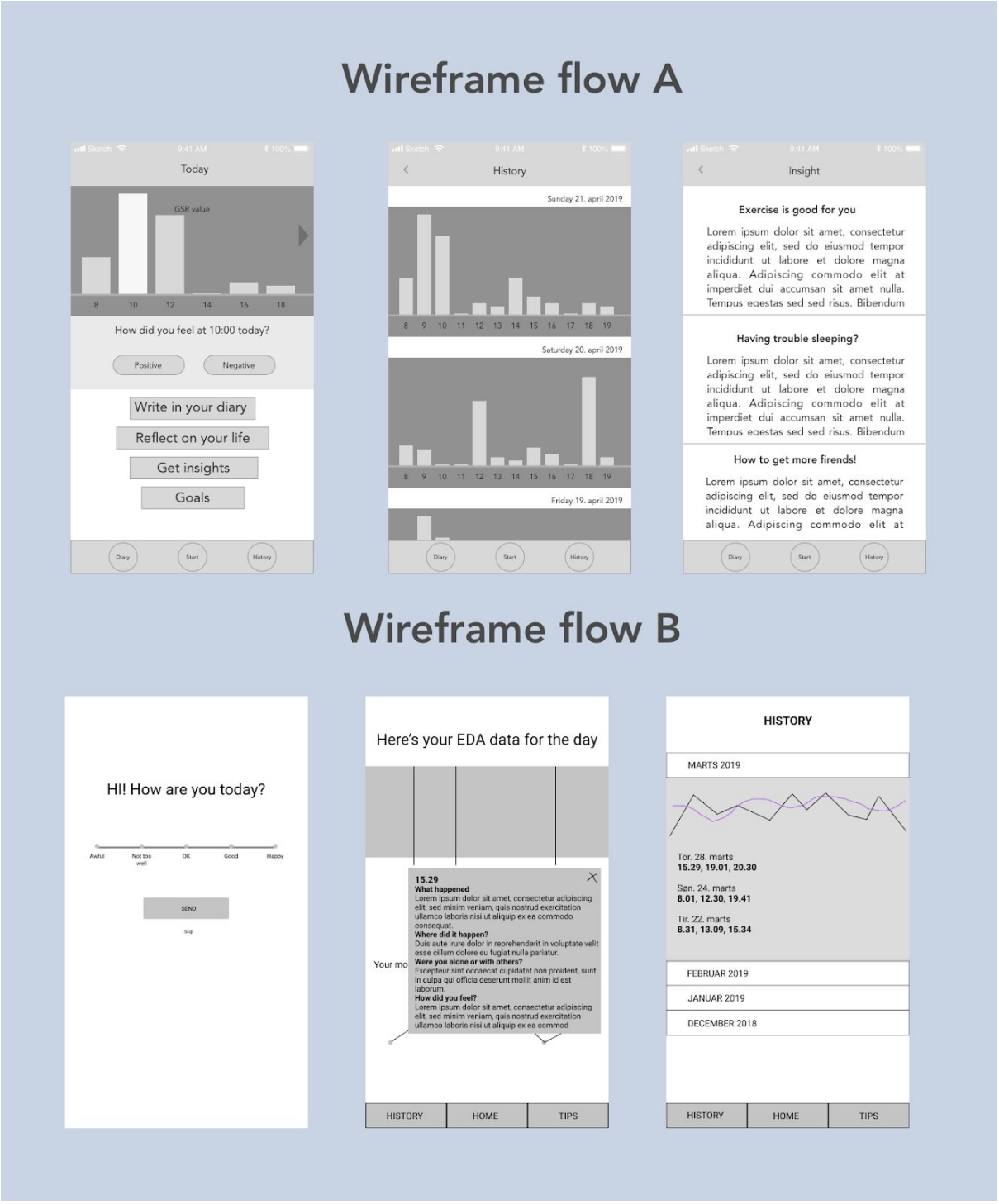
Figure 6.e shows a sketch where the user sees the GSR graph for the day with some timestamps marked where the GSR was increased. The user is able to fill out what happened at that time. After writing it down, the user can press ‘analyze’ where after the app will try to analyze the data to help the user discover patterns and get more personalized tips. The sketch also shows a history page, shown in the middle of the sketch, where the user can get an

overview of how they are doing. The sketch in figure 6.f shows an app where the user is able to note how they are feeling by using the slider at the top of the page. Furthermore, they can see their GSR graph with markers from pressing the button throughout the day. The user can press the + sign by each of the markers which will send them to the screen shown to the right in figure 6.f, where they can see the graph still as well as fill out how they are doing based on the questions from the Day Reconstruction Method, that we used during the study. The user is also able to log an episode for a different time than when they have pressed the button on the wristband, by pressing the purple button in the middle of the screen to the left in figure 6.f. This will once again send them to the right screen.

After sketching and exploring different aspects of the design ideas, we chose to go with a mobile app that does not show live feed of the GSR data. This was done due to the knowledge that getting a live feed may make mentally ill patients feel like they are constantly being monitored, which would not work in their favor. This knowledge came from the Professional Forum meeting held at the psychiatric ward.

Wireframes

Wireframes are mock-up screens of a design to illustrate what functionality will be a part of the design and how the system will interact. Wireframe flows depicts how functionalities in a design will interact with each other. Wireframes can be of a higher fidelity than sketches and is often specifically focused on the flow of the app (Preece et al., 2015). The wireframe flow is used in this project to illustrate and communicate how the design and flow of the app will function. Two alternative wireframe flows were created by the researchers, and in collaboration, a final wireframe for the app was ready to be made into an interactive prototype. The two wireframes are seen in figure 6.h.



Figur 6.h Wireframes

6.3 Design Guidelines

As mental healthcare technologies are a new area of design with specific concerns, it is relevant to consider certain guidelines when designing these systems (Doherty et al., 2010). Doherty presented an initial set of guidelines for designing mental health technologies based on existing literature as well as previous software development projects for supporting mental health. They divided them into three groups of guidelines: design process, design recommendations and evaluation. In terms of the design process, Doherty (2010) argued the

importance of including both HCI professionals as well as clinical professionals in the design process. Though the interdisciplinary work can be challenging, it is also rewarding, as clinical professionals might have relevant and insightful concerns regarding potential drawbacks of using the technology as part of the treatment.

As these guidelines are initial they should not be seen as prescriptive rules to follow, but rather as formative as they raise awareness of certain design concerns worth considering when designing mental health technologies (Doherty et al., 2010). We have considered several of these design guidelines when discussing the use of the system in the current treatment and also when designing the corresponding app, which will be presented in the section 6.4 Interactive prototype.

6.4 Interactive Prototype

The interactive prototype was built using Sketch app (Sketch App, 2019). Screenshots of the app is presented with a description but a larger version of the screenshots is available in Appendix 10. The app is in the early development state, and the description is therefore on a conceptual level. A flowchart was made to present the flow of the app, which is seen in figure 6.i. The interactive prototype includes the following features:

- Start
- Episode wizard
- History
- Insight
- Reflection
- Diary
- Goals
- Settings

Some of these features are furthermore in line with what Li et al. (2011) found to be essential questions people ask about their data when using technologies for self-reflection. These questions are related to: Status, History, Goals, Discrepancies, Context and Factors and furthermore divided into two phases: Maintenance and Discovery. Li's results will further support the reasoning behind the included features in the proposed application.

In the following each feature will be described along with where the feature came from, and what guideline it fulfills.

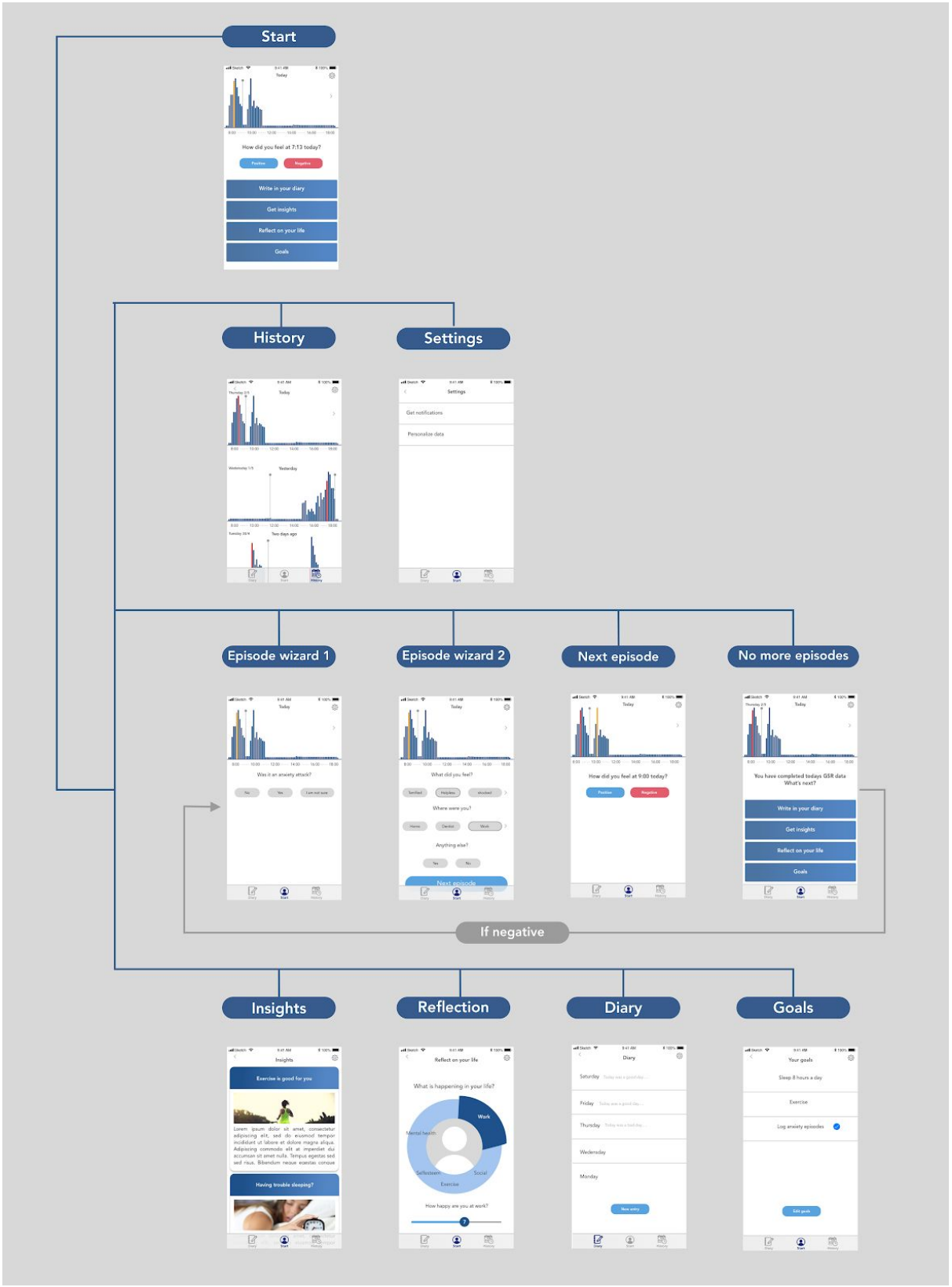


Figure 6.i Flowchart of the mobile application

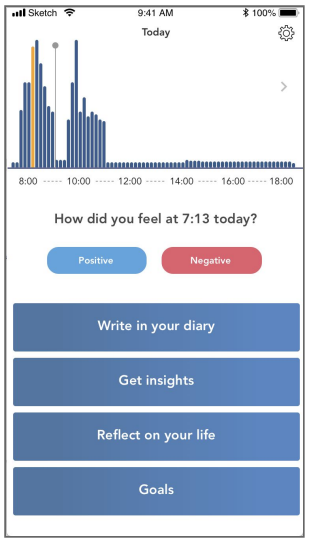


Figure 6.j Start

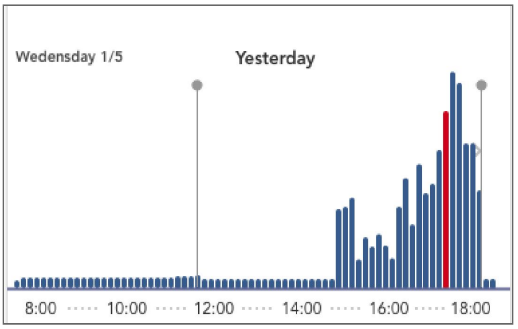


Figure 6.k GSR graph

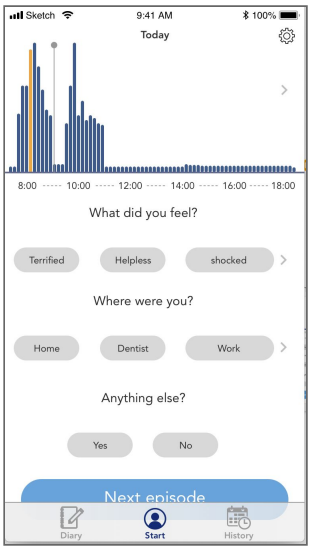


Figure 6.l Episode wizard

Start

The start screen is seen in figure 6.j, and will be the first screen the user is presented with when opening the app. The first thing the user sees is the GSR graph in the top of the screen, figure 6.k gives a better view of the GSR graph. The GSR graph has blue pillars and depicts the GSR levels throughout the day. On figure 6.k you see a timeline underneath the GSR graph which goes from 8:00 to 18:00. This means that the user has worn the wristband since she got up at 8 o'clock and is now looking at today's activities at roughly 18:00 in the evening time. On the GSR graph there is also a grey line with a dot on top. This grey line shows when the user has marked an anxiety episode herself during the day. There is an orange pillar marked at roughly 8:00 in figure 6.j. This is the underlying system telling the user that there was a spike in the GSR level. The user is asked: *How did you feel at 7:13 today?* This related to the sudden spike in GSR, and 7:13 is where the system sees it began. Since GSR does not tell emotions but just the level of arousal, it is necessary for the user to classify whether the spike in GSR at 7:13 was positive or negative. These buttons takes the user through the Episode Wizard. From the start page it is also possible to access: Diary, Insights, Reflection and Goals as seen in figure 6.j. To have a collection of the tools usually used within CBT right at hand, was a wish from the participants as well as the psychiatrists. With this app, the participant has the GSR data and diary on the phone, and giving the Episode Wizard, the user can easily note possible anxiety episodes.

Episode Wizard

The episode wizard is the core of the application and is a wizard the user can go through in order to fill out anxiety episodes, as is seen in figure 6.l. From the Start page, the user is first asked if the increase in GSR levels shown in the GSR graph is negative or positive. The

framing of negative and positive is based on the three-dimensional model *Pleasure-Arousal-Domination* (PAD) for describing and measuring emotions (Bruun and Ahm, 2015). Arousal spans from excited to relaxed on this scale, but can both describe a positive emotion such as joy or a negative one like stress or anxiety. The GSR sensor is only able to determine the arousal level and not how pleasant the emotion is which spans from negative to positive in the PAD model (Lang, 1995). The question of whether the spike was positive or negative allows the patient to subjectively determine how pleasant the emotion was.

If the user selects positive, the graph will move on until there are no more graphs spikes to examine. If the user then wants to add another episode that the system did not detect, the user is able to touch on the graph to add an individual episode. When the user selects a negative experience with a timestamp, i.e. 7:13 as shown in figure 6.j, the user will go through the Episode Wizard as seen in figure 6.l. The Episode Wizard asks questions, on the specifics on the episode, i.e. *What did you feel? Where were you?* To answer the question in a quick and easy way the app suggest some predefined tags. For instance to the question *What did you feel?* The predefined tags are: *Terrified, helpless or shocked.* It is not always possible to know precisely what the user feels, the user is therefore able to add their own tags in order to further customize the experience. The app will then remember the customized tags for later use. This corresponds to the design guidelines established by Doherty (2010): *Design for client engagement* and *Make the system tangible.*

The questions and terminology are well known in DRM as well as it is a cornerstone in CBT. By using CBT methodologies in the app design, the app corresponds to the design guideline: *Build on accepted theoretical models of mental health care.* When an anxiety episode has been logged, the before orange pillar now turns red as seen in figure 6.k. The episode Wizard became essential in the app because of the results of the field study that showed how the patients could write down episodes and get them confirmed by the GSR graph. Maya also found that looking at the graph first could support her writing down her episodes.

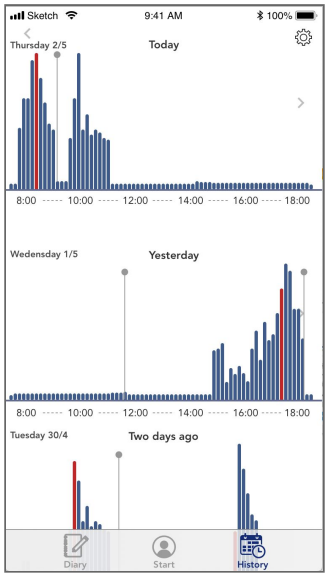


Figure 6.m History

Figure 6.n Reflection

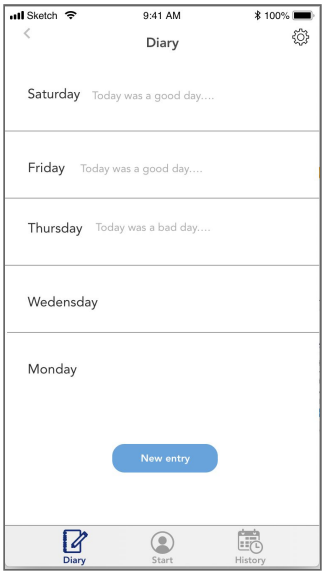


Figure 6.o Diary

History

The history tab is seen in figure 6.m and keeps track of previous days. It is possible to click the graphs to get a closer look and watch or edit the happenings of the previous days. In a study by Li et al. (2011) they found that history is important when it comes to personal reflection apps, as it helps get an overview and create a pattern over data. The need for a history overview differs depending on the users' need for discovery or maintenance (Li et al., 2011). The need to have a history of previous GSR sessions came from the result of the field study, i.e. in section 4.2 Understanding Oneself and Patterns, where Mathilda is able to see patterns of her anxiety by looking at the data from previous days. This function is leaning towards the: *Make the system adaptable and sustainable* design guideline, because it helps users keep track and get an overview of their anxiety episodes, also to be able to incorporate the GSR data in therapy. This was a wish from the participants as seen in Field Study Results section 4.4 Use in Treatment, where Mathilda discusses her GSR findings with her psychiatrist, though she is not able to bring the actual data with her. The need to save the anxiety episodes is an important aspect of incorporating this technology in therapy thus fulfilling the guideline on: *Provide flexibility in the delivery of support.* The eligibility to have the history tab also comes from the psychiatrists, who found the function useful, when bringing the GSR data into therapy.

Reflection

The reflection tab is seen in figure 6.n and is an addon for further possibility of exploring yourself and reflect on mental health. The idea is that there are some sections, i.e. mental health, self esteem, exercise, social and work. The user can now rate the experience of those sections on a scale of 0-10 by dragging a slider. On figure 6.n the user scenario is that the user has selected work and giving it a seven on the scale. This type of interaction lies heavily on the guidelines: *Design for client engagement and Make the system tangible.* Having a

mental disorder affects many aspects of your life and it is therefore important to consider the patient's whole life. This data can also be presented to the psychiatrist in sessions, in order to gain more in depth understanding of the patient's full life experience.

Diary

The diary tab is seen in figure 6.o, and is where the user can add additional information about how they are feeling. Writing in a diary-like format is an essential part of CBT and having the opportunity to write in your phone in the app is something the participants in the field study already express doing. Therefore providing the user with the ability to write notes directly in the app corresponds with the participants wish to have a more accessible tool at hand, as seen in Field Study Result section 4.7 Design Considerations. Writing custom tags is already a part of the Episode Wizard but since the Episode Wizard does not provide for a more detailed and freehand description of the users day, the diary was added. The diary has a list of the previous days and a button for new entries.

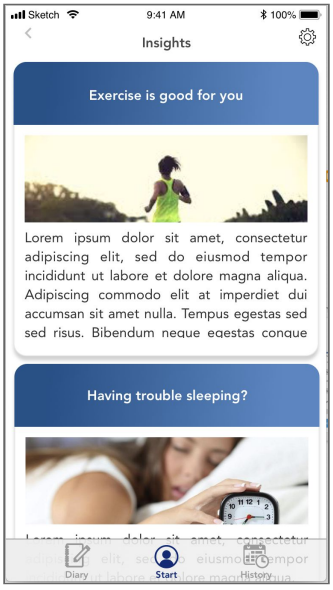


Figure 6.p Insights

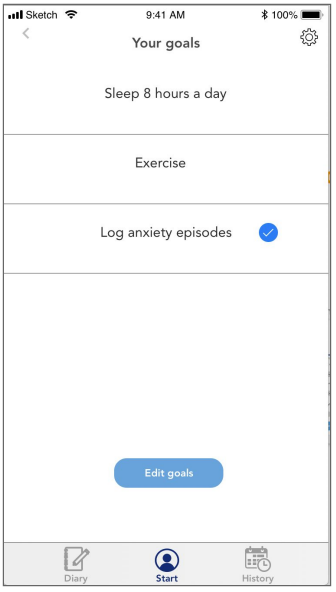


Figure 6.q Goals

Figure 6.r Settings

Insight

The insights tab is seen in figure 6.p and is a news feed of advice and tips for the user. The insights are supposed to derive from what the user is experiencing. So for instance if the user reflect her life that she is not getting enough sleep, the insights will provide content that matches this behavior. This is related to the guideline: *Provide flexibility in the delivery of support,* due to the personalization of the content.

Goals

The goals tab is presented to the user in order to give the user the opportunity of establishing goals to keep track during the day or on a long term. The goal can for instance be to log

anxiety episodes as depicted in figure 6.q. In figure 6.q the goal of logging anxiety is selected. Goals are a great way to help users keep track of progress, and in the study by Li et al. (2011), it was found that setting up goals can have an effect on how the user pursues them, and what data they look for to achieve their goals. This corresponds to the design guideline: *Provide flexibility in the delivery of support,* because the user have multiple ways of interacting with the system functionalities.

Settings

The Settings feature is seen in figure 6.r and is located everywhere in the app in the top right corner. In settings you can adjust notifications and personalize data. The need for a setting derived from the field study, in which it was made clear that the participants had many individual requirements for using the app. For instance Mathilda is keen on having the app made customizable to fit her needs, as seen in Field Study Results under section 4.7 Design Considerations. The guideline: *Provide flexibility in the delivery of support,* also suggests that we need technology that adapts to the person and specific case, rather than one size fits all.

6.5 Evaluation

An evaluation was conducted with one of the participants from the field study, Maya. She was presented with the interactive prototype, and a case description, as seen in the interview guide in Appendix 11. All of the quotes in this section comes from Maya, to see a full transcript of the evaluation see Appendix 12. Also the two psychiatrists, Anna and Camilla were presented with the prototype. This is elaborated later on.

Overall the app was well received by Maya. She seemed to understand the app and its functionalities. The thing she was most excited for, was that the GSR data is now available on the phone, as she says:

"(The app is) A lot better... Because you don't have to go on you computer ... And also because, it is more personal, this, in regards to all of the diary and reflection of your life and goals and such…"

With the E4 Connect desktop application that Maya tried during her participation in the field study, it was burdensome for her to open a computer and find the fill-out sheet and fill it out by hand. With the new app it is all located on your phone, that you carry around with you. Maya did experience some distress and confusion when using the Episode Wizard. For instance there were some confusion on what the grey line with the dot on top means: *"I just thought that the grey arrow was the timestamp"*. This indicates a misunderstanding of the GSR graph. Also she was confused with where the app suggested there was an anxiety episode. After trying the GSR graph more she ended up having an alternative idea to the grey line:

"Well I thought to make it less boring, you could add a star or a flower or just something small right there. Because it does not need to be so tiny tiny as it is here."

Maya had an idea to incorporate something a little more visible to indicate when the user had pushed the button on the wristband. She also mentioned that it would not matter that the object is larger than the grey line now, because an anxiety episode may last for several minutes. Maya did have suggestions for what she thought would make a better app. For instance she would prefer that the History tab could expand when you tilted the device. This could give more room to show GSR data, and might make it easier for the patient and psychiatrist to view data on the phone.

There were also some concerns on the challenges with technology. For instance Maya expressed that she needs to trust that the technology is functioning:

"...The hassle you can end up having with it. I am not a fan of that. But if it works, it works. Then there is no problem."

When Maya participated in the field study, she experienced that sometimes the system failed to synchronize. This is a annoyance to the user and becomes a disturbance in the use. When it comes to ethical questions and technology trust, Maya also had a concern towards automatic identification of anxiety episodes.To the question whether the app should identify anxiety episodes on its own, Maya answered:

*"Well I believe that it should be careful on telling you that. If for instance you did not have it (Anxiety), you should not be imposed that you could have had it...On the other hand, it could be positive to show, that it (*GSR*) went up here and you actually managed to control it without thinking about it."*

It is important to understand the user's reality and be careful about declaring every GSR spike an anxiety episode. Therefore it is crucial to remember to: *Make it clear that data are secure,* as one of the design guidelines by Doherty (2010) sounds, in order to not make a situation worse.

To summarise the evaluation with Maya:

- She found the app to be better than the E4 Connect desktop application
- The GSR graph can be confusing
- Concerns about the technology as to whether it will work properly

6.6 Evaluation with Psychiatrists

The interview with the psychiatrists ended with a quick introduction of the prototype, just to get a quick opinion on the matter. Since Doherty (2010) explains that designing with the therapists in mind is important, we found it relevant to introduce the app to them to get their

thoughts on it. The psychiatrists both thought of the app to be good, since it is on the patient's mobile and can be carried around, and the anxiety episodes can be described right after it occurred, as Anna says:

"So they already have something, so that they do not need to remember to find a piece of paper. That is a bit tedious. And everybody is already looking at their phone when you do anything." - Anna

Also the fact that the patients can bring their data with them in therapy is a great asset of the app. Anna explains in this quote how she usually uses this technique of bringing the anxiety episodes that has happened in the time between sessions, into the therapy session:

"... I think that it would be fine to bring it up. I usually do that with the regular week assignments schedule." - Anna

In relation to this, Camilla also expressed how bringing the GSR data overview into sessions can be beneficial, and easier than to explain it to the psychiatrist:

"...It could be that they (the participants) think it is easier to show me the graph rather than explaining how the graph looks." - Camilla

The fact that it is the patient who shows the results to the psychiatrist corresponds to the design guidelines: *Do not place burdensome time demands on therapists* and *Consider the responsibilities placed on therapists.*

6.7 Summary

An app was created in order to make a mobile solution to the desktop application: E4 Connect, that contains basic methodology from CBT. The app was created through a prototyping process going from lo-fi prototyping in brainstorming and ending up with a semi-functioning higher fidelity interactive prototype. The app was evaluated with one previous participant in the field study, and two psychiatrists. The results indicates that the new app is more usable than the desktop application, with the mobility to be the greatest aspect. But the new app still have some limitations, such as the GSR graph is unclear and further evaluation and development is needed.

7. Discussion

This chapter will include a discussion of potentials, challenges as well as future directions for introducing sensor technology into the current treatment of anxiety on the basis of the results of the field study, the interviews with psychiatrists, the design and evaluation of the mobile application as well as related work.

The three main findings from our study are that the system can: Help patients recall episodes, help patients confirm episodes and help patients disconfirm episodes. These three areas will be discussed in the following, but first, a discussion of the reasoning behind introducing technology into anxiety treatment and the importance of relating subjectively to the object data. At the end of the chapter, there will be a discussion of design considerations for the mobile application as well as future directions and limitations for the presented study.

7.1 Improving Effectiveness of Treatment

As mentioned previously, there already exists well-established treatment programs for anxiety disorders with Cognitive Behavioural Therapy (CBT) being the golden standard. However, several issues prevent those in need from receiving treatment, including access to treatment due to i.g. geographical, attitudinal and financial limitations (Bry et al., 2018) resulting in nearly two thirds of mentally ill people not seeking help (Barak and Grohol, 2011). Furthermore, for those who do seek treatment, the drop-out rates of CBT programs are high (Winslow et al., 2016). This is where introducing technology in treatment might be useful, both in terms of improving the effectiveness of treatment to limit the dropout rates and to make the treatment more accessible to a wider population (Doherty et al., 2010).

In Levy et al. (2016) and Boukhechba et al.'s (2018b) studies, as presented in chapter 2: Related Work, they focused on both making treatment more accessible through computerized treatments without the need for a therapist to be present, and by using sensors to assess the effectiveness of these treatment programs. When we presented the idea of using the system before treatment with a psychiatrist, there was raised concern at the Professional Forum, as it might make the patients more nervous, possibly resulting in them not wanting to receive treatment afterwards. As the technology is not able to differentiate between different types of arousal, it is highly important that the patient relates subjectively to the data to assess whether it is anxiety or not. The psychiatrists were concerned that patients, who did not receive treatment already, were not able to do so and might therefore misinterpret the data and think they were having anxiety a lot more than they actually were. They found it important that the psychiatrists were a part of the introduction of the technology, which will be elaborated later in this chapter. Therefore, for this case, it might be more relevant to consider how the system can help improve the effectiveness of the current treatment rather than making it more accessible.

7.2 Burdensome Homework in Current CBT

The high dropout rates in current CBT programs might be due to the homework assigned during sessions which can be burdensome for the participants. Though the homework pose a great burden on participants, its use in treatment has also proven to be more effective than when not completing homework (Kazantzis J. et al., 2010). The psychiatrists, we interviewed, also mentioned how the patients did not always complete the homework, i.g. because of lack of motivation or because they forgot to do it. According to the psychiatrists, not completing homework makes it difficult for the patients to recall all episodes or details about these, during sessions, which could put a lot of pressure on the patient in that situation. The patients Mathilda and Caroline also mentioned how it could sometimes be difficult to remember all episodes.

Problems regarding memory and burdensome self-reports is also known in HCI in studies for assessing the user experience of technology such as studies using the Day Reconstruction Method (DRM) or the Experience Sampling Method (ESM). In DRM studies, participants are asked to fill out questionnaires at the end of the day about their experiences (Kahneman et al., 2004) while in ESM studies, participants are prompted several times a day to fill out questionnaires about how they are feeling in that moment (Scollon et al., 2003).

In DRM studies, the pitfalls include memory bias and peak-end rule, where the user tends to remember the most intense emotion or the one closest to the end more clearly than others (Redelmeier and Kahneman, 1996). This was also touched upon by one of the psychiatrists who said that patients might remember the most intense episodes when they are asked to recall in sessions, but other episodes might also be interesting to remember. They furthermore explained how the patient would most likely be affected by their current situation when recalling episodes, which might not give a correct image of their experiences. Boukhechba et al. (2018b) also stated how self-reports were prone to biases both because of human memory but also because of the patient's desire to show improvements. When Bry et al. (2018) made a survey of anxiety apps for youth they found that many of the existing apps on the market were based solely on self-reporting and thus not making use of the advanced technologies in mobile phones that could provide objective data to support self-reports, such as GPS.

ESM aims to limit the memory bias by prompting participants to fill out questionnaires throughout the day. Though it limits memory bias, it does put a significant burden on the participants (Scollon et al., 2003). This was also found in the related work done by Bauer et al. (2018) who experienced high dropout rates from their study which included several prompted self-reports throughout the day.

Cued Recall Debrief (CRD) is another method in HCI where he participants are provided cues collected passively throughout the day to help them recall their emotions (Omodei et al., 1994). In the context of anxiety and CBT, these cues could be the physiological

measurements collected through GSR, but could also be data from other contextual sensors such as GPS. In our field study, it was discovered that these measurements could not only help the participants recall their episodes, but also help confirm the episodes they did remember. Some of the psychiatrists also mentioned how it could potentially be used to disconfirm episodes as some patients feel they have anxiety constantly, which the psychiatrists said was doubtful. However, it was also discussed how the system might not be suitable for exactly these patients as they were already so focused on the body that it might have a negative effect. These were the main findings from the study. An important aspect of all the main findings is that it is important that the patient relates subjectively to the objective data measured through GSR, which will be elaborated in the following section.

7.3 Importance of Subjective Reflection of Objective Data

GSR sensors measure the fine-grained change in skin conductance (Poh et al., 2010) which has been proven to be a sign of increased arousal. Arousal can both be seen as a positive emotion of excitement or engagement or as a negative one like stress and anxiety, and the GSR sensor is not able to determine which one caused the change in skin conductance (Lang, 1995). The psychiatrists therefore stated the importance of the patient being able to subjectively assess the data to determine whether the increase in GSR was a sign of anxiety or not. This was something the psychiatrists could support the patients in doing. It was especially important for patients who felt they experienced anxiety constantly and for those who was completely cut off from the body, as Camilla stated in the interview.

Boukhechba et al. (2018b) hypothesized how the correlation between HR data and accelerometer could help determine whether the increased HR was due to a negative mental state such as stress or anxiety or due to physical activity. This could potentially also be done with the correlation between GSR and accelerometer. Boukhechba found that there was in fact a pattern in the correlation between HR and accelerometer and the changes in negative mental state. However, they did not account for the possibility of positive arousal not due to physical activity. During our field study, the participants said they did not have difficulty determining whether the peak in GSR was due to physical activity as they knew at which times they had been working out.

Hollis et al. (2018) found that users tend to trust affect detection algorithms as they have difficulty understanding them and thus make correct conceptual models of how they work. We also saw this in the field study where Mathilda for instance said that she was not able to see how the data could be fake when it was objective measurements with concrete numbers. It can be problematic if the users trust the system blindly as there might be many different factors that influence the skin conductance other than anxiety. Furthermore, it is not certain that everyone who has anxiety, experience an increase in skin conductance. For instance, Maya had an episode where the graph did not spike, but she knew she had had an episode and Eric was not able to see an increase in skin conductance even though he had perceived his anxiety level to be increased during the period where he wore the wristband.

While Boukhechba et al. (2018b) were studying how physiological measures can potentially eliminate self-reports all together, we argue that there is still a need for subjective self-reports to interpret the objective data. The subjective self-reports in combination with the objective data can then potentially help patients recall episodes, confirm episodes or disconfirm episodes.

7.4 Main Findings

The three main findings from our study was how the system could help patients recall anxiety episodes, confirm episodes and disconfirm episodes. The purpose of using the system would depend on each patient, as the psychiatrists said there were many different types of patients who might benefit from using the system in different ways. In the field study, we also experienced a big difference in the participating patients in terms of personality and how they responded to using the system. In the following, each of the main findings will be discussed.

Recalling episodes

The system can support recollection of episodes by providing the patients with cues similar to the ones used in CRD studies, as discussed previously. The patients said it could sometimes be difficult to remember episodes, and the psychiatrist said as well that the patients' memory and concentration were often affected by their disorder. None of the related anxiety studies, presented in chapter 2: Related Work, used sensor data to support patients in recalling as they did not show the sensor data to the patient but rather used it to either predict their anxiety level (Boukhechba et al., 2018a) or to determine the effect of a treatment program (Levy et al., 2016; Boukhechba et al. 2018b). Bauer et al. (2018) were the only one who used the sensor data actively in treatment but this solely consisted of GPS data as a way to determine isolation and no physiological measures.

Confirming Episodes

Another use case for the system is to help confirm episodes as the GSR spikes could show the patients that a change happened in the body and not only in the mind. Three of the patients in our study found the confirmation comforting. Once again, none of the related work articles used the sensor data for helping patients confirm data. However, they did study how the sensor data correlated with the participants' mental state. For instance, Boukhechba et al. (2018b) studied the correlation between HR and accelerometer and subjective measures of negative mental state, as described previously in section 7.3. Boukhechba et al. (2018a) used GPS data to predict participants' social anxiety level based on their subjective score on a scale for social anxiety.

The psychiatrists said that they currently confirm the patient's anxiety episodes by asking the patient about the physical symptoms such as increased HR, trembling or sweating. They said if the patient had a number of these it was most likely a sign of anxiety.

Being able to confirm episodes could also support patients in discovering patterns of their anxiety like Mathilda did during the field study. This could help uncovering when and where the anxiety happened. Anna, the psychiatrist, said it was beneficial for the patient to learn what triggered their anxiety so they could make an effort to prevent it.

Disconfirming Episodes

At the Professional Forum, attention was drawn to how the system might also be useful in disconfirming episodes for patients who feel they have anxiety constantly. This could help both the patient and the psychiatrist uncover when it was anxiety and when it was not. Using sensor data to disconfirm episodes were not studied in any of the related work papers. Garcia-Ceja et al. (2018) who did a review on mental health monitoring systems, also found that studies typically use sensor data for either associating, detecting or forecasting a mental state and not how the sensor data could be used to disconfirm a perceived mental state.

The psychiatrists also stated how these types of patients might not be suitable for using the system, as it would draw even more attention to the body, possibly making them feel more anxious. Therefore, it was important that the psychiatrist assessed the patient's' suitability for using the system and made sure to inform the patient that the GSR did not differentiate between different types of arousal.

7.5 Design Considerations

In this section some future design considerations for the mobile application that was designed will be discussed based on our findings and related work. This includes the importance of including psychiatrists in the design process, the effect of the framing of the GSR data as well as the potential of just-in-time interventions based on GSR data.

Involving Psychiatrists

As previously mentioned, it is important to include psychiatrists or other health professionals when designing mHealth technologies for mental health (Doherty et al., 2010). Only one of the four papers presented in chapter 2: Related Work included health professionals in their study. Bauer et al. (2018) included both patients and caregivers in their study, as the design they made was intended to be used by both. Though they included caregivers when evaluating their design, they did not involve them actively in the design process.

When looking at the existing mobile applications for anxiety management on the market, it is also evident that few of these use evidence based treatments or are validated by health professionals (Bry et al., 2018). The importance of including psychiatrists in our study was clear, as it gave us insights on concerns and potentials for using the system in treatment that we had not considered previously. For instance, these included the individual assessment of patients, the potential double-edged sword of body awareness and how the system should be introduced in the current treatment.

Framing of GSR data

Hollis et al. (2018) studied how three different framings of GSR data affected participants' self-judgment of emotions. The framings were Negative ("stress"), Positive ("alert and engaged") and Control (no frame). They found that the framings greatly influenced the participants' perception of their emotions. It is therefore important to consider the framing when presenting the GSR data to the patients. This is also in line with what the psychiatrists said, in terms of teaching patients that not all arousal is a sign of anxiety. In the proposed design of the mobile application, we leave it up to the users to determine whether a spike on the graph was negative or positive, and then afterwards answer if it was anxiety or not.

Just-in-time interventions (JITIs)

An ongoing discussion among the psychiatrists were whether they should try and uncover the patients' trigger in order to be able to predict episodes and prevent them. There was some disagreement regarding this between the psychiatrists. The presented system might have the potential to help patients uncover the trigger through uncovering of patterns in when and where the anxiety happens. The motivation for finding the trigger would be to understand more about one-self, be able to prevent or predict an episode and potentially provide just-in-time interventions (JITIs). Even though there was disagreement regarding finding the trigger, the psychiatrists all agreed that an important element of treatment was for the patient to seek out the anxiety to learn to handle it.

Predicting a mental state was also seen in Boukhechba et al. (2018a) in relation to predicting social anxiety levels. It was also touched upon in Garcia-Ceja et al.'s (2018) review on mental health monitoring systems though not specifically aimed at anxiety but mental health in general. Garcia-Ceja stated the difficulty of predicting an upcoming mental state, but also highlighted the great potential for introducing timely treatment if prediction was possible. In Boukhechba et al. (2018b) they also discussed the potential of providing JITIs at the correct moment based on sensor data.

In general, JITIs have gained interest in the literature as a potential for delivering treatment content through the user's mobile phone in the moment anxiety is experienced (Bry et al., 2018). However, the timing of such interventions need to be considered if they are solely based on physiological data, once again due to the many factors that can influence arousal (Sarker et al., 2016).

Being interrupted right before an anxiety episode occurs might also be a distraction to the patient, and furthermore, in that moment they might not be able to interpret or act on the information they are receiving (Hollis et al., 2018). This is in line with what Maya said that she did not want anything else disturbing her if she was having anxiety as she was already so engaged in the episode and in herself at that moment. Mathilda also said she did not want an alert if the wristband detected an upcoming episode, however, could see a potential in the technology providing her with a countdown to when it would be over.

It could be argued that exactly for the reason Maya stated, it might be beneficial for the patient to get a notification that an anxiety episode was on its way to give the patient an opportunity to leave their current situation if possible. Further evaluations focusing on JITIs and alerts are needed to investigate the potential further.

7.6 Future Directions

The future work suggested is related to further study of the technology with participants and improving the mobile application.

Further Research Needed with More Participants

It is important to note that the field study conducted in this paper is preliminary and using this technology in treatment of anxiety is in the discovery phase. Moving forward there is a need to conduct further research with patients. For instance by conducting studies over longer periods of time to study the development of the user experience when the novelty effect wear off. It is also important moving forward to conduct further research with caretakers and psychiatrists. The interviews conducted in this study yielded sides of the technology not thought of by the researchers and much more is out there to be learnt.

From the interviews with the psychiatrist, it was found that using GSR as cues in combination with CBT could be beneficial for other mental illnesses as well, other than anxiety. Other mental illness could include: depression, OCD or stress. Future work could therefore also address researching use of GSR technology in combination with CBT, for other mental illnesses and user groups.

Improvement of Mobile Application

The mobile application developed for this project is limited in functionality and needs further research and design before deployment. During the evaluation of the app, some design considerations did occur, but it was not possible given the time, to consider these findings. In the future we want the app to support more sensors such as: GPS, accelerometer and HR. Development of an algorithm that creates patterns is also crucial prior to deployment, since no technical background has been tested for that so far.

7.7 Limitations

A limitation was working with mentally ill patients who are very difficult to get to participate in the study. For this study it was relevant to use representative participants suffering from anxiety, but finding these patients was difficult. We tried soley recruiting through the psychiatric ward, but only two patients agreed to participate. An obstacle might be the fact that participants had to be part of a new study, while already being in unfamiliar environment by just starting up treatment. The participants we found through Facebook were not in treatment, and their anxiety was more in control than the participants recruited from the psychiatry. This meant that there was a difference in the participants' use of the technology.

Another limitation was the prototype of the mobile application. When evaluating the app, both the participant Maya and the two psychiatrists had difficulties in navigating through the prototype, giving the limited interactions.

8. Conclusion

We researched existing literature, conducted a field study with four anxiety patients, participated in a meeting with a group of psychiatrists as well as interviewed two psychiatrists individually to answer the following research question:

How do anxiety patients experience reflection of continuously measured GSR sensor data, with the purpose of understanding their anxiety and how can this be a potential part of anxiety treatment?

We found that the use of the sensor technology as part of anxiety treatment was highly dependent on each individual patient. The three main use cases for the technology were: Help patients recall episodes, help patients confirm episodes and help patients disconfirm episodes. Furthermore, considerations of how and when to introduce the technology in a treatment program were also an important aspect, as the technology might not be useful for everybody. For the technology to be useful it is necessary that the patient, potentially in collaboration with a psychiatrist, is able to subjectively assess the objective data collected through the sensors. This is necessary as the GSR sensor does not differentiate between different types of arousal.

By providing the patients with GSR cues collected throughout the day the technology can help support patients in recalling episodes because this can be difficult, as also seen in DRM studies in HCI. It can also help them confirm the episodes they do remember as they can sometimes be unsure whether it is anxiety or not and also because it is comforting to know that something actually does happen in the body and not solely in the mind. Opposite, the system can also help disconfirm episodes for patients who find they are having anxiety constantly. All three aspects can help patients better understand how and when they experience anxiety and possibly also learn why.

When introducing the technology into a treatment program, the psychiatrists need to get to know the patient beforehand to determine if the technology can be useful for them and with what purpose. For some patients who ruminate a lot and think everything is relevant, the technology might not be suitable as it might bring too much attention to the body and the anxiety. The psychiatrists referred to this as a double-edged sword as it was positive to understand and be connected to the body but too much attention could be harmful for some patients.

The patients who participated in our study were all able to read and understand the data through the GSR graph, however found the setup of the E4 Connect desktop application to be

too intricate. A mobile application was therefore designed based on the following follow-up research question:

How can we design tools that supplement this?

The idea behind the application was to make it more mobile and personalized. In the app, the patients are able to see solely the GSR graph from the wrist-worn sensor and can write down notes for each spike that appears. They are also able to write in a diary, set up goals, see a history of their data and reflect on their life. The app was evaluated with one of the patients from the field study.

Our findings show preliminary results that points to the potential of introducing sensor technology in anxiety treatment to support patients' self-reporting of episodes. Further evaluations with more patients is needed to understand more on how patients experience using a system for monitoring their anxiety.

9. References

1. Adams Z.W., McClure, E.A., Gray, K.M., Danielson, C.K., Treiber, F.A. & Ruggiero, K.J. (2017) Mobile devices for the remote acquisition of physiological and behavioral biomarkers in psychiatric clinical research, *Journal of psychiatric research*, vol. 85, pp. 1-14.

2. Barak A. & Grohol J. M. (2011) Current and Future Trends in Internet- Supported Mental Health Interventions, *Journal of Technology in Human Services*, 29:3, 155-196, DOI: 10.1080/15228835.2011.616939

3. Bauer, A.M., Iles-Shih, M., Ghomi, R.H., Rue, T., Grover, T., Kincler, N., Miller, M. & Katon, W.J. 2018, "Acceptability of mHealth augmentation of Collaborative Care: A mixed methods pilot study", *General hospital psychiatry, vol. 51*, pp. 22-29.

4. Beck, J. S. (2011). Cognitive behavior therapy basics and beyond. (2 ed.). Guilford Press.

5. Boukhechba, M., Chow, P., Fua, K., Teachman, B.A. & Barnes, L.E. (2018a), "Predicting social anxiety from global positioning system traces of college students: Feasibility study", *Journal of Medical Internet Research*, vol. 20, no. 7..

6. Boukhechba, M., Gong, J., Kowsari, K., Ameko, M.K., Fua, K., Chow, P.I., Huang, Y., Teachman, B.A. & Barnes, L.E. (2018b), "Physiological changes over the course of cognitive bias modification for social anxiety", *2018 IEEE EMBS International Conference on Biomedical and Health Informatics, BHI 2018,* pp. 422.

7. Brandt E. (2006). How Tangible Mock-Ups Support Design Collaboration. *Springer Science + Business Media B.V*.

8. Bruun, A., & Ahm, S. (2015, September). Mind the gap! comparing retrospective and concurrent ratings of emotion in user experience evaluation. In *Human-Computer Interaction* (pp. 237-254). Springer, Cham.

9. Bruun, A., Law, E. L., Heintz, M. & Eriksen, P. S. (2016) Asserting Real-Time Emotions through Cued-Recall: Is it Valid? *NordiCHI '16*.

10. Bry J., Chou T., Miguel E., Comer J. S. (2018). Consumer Smartphone Apps Marketed for Child and Adolescent Anxiety: A Systematic Review and Content Analysis. *Behavior therapy, 49(2), 249-261.*

11. Chen, C. V., George, S. A., Liberzon, I., 2017. Stress and Anxiety disorders. Book chapter In: *Hormones, Brain and Behavior*, p. 251–274

12. Doherty, G., Coyle, D., & Matthews, M. (2010). Design and evaluation guidelines for mental health technologies. *Interacting with computers*, *22*(4), 243-252.

13. Empatica, 2019 https://www.empatica.com/en-int/company/ , [accessed June 2019]

14. Gao, Y., Li, H., & Luo, Y. (2015). An empirical study of wearable technology acceptance in healthcare. *Industrial Management & Data Systems*, *115*(9), 1704-1723.

15. Garcia-Ceja, E., Riegler, M., Nordgreen, T., Jakobsen, P., Oedegaard, K. J., & Tørresen, J. (2018). Mental health monitoring with multimodal sensing and machine learning: A survey. *Pervasive and Mobile Computing*.

16. Gerken, J. and Reiterer, H. (2009) A Taxonomy for Longitudinal Research In HCI. *Mensch & Computer 2009: Grenzenlos freil*. Oldenburg Verlag.

17. Hassenzahl, M. (2001). The effect of perceived hedonic quality on product appealingness. *International Journal of Human-Computer Interaction*, *13*(4), 481-499.

18. Hollis, V., Pekurovsky, A., Wu, E., & Whittaker, S. (2018). On Being Told How We Feel: How Algorithmic Sensor Feedback Influences Emotion Perception. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *2*(3), 114.

19. Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N., & Stone, A. A. (2004). A survey method for characterizing daily life experience: The day reconstruction method. *Science*, *306*(5702), 1776-1780.

20. Karapanos, E., Zimmerman, J., Forlizzi, J., & Martens, J. B. (2009, April). User experience over time: an initial framework. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 729-738). ACM.

21. Kazantzis, N., Whittington, C., & Dattilio, F. (2010). Meta‐analysis of homework effects in cognitive and behavioral therapy: a replication and extension. *Clinical Psychology: Science and Practice*, *17*(2), 144-156.

22. Lang, P. J. (1995). The emotion probe: studies of motivation and attention. *American psychologist*, *50*(5), 372

23. Lazar J, Fend J. H., Hochheiser H. (2010). Research Methods in Human-Computer Interaction, John Wiley and Sons.

24. Levy, F., Leboucher, P., Rautureau, G. & Jouvent, R. (2016) E-virtual reality exposure therapy in acrophobia: A pilot study, *Journal of telemedicine and telecare*, vol. 22, no. 4, pp. 215-220.

25. Li, I., Dey, A. K., & Forlizzi, J. (2011, September). Understanding my data, myself: supporting self-reflection with ubicomp technologies. In *Proceedings of the 13th international conference on Ubiquitous computing* (pp. 405-414). ACM.

26. Mohr, D.C., Zhang, M. & Schueller, S.M. (2017) *Personal Sensing: Understanding Mental Health Using Ubiquitous Sensors and Machine Learning*.

27. Mørch, M.M. & Rosenberg, N.K. (red.) (2005): Kognitiv terapi – modeller og metoder. *Hans Reitzels Forlag*.

28. Omodei, M. M., & McLennan, J. I. M. (1994). Studying complex decision making in natural settings: using a head-mounted video camera to study competitive orienteering. *Perceptual and motor skills*, *79*(3_suppl), 1411-1425.

29. Poh, M. Z., Swenson, N. C., & Picard, R. W. (2010). A wearable sensor for unobtrusive, long-term assessment of electrodermal activity. *IEEE transactions on Biomedical engineering*, 57(5), 1243-1252.

30. Preece, J., Rogers, Y., Sharp, H., 2015. Interaction Design: Beyond human – computer interaction. Wiley

31. Redelmeier, D. A., & Kahneman, D. (1996). Patients' memories of painful medical treatments: Real-time and retrospective evaluations of two minimally invasive procedures. *Pain*, *66*(1), 3-8.

32. Rogers Y., Marshall P., (2017). Research in the wild. *Morgan and Claypool Publishers*

33. Sarker, H., Tyburskir, M., Rahman, M.D.M., Hovsepian, K., Sharmin, M., Epsteinr, D.H., Prestonr, K.L., Furr-Holden, C.D., Milam, A., Nahum-Shani, I., Al'Absi, M. & Kumar, S. 2016, "Finding significant stress episodes in a discontinuous time series of rapidly varying mobile sensor data", *Conference on Human Factors in Computing Systems - Proceedings*, pp. 4489.

34. Scollon C., Kim-Prieto, Chu, Diener E. (2003) Experience Sampling: Promises and Pitfalls, Strengths and Weaknesses. *Journal of Happiness Studies*, Vol.4(1), pp.5-34

35. Sketch app (2019), https://www.sketch.com [accessed June 2019]

36. Tolin, D. F. (2010). Is cognitive–behavioral therapy more effective than other therapies?: A meta-analytic review. *Clinical psychology review*, *30*(6), 710-720.

37. Whiteford, H. A., Degenhardt, L., Rehm, J., Baxter, A. J., Ferrari, A. J., Erskine, H. E., ... & Burstein, R. (2013). Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *The Lancet*, 382(9904), 1575-1586.

38. Winslow, B. D., Chadderdon, G. L., Dechmerowski, S. J., Jones, D. L., Kalkstein, S., Greene, J. L., & Gehrman, P. (2016). Development and clinical evaluation of an mhealth application for stress Management. *Frontiers in psychiatry*, 7, 130.