Designing Benchmark Application for Stress Elicitation.

Master thesis

In the study program

Human-Computer Interaction

Submitted by

Charles Dolapo Ajiboye.

Matriculation Number: [redacted].

On January 2nd 2023

at

University of Siegen, Germany

First Examiner: Prof. Dr. Claudia Müller

Second Examiners: Prof. Dr. Volker Wulf

Prof. Dr. Maria Maleshkova

Supervisors: M.A Richard Paluch

MSc. Florian Grensing

MSc. Jennifer Daniel Onwuchekwa

CONFIRMATION

Hereby I confirm that I have composed the present thesis independently. I only have used the sources and means specified in this thesis. Especially from the internet, I only have used the denoted references. I have taken note of the section in the examination regulations concerning attempts to cheat. I confirm that the electronic version of the thesis which I deliver is identical to the printed version with respect to the content. I agree that an electronic version of the thesis will be stored for purposes of inspection of plagiarism.

[redacted] [redacted]

Date (Signature)

Abstract

The design and development of digital tools to fast-track and automate research processes has been a topical issue in the world of research over the years. Stakeholders and research experts have been working assiduously to come up with superior ideas that will give rise to the development of cutting-edge technological tools for research and experimental purposes.

In this study, the focus is on how a digital tool that will serve as a benchmark App for inducing/elicit stress can be design with the use of User centered design principles.

To actualize this, a literature review was first conducted to understand the social practice and also identify some underlying issues that might exist. Thereafter, the design case study framework consisting of three phases (pre study, design and appropriation) was implemented all through the design process. The insights gathered during the qualitative interview phase and the user needs and pain points discovered during the user workshop phases were all implemented in the High fidelity prototypes that were generated for the final App design. The designed interfaces were then tested and evaluated by users using the thinking aloud method so as to discover usability issues that needed to be fixed for the improvement of the overall user experience of the product. In line with the results obtained during the evaluation, recommendations were given on how this work can be further developed to have a fully functional application that is appropriated and tailored to meet the specific needs of the users.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

EEG: Electroencephalography

HR: Heart rate

Sp02: Arterial oxygen level.

EDA: Electro-dermal activity

HRV: Heart Rate Variability

SRI: Stress Response Inventory

PSS: Perceived Stress Scale

SSRS: Stress Self Rating Scale

SCWT: Stroop color and word test

VR: Virtual reality

SC: Skin Conductance

ANS: Autonomous nervous system

SCL: Skin Conductance Level

SCR: Skin conductance responses

DCS: Design Case study

Lo-fi: Low fidelity Prototype

Hi-fi: High fidelity prototype

UI: User interface

HCI: Human Computer Interaction

UCD: User centered design

1. INTRODUCTION

One of the many difficulties that humans have to grapple with daily is stress. Over the years, the topic of stress is a psychological concept that has piqued a lot of scientific interest. (Parent, et al., 2020). The attention it has attracted could be a result of the effect it has on the performance or health of an individual. Numerous definitions for stress exist but according to Webster's New World Dictionary of American English (2000, 4th ed.) the word stress is derived from the Latin word *strictus* meaning "hardship, adversity or affliction". Later on, it was called became stresse in Middle English and estresse in Old French (Ramos & Sharma, 2022).

In the health sector, the advancement in technology has led to an increase in the number of medical instruments or devices used to monitor stress levels and capture physiological data from people. This advancement has given rise to the development of devices such as microelectromechanical, electromyogram, and electroencephalogram (EEG) based brain sensing platforms with sensors that enable the monitoring of various human physiological stressors (Park, et al., 2015). However, despite the advancement in health technology and investigations carried out for the determination and management of stress, at the time of conducting this research, no mobile App has been designed to serve as a benchmark for researchers in the academic and digital health sphere to be utilized in subjecting research participants to stress conditions, elicit stress while capturing physiological data such as Heart rate(HR), arterial oxygen level (Sp02), temperature, electro-dermal activity (EDA) and acceleration with the use of non-intrusive wrist-worn biosensors devices. This gap in research is what this thesis intends to fill in other to revolutionize and standardize the physiological and neurological data-capturing process for researchers.

1.1 Problem and Research Motivation

Wearable sensors or devices allow us to monitor different human physiological activators in an efficient, non-invasive, and inexpensive manner (Birjandtalab et al., 2016). . However, at the time of writing this thesis, there is currently no App or software which synchronizes physiological data gotten from the wearable devices with the software used to induce the stress in real-time. Furthermore, another research gap discovered in the current state of assessing and visualizing the neurological status of research subjects is that multiple software for the wearable devices will have to be installed on a computer system which therefore increases the touch points of the physiological data thereby reducing the quality of the data.

The motivation for designing this App is to create one platform on which researchers can create and simulate stressful activities (Physical, emotional, cognitive, and relaxing) that they want the participants to undergo, gather data from research participants, and then an initial look or assessment of the captures data all on the same App. This will be a massive improvement that will enhance and quicken the research process and also reduce the number of touch points that the captured data will have to go through. The accomplishment of this goal will also greatly expedite the neurological data capturing and analysis process for researchers in the digital health space.

This thesis work is predicated on the following research question:

How can a benchmark app be used for the elicitation of stress and capture of physiological data in the field of digital health and life sciences?

To provide an answer to the research question above, I have made use of the Design Case Study framework in the format below.

1. <u>Pre-study</u>: Gain an understanding of the work context and practices in the field of digital health while also identifying the pain point and needs of researchers in this field
   <u>Action</u>*: Conducted qualitative research (e.g In-depth interviews, focus groups, co-design workshops)*
2. <u>Design Phase</u>: Conduct a co-design workshop to generate design insights and ideas
   Action*: The sketching of both low and High fidelity Prototypes with the design Software Figma and Miro.*
3. <u>Appropriation Phase:</u> Utilized the design created in the design phase in a naturalistic environment while observing how users interact with it to see if it will affect their social practice.
   <u>Action</u>*<u>:</u> Conducted a thinking aloud workshop combined with usability testing to observe users' interaction with the design and identify usability issues.*

The main goal here is to simplify the data collection process of researchers in the digital health field with the mobile app and also facilitate the visualization and assessment of neurological status. This goal is intended to be actualized with the use of the design case study framework of Wulf et al., (2011) which comprises three phases: Pre-study, development, and evaluation.

1.2 Structure of Thesis

The following will be the format for the section of the thesis. Chapter 2 gives a literature review to establish the current state of the art in the research area of assessing and visualizing the neurological status of research subjects under different stress conditions. Chapter 3 will do a deep dive into the methodology implemented for this research work which is the Design case study framework. Subsequently, ideas that were generated and conceptualized for the App design will be recorded in Chapter 4.

Chapter 5 documents the App concept's implementation and the usability testing of the high-fidelity mockup with users. Chapter 6 contains an evaluation, discussion, and also reflection on the results obtained during the research while also stating some limitations and also recommendations for future work.

2 Background and State of the Art

Recent advancements in wearable sensors enable us to monitor several human physiological activators effectively, non-intrusively, and affordably. EEG signals are usually collected using surface or implanted electrodes to monitor and evaluate brain activity. EEG-based monitoring, however, is neither convenient nor comfortable during routine daily tasks.

Modern wearable devices can measure heart rate, skin temperature, and wrist movement. Since they are portable and simple to use, wrist-worn devices are appropriate health monitoring instruments for everyday use (Empatica, 2015). Additionally, the wrist can be used to monitor changes in skin conductance, which signify changes in the sympathetic nervous system (Poh M. Z., 2011).

This section of the thesis will review already existing literature on “Stress”, and the different categories while outlining stress types and stress levels measuring techniques. Additionally, it will analyze related studies that have attempted to use technological tools to elicit stress or to evaluate and visualize the neurological condition of research participants under various stress settings.

2.1 Background and Stress Theory

Stress is a broad and ambiguous concept with several definitions that are also attributed to different situations. (Shahsavarani, Azad Marz Abadi, & Hakimi Kalkhoran, 2015). In earlier studies, various definitions from various angles were put up to provide an overall explanation of stress but in reality, there is no widely accepted definition for it. Humphrey (2002) defines stress as a factor that makes people feel like it's difficult to adapt to and maintain homeostasis with their surroundings, both inwardly and externally. Stress is defined simply and succinctly in literature as any influence of a living thing's internal or external environment that upsets its equilibrium or homeostasis. (Ashayeri, Shahsavarani, Lotfian, & Sattari, 2013) .

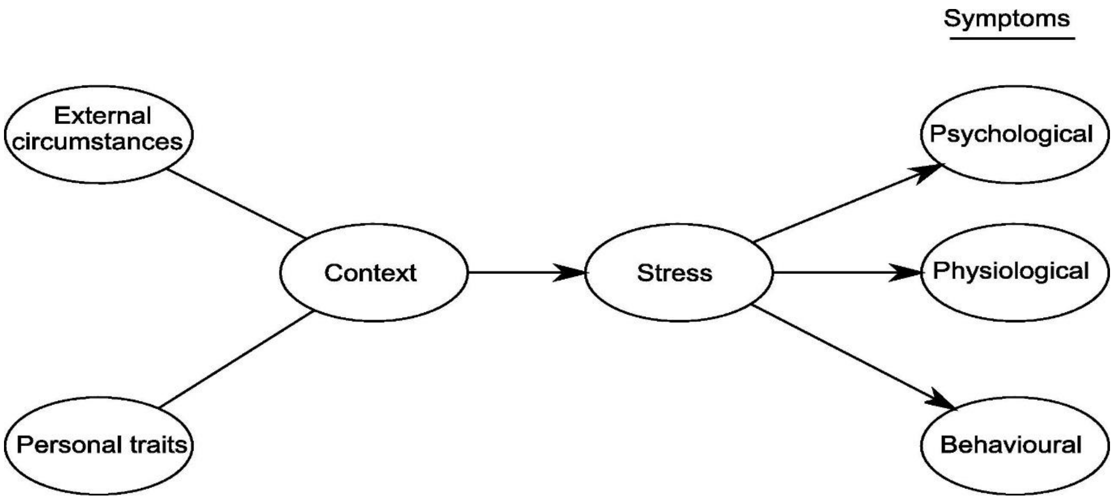


Figure 1.The multimodal nature of stress. (Carneiro et al., 2012.

From an extensive review of existing literature, it was discovered that stress could be defined according to:

- Stressor's nature (physiological stress and psychological stress)
- The impact on the individual (positive impact and negative distress)
- The length of time that the stressor was present (acute stress, which is short-lived, and chronic stress) (long-term)
- The stressor's characteristics or nature (physiological, psychological) (Shahsavarani, Azad Marz Abadi, & Hakimi Kalkhoran, 2015).

In line with the aforementioned definition of stress, it can be seen that stress is a typically individualistic experience that is dependent on the event or the particular psychological determinant that triggers a specific response (Fernand, 2007). A stressor is a source of stress that causes a certain set of reactions or responses in the body of a living creature. (Davis, Eshelman, & McKay, 2008). Any disruption, whether physical or psychological, that is thought to have the potential to endanger the stability of the organism's internal environment is referred to as a stressor (Bjorntorp, 2001). These stressors could be psychological challenges such as personality traits, hassles of life, or changes to our way of life. It can also be environmental demands such as temperature changes or increments in noise levels. On a larger scale, stress is unavoidable, yet chronic or acute exposure to stress has been linked to several health issues, including respiratory disorders, cardiovascular diseases, and autoimmune diseases (Schneiderman, Ironson, & Siegel, 2005). Aside from health considerations associated with stress, some researchers have also identified some links between stress and the performance of humans. It was discovered that stress has a negative effect on cognitive performance such as memory. High arousal, according to authors, may improve memory consolidation but impair memory recall (Wolf, 2009). According to a study, many workers are subjected to harmful levels of stress at work in different nations of the world (Jones, Latreille, & Sloane, 2016), which results in a substantial rise in absenteeism and intention to leave the company (Jamal, 2007). Work-related stress has been described as "the emotional, cognitive, behavioral and physiological reaction to unpleasant and noxious characteristics of work, work environments, and work organizations. High levels of alertness, distress, and frequently a sense of not being able to cope, characterize it (Mishra, et al., 2011). In the work environment stress is also associated with lower job performance (Jamal, 2007). Some findings in research have suggested that stress might be of great benefit in some scenarios. Researchers looked into the connection between stress (i.e., time constraint) and team communication using a crisis management simulation. From this simulation, it was discovered that stress can increase communication quantity and efficiency. However, it also discovered that frequent request for information is associated with poorer task performance (Pfaff, 2012).

2.2 Stress types and levels

The terms "eustress" and "distress," which refer to positive and negative stress, respectively, were distinguished by Selye (Selye, 1956). Positive adjustments or demands that are easy to deal with or adjust to the new environment are when eustress manifests. It may enable us to accomplish our objectives and boost productivity (Colligan & Higgins, 2006). Stress may be extremely destructive and have detrimental effects.

Furthermore, there are three levels of stress that can be distinguished depending on how long a person is exposed to stressors. Acute stress is the body's natural "flight-or-fight" reaction to brief exposure to stimuli and is not thought to be harmful (Bakker, Pechenizkiy, & Sidorova, 2011). When stressful events happen more frequently but then stop, this is known as episodic stress. It is linked to a life that is extremely hectic and stressful (Bakker, Pechenizkiy, & Sidorova, 2011). Finally, when stressors are enduring and persistent, such as family issues, work stress, or poverty, chronic stress which is the most detrimental, occurs (Colligan & Higgins, 2006).

To stop stress from reaching its peak and to lessen the risks, it is possible to identify and manage it while it is still acute or episodic (Sharma, Dhall, Gedeon, & Goecke, 2014). Every aspect of an individual's life, including emotions, actions, cognition, and physical health, can be impacted by stress. The body as a whole is not immune to stress, but because everyone responds to stress differently, there may be a variety of stress-related symptoms.

For the purpose of this research and to properly elicit stress, stress will be categorized into three groups:

- Physical Stress
- Emotional Stress
- Cognitive Stress

2.2.1 Physical stress

Physical stress is described as a force transferred to a specific area of biological tissue or as a reaction to stresses and demands from the environment (Azher, Noushad, & Ahmed, 2014).

Such physical stress is experienced when the demand on an individual exceeds the resources contained in the body to measure up with demands. Some level of physical stress is considered to help boost the immune system, it has been proven in research that it can help improve the rate at which the heart works and protect your body from infection (Azher, Noushad, & Ahmed, 2014)

Generally, physical stress is categorized into two:

- Positive Physical Stress
- Negative Physical Stress

2.2.1.1 Positive physical stress

An action that encourages positive striving and emotions, which results in good health, excites people, and yields favorable results is known as "positive physical stress" (Nelson & Simmons, 2003). Activities such as lifting weight, jogging, and motivation to take up new challenges are considered as positive physical stress. In the study of stress, the differentiating factor between positive stress and negative stress is considered to be subjective. A certain level of excitement is expected to be involved before stress is termed positive physical stress. While some stress activities might be seen as scary and overwhelming for some individuals, for some others it might be a fun activity. Shown in figure 2 are some effects and benefits of positive physical stress.

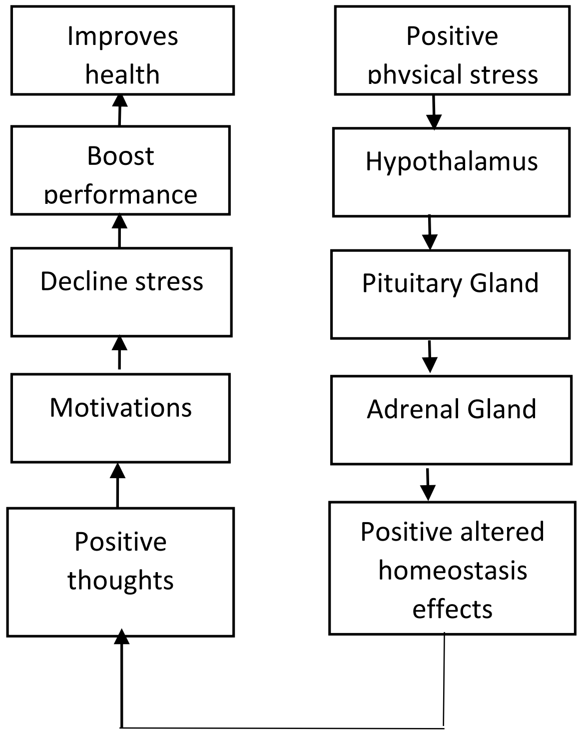


Figure 2. Positive Physical Stress. Taken from (Azher et.al., 2014)

2.2.1.2 Negative Physical stress

Negative physical stress is defined as any physical stressor that activates the hypothalamic pituitary adrenal (HPA) axis for an extended period of time and causes a surge in the stress hormone levels that cause tension and may result in illness, pain, unpleasant feelings, anxiety that impairs performance, and overexertion (Ahmed, Noushad, & Azher, 2014). In such a situation, the blood level and adrenaline of the individual is increased to a higher level over an extended period. Failure to put this in check might lead to complications such as burnout or panic attacks. Some other health conditions such as depression, gastrointestinal diseases, high blood pressure, heart diseases, and fatigue have been associated with chronic negative physical stress that lasts for an extended period of time. Shown in figure 3 are some of the effects of negative physical stress.

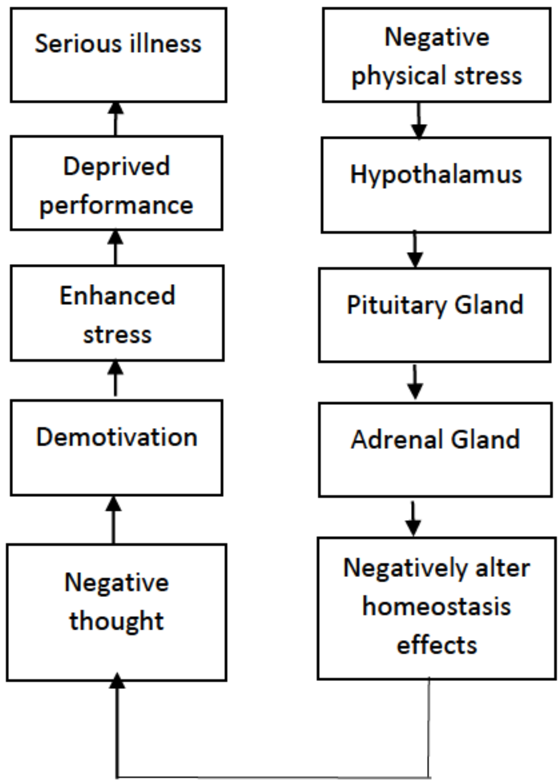


Figure 3. Negative Physical Stress. Taken from (Ahmed, Noushad, & Azher, 2014)

2.2.2 Emotional Stress

Emotion is considered a basic characteristics of human behavior (Gendron, 2010) and it plays an important role in communication. In the field of Human Computer interaction and psychology, emotion has been subject to extensive research (Beale & Peter, 2008). In recent times, researchers have explored different techniques and methods to elicit stress (Jallais & Gilet, 2010) (Mardini, et al., 2018). These techniques used to elicit stress are called emotion elicitation techniques in research local parlance. These often-employed strategies or tactics include thinking back on past experiences and using visual, auditory, or other forms of stimulus (Braun, Weiser, & Pfleging, 2018) (Mauss & Robinson, 2009). Compared to other types of stress, emotional stress can be more intense and painful (Freshwater, 2018). In general, different stress exposures, such as significant life events, persistently stressful situations, and everyday difficulties, can cause such stress. Certain objective characteristics of a stressor affect how likely it is to cause emotional stress. In most cases, emotional stress originates from stressors that are out of the control of an individual (e.g., a death) and these might affect central aspects of an individual's life (Dohrenwend, 2000). Heart disease and other physical health issues have also been linked to emotional stress (Brotman, Golden, & Wittstein, 2007). In order to reduce the negative effects emotional stress has on the health, various psychosocial stress management interventions have been developed. These interventions are developed to enhance positive coping methods, exercise, relaxation techniques usage, and cognitive strategies for managing stress. Several stress management techniques have been demonstrated to improve emotional and physical consequences (Blumenthal, et al., 2005).

2.2.3 Cognitive Stress

Cognitive stress is defined as feelings of fear, anxiety, persistent worry, and racing thoughts. (Calvo & Gutiérrez-García, 2016). In general, it is referred to as any stress that has a negative impact on the mind or brain of an individual. Stress is something that is often felt and it is normal for it to have some cognitive component in it. In contrast, neurological disorders are what cause strains on the nervous system. One of the most significant factors that could affect mental health is stress. Because the brain responds to a stimulus like a muscle, several studies have also revealed that cognitive stress may have long-term impacts on how the brain works. The amygdala, which regulates your alertness to danger, is used more intensely than other sections of the brain when you're anxious (Walker, 2021). They seem to be working out while other areas are getting less attention (Harvard, 2021).

2.3 Measurement of Stress

For many years, health researchers have looked for a single biological sign that someone is "under stress." However, there is no single stress-specific biomarker (Crosswell & Lockwood, 2020). This result was reached because acute stress is not always the only state that causes physiologic changes, such as an increase in blood pressure and heart rate. Some other existing acute stress states such as the feeling of excitement, and exercising also evoke biological responses which are considered to be somewhat similar to those triggered by negative acute stressors such as blood pressure and an increase in heart rate. There are nevertheless compelling reasons to include these biomarkers in research studies on stress and health even though evaluating stress-related biomarkers may not provide a perfect prediction of whether someone is under stress or not (Crosswell & Lockwood, 2020). The following part of this section will elaborate on various stress measurement techniques that have been explored by different researchers under varying settings, situations, and conditions.

2.3.1 Physiological measures

Numerous research centered around detecting physiological stress in different conditions have been conducted. Most of these research works were conducted in laboratory settings which offers the possibility for both the context and the stressor (duration, timing, frequency) to be strictly controlled (Kusserow, Amft, & Tröster, 2012). Numerous possibilities are opening up for continuous, ambulatory stress monitoring as a result of the widespread use of wearable technology, and during the past five years, research in this area has grown significantly

(Smets, De Raedt, & Van Hoof, 2018). For example, in a setting that is controlled the possibility of physiological signals to detect stress has already been demonstrated (Smets, De Raedt, & Van Hoof, 2018). Numerous stress-inducing stimuli, such as the Stroop color word test, mental arithmetic, public speaking, computer work, or a cold pressor test, have been employed (Karthikeyan, Murugappan, & Yaacob, 2011). Physiological responses are non-voluntary movements or responses that are very difficult or impossible to identify by external observation since they are a natural aspect of the operation of a living organism or bodily part (Physiology, 2015). An advantage with the Physiological measures of stress measurement is that they are objective and can be taken in real time over a long period of time without interruptions. As a result, physiology-based models have been developed in a number of recent research, sometimes with a somewhat high detection accuracy (Smets et el., 2019). Some common physiological parameters used in measuring stress levels are heart rate variability, respiration both rate and volume, blood levels, and EEG (electroencephalogram) (Reisman, 1997). Furthermore, when measuring stress a person's internal emotional experience can be described in terms of its intensity and quality using physiological signs (Zimmermann et al., 2003). Although considered a trustworthy method for determining the physiological reaction to stress, hormone testing has several limitations. Some physiological cues, like Heart Rate Variability (HRV) and salivary cortisol levels, have correlations that have been demonstrated (Seo & Lee, 2010), while others still require more study.

2.3.2 Psychological measures

Psychological measure of stress do not require carrying out an action; rather, they are viewed as or connected to the mind or mental activity (Psychological, 2015). Generally, psychological stress can be assessed by self-report questionnaires or by having a psychologist interview you. The former is regarded as a dependable strategy and is one of the most popular ways to gauge how stressed out people are. Examples include the Stress Response Inventory (SRI), the Perceived Stress Scale (PSS), and the Stress Self Rating Scale (SSRS) (Hayashi et al., 2012). However, these questionnaires only provide information on the user's current levels of stress; they do not provide information about the sources of that stress or how those levels have changed over time. These tests can be performed occasionally, but they might not be able to pick up on little changes that might point to a serious issue in its early stages. In reality, they are only administered when the patient or those around him realize or suspect the seriousness of the condition, which in the vast majority of cases is too late. In addition, questionnaires demand the user's complete attention because they are subjective. We are not always aware of our true stress levels, and techniques like self-report questionnaires may occasionally result in an inaccurate stress level measurement because "people can suffer lapses in recall concerning the emotional tone of a day in as little as 24 h" (McDuff et al., 2012).

When carrying out psychological test for stress measurement, the reaction of the research subject is a vital indicator used by the research to determine the stress level of the participants. Some psychological reactions could include the escalation of strong negative feelings like anger, anxiety, irritability, or depression (Milczarek et al., 2009) and can also heighten our emotional reactions, making us feel more stressed, irritable, and hostile, which can have an impact on our interpersonal interactions (Peternel et al., 2012).

2.4 Stress Elicitation techniques

Selecting the most effective strategy to elicit stress is one of the most difficult aspects of studying it. Some researchers find it convenient to elicit stress in a laboratory and ambulatory settings while others think it is more appropriate to make use of technological tools in form of wearable devices to capture neurological data from humans when eliciting stress.

From a human-computer interaction perspective, the user's stress is more likely to be moderate in intensity and mental in nature (as opposed to physical, intellectual, emotional, and perceptual stress) (Zhai & Barreto, 2006). However, in recent times further research has been conducted with the purpose of developing techniques and technological tools that can be used to elicit stress, capture data from research participants and then analyze the data to fully understand or assess the neurological status of an individual under varying stress conditions. Common techniques that have been used are the Montreal Imaging Stress Task, Trier social stress test(TSST), stroop color word test(SCWT), viewing videos with stressful content, virtual reality, Augmented reality and machine learning techniques just to mention a few are some of the techniques that have been used to achieve this. This part of the thesis will elaborate on some of the aforementioned techniques, investigate their shortcomings and then outline some of the areas in which this work will build upon to further contribute to the field of human computer interaction and digital health.

2.4.1 Computer based Stress elicitation

With reference to a stress elicitation experiment carried out by Zhai and Barreto, where a computer based interacting environment was established to elicit moderate mental stress and allow participating research subjects experience a similar stress effect at predetermined times. (Zhai & Barreto, 2006). To achieve this, a computer game which is based on the classical Stroop color and word test (SCWT) was developed to trigger mental stress as the participating subjects operate the computer. SCWT is a neuropsychological test majorly used in clinical and experimental settings. The test demands that the colour font of a word designating a different colour be named. This technique is typically preferred by researchers doing stress elicitation experiments as a psychological or cognitive stressor to elicit strong emotional reactions and increased levels of physiological responsiveness, notably autonomic reactivity (Renaud & Blondin, 1997). Some other researchers consider it as an effective method in assessing the ability to inhibit cognitive interference that occurs when the processing of a specific stimulus feature impedes the simultaneous processing of a second stimulus attribute (Scarpina & Tagini, 2017).

While conducting the experiment, the stroop test was embedded into an interactive interface which requires the research subject to click on the displayed button that shows the right answer. The right answer is shown on one of the five buttons displayed on the screen. The participants are not required to state it verbally. To increase the physiological stress reactions, task pacing was added to the stroop test by only waiting a maximum of 3 seconds for a user to feedback or response (Renaud & Blondin, 1997). Within the 3 seconds time, if the participant cannot make a decision, the screen will change to the next trial. Shown in figure 4 is a typical example of the test interface for stroop test conducted on a computer.

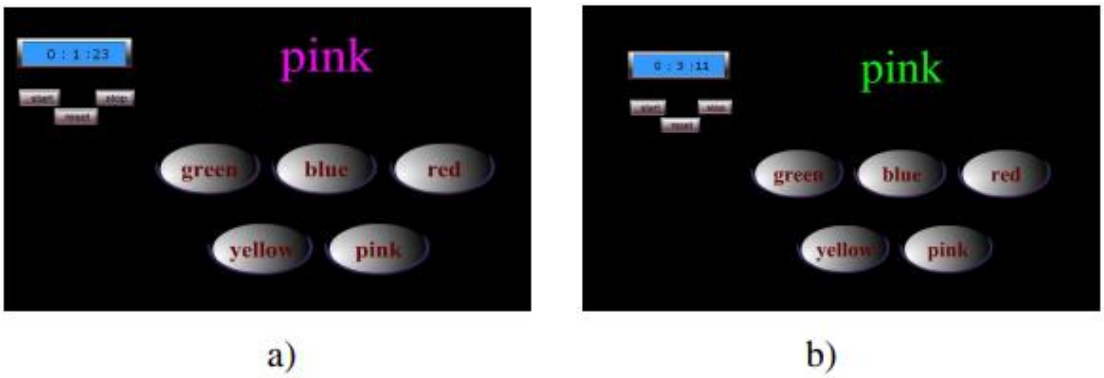


Figure 4. Sample Stroop trials: (a) Congruent: “pink” in pink; (b) Incongruent: “pink” in green (Taken from (Barreto, Zhai, & Adjouadi, 2007)

2.4.2 Stress elicitation in laboratory & ambulatory settings

Due to the varying effect stress has on humans, researchers have given attention to the subject in a bid to propose models that will elicit and detect stress from a physiological perspective. Smets et al.,(2019), conducted a review of 25 papers with the purpose of investigating the research question. In accordance with (Smets et el., 2019) description, a larger number of the reviewed papers has laboratory settings as their focus, on the other hand, a few of the research works attempted to detect and elicit stress in an ambulatory settings. For example, a recent research work conducted in a laboratory setting made use of portable wristband to record heart and electro dermal activity in order to detect affective stress triggered by Trier social stress task. This research reported to have achieved an area under the receiver operating characteristics curve of 0.87 (Ollander, Godin, Campagne, & Charbonnier, 2016). In a similar fashion, another research work made us of ECG and EEG in order to classify the affective state of individuals playing a survival horror game (Vachiratamporn, Legaspi, Moriyama, & Numao, 2013). In an ambulatory setting, some recent examples were analyzed where portable sensors such as (wristband, chest strap) were used to elicit and detect stress in ambulatory settings (Hovsepian, et al., 2015) (Gjoreski, Gjoreski, Luštrek, & Gams, 2016). It was discovered that the models reached a level of 0.72 coefficient or 76% classification accuracy at detecting self-reported stress (i.e. two class classification) (Parent, et al., 2020). Despite the advancement made in ambulatory method of stress detection and elicitation, there still exist numerous challenges on the way to achieving a robust stress detection model. In line with Smets suggestion, the most conspicuous limitation to stress detection model is physical activity and movement (Smets et el., 2019). Some models do not incorporate the capability of stress prediction when physical activity is identified in their setup (Hovsepian et al., 2015), which biases error rate metrics and makes it difficult to detect stress during physical exercise. In exchange for manual input, the configuration for some other models can accept contextual data, such as physical activity, enhancing accuracy (Gjoreski, Gjoreski, Lutrek, & Gams, 2016). From the aforementioned research scenarios, it can be deduced that several factors differ for each study which makes it difficult to capture the performance of each model investigated in the studies. The first challenge is that it is allowed to use different stressors. Secondly, the models make use of varying physiological modalities and in some other settings varying combinations of modalities. Another hurdle is that models differ in their classification scheme and classifications levels. This difference could be within participants, or between participants.

From observations in the study, it was discovered that a good number of the studies favor models that differentiate between a state of rest and a stressful activity or task. However, a good number of the studies use regression models to measure stress levels or classify the multiple levels of stress into (low, medium, high) (Hovsepian, et al., 2015). Irrespective of the setting i.e. either Laboratory or ambulatory, movement of body parts or physical activity in general are known to affect physiological measures in three different ways.

1. Physical activity has a strong influence on physiological measures
   (When a person performs physical activity, the body triggers a series of physiological mechanism to shift from a rest state to an active state. A rise in heart rate brought on by physical effort makes it clear (Bernardi, et al., 1996). The heart must beat more quickly in order to provide more nutrients to muscle cells, obtain more oxygen, and expel CO2 from the lungs since physical activity requires energy (Parent, et al., 2020).

2. Due to a shift in psychological states when an individual engages in physical activity, physiological measures will be greatly influenced. Physical exercise, such as lifting boxes, has been shown to deplete mental resources (DiDomenico & Nussbaum, 2008).There is also proof that physical activity reduces long term stress and anxiety (Pedersen & Saltin, 2015).

2.4.3 Stress elicitation through Gamification

Video games have been used for some years in the field of psychophysiological studies (Carroll, Turner, & Rogers, 1987). However, the potential of video games as stressors has not been fully explored. According to a study, playing video games can have the same stress-inducing effects as performing the Trier social stress task (Guitard, et al., 2010). Some other studies are of the opinion that they possess the potential to stress research subjects differently from the commonly known stressors which has given rise to the call for more research to be conducted on the subject (Porter & Goolkasian, 2019). In an experiment conducted by Parent et.al., stress was modulated by switching between two video games which were a non-stressful one serving as control condition and a stressful one (Parent, et al., 2020). The game which was non-stressful is called Timeframe. Timeframe is a readily available puzzle/exploration game developed by Random Seed Games (Random Seed Games, 2015). The stressful game used for the experiment was Outlast. Outlast is also a readily available survival game developed by Red Barrels (Red Barrels Games, 2013). For the Timeframe game, the players are required to move around an abandoned city and explore the ruins while also looking out for artifacts. Some elements which made timeframe suitable as a non-stressful game are the following: Firstly, the person in the game cannot be killed which means there is no threat involved in the game, secondly the environment is peaceful and well illuminated with soft background music and lastly the players will be informed that the number of artifact they are able to find will not count towards getting any point. This helped in decreasing stress for the players. In the case of the stressful game Outlast, the goal in this game is to move around a creepy asylum and evade some dangerous inmates in the vicinity. In this game, the players do not fight. The task is to avoid and get free form enemies. Some elements or features in the game that increases stress are a horror environment with eerie sound and music, poorly lit areas in the game which requires the players to make use of night vision mode. The room where the experiment was conducted was also poorly lit to increase stress. Both games were played with the Xbox controller one. Figure 5 and 6 below show images of the game scenarios

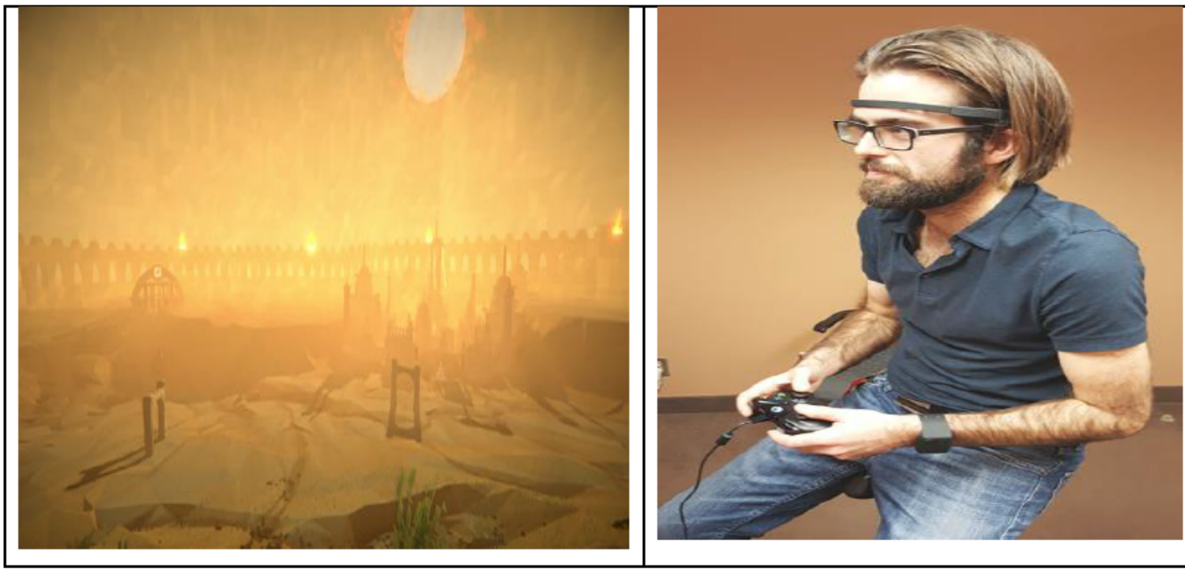


Figure 5. Time frame game environment (left). Player with console and wearable device (right).

Figure 6. Outlast game scenario showing normal (left) and night vision mode (right)

2.4.4 Stress Elicitation in Virtual Reality (VR)

A computer method called virtual reality (VR) generates an artificial environment with realistic sounds, visuals, and other feelings (Cho, et al., 2017). It entails simulating a realistic three-dimensional environment with the use of generated realistic images, sounds and other sensation (Blascovich, et al., 2002). In general, VR technology produces environments that are somewhat reminiscent of the actual world, and as a result, it finds use in a variety of industries. For instance, in the realm of medicine, virtual reality (VR) equipment are utilized for the treatment of post-traumatic stress disorder (PTSD) as well as social adaption training for social phobias (Rothbaum , et al., 2001). Researchers also make use of VR devices in their experiments to observe the corresponding responses they get from the created environments (Jönsson, et al., 2010) (Crifaci, et al., 2012) (Bullinger, et al., 2005)

In recent times, the advancement in virtual reality technology has given rise to questions such as "Can studies which involve the human subjects be carried out in virtual reality"? (Makela, 2020). However, Rivu et al., (2021) in his research proved that conducting studies in VR brings some potential benefits. One of such benefits is that researchers gain full control of the virtual world with the possibility of modifying the environment, the conditions and the variables at will (Rivu, et al., 2021). Furthermore, conducting research in VR also gives researchers a high degree of freedom as it can be conducted remotely (Mäkelä, et al., 2021).

2.4.4.1 Emotional Stress elicitation in Virtual reality

In the field of psychology, emotions have received a considerable attention for more than a century (Rivu, et al., 2021). The use of VR to elicit stress and also induce emotion was explored by (Estupinan, et al., 2014). In this pilot study, an investigation was done to on how valence and arousal of emotional stress in VR correlate with the real world. From their findings it can be seen that arousal is higher in virtual reality when images are used to elicit emotional stress (Estupinan, et al., 2014). The strengths of VR as a tool for investigating emotions was further identified by Chirico et al., (Chirico, Yaden, Riva, & Gaggioli, 2016). In this work, the exploration of how an intense emotion such as '*awe'* was elicited in VR was done. Their findings favors VR as a great tool which allows researchers generate a wide range of stimuli to elicit emotions which are ordinarily difficult to actualize in the real word (Chirico, Yaden, Riva, & Gaggioli, 2016). In the field of Human Computer Interaction (HCI) emotional stress elicitation is an under research area (Rivu, et al., 2021). To investigate further in this topic, Rivu et al.,(2021) conducted an exploratory assessment of several emotional stress elicitation methods in VR. The goal for the study was to have an in-depth understanding of how various elicitation methods can influence a range of emotions in VR. To carry this out, a mixed design user study (N=39) was conducted in which participants were exposed to the major four emotion elicitation methods (audio, video, image, autobiographical recall) as well as the major four basic emotions (sad, happy, excited, and angry) in identical virtual and real-world settings. The findings from the research are stated below:

- Emotional stress elicitation methods work similarly in both the real world and VR
- The four elicitation techniques (audio, video, image, autobiographical recall) had a very small disparity between the real world and the virtual world.

- Emotions you elicit don’t fade out immediately, which shows it has a minimum waiting period. (Rivu, et al., 2021)

This understanding is useful for our research because it gives insights into the differences between emotional stress elicitation methods and also provides evidence on how various emotions can be elicited with a technological platform as VR. Figure 7 depicts a stress elicitation experiment that was conducted in a lab setting using virtual reality technology

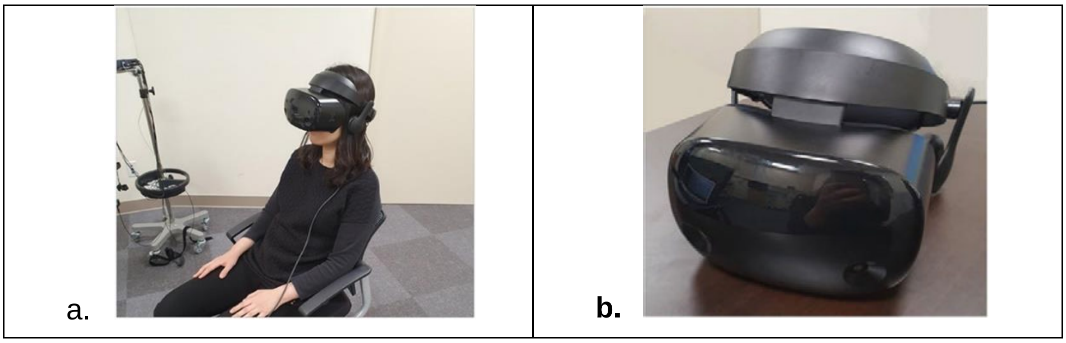


Figure 7. (a) VR relaxation session. (b) VR device.

2.5 Assessment of Neurological Status with Wearable Devices

Due to their utility for widespread healthcare monitoring, wearable gadgets have gained interest over the past 20 years (Wang et al., 2020). (Peake, Kerr, & Sullivan, 2018).With wireless connectivity, these devices enable low-cost, disposable, long-term and self-monitoring solutions to be deployed at home and hospital settings (Chen, et al., 2019). Ailments and conditions that could previously only be observed at significant expense and the subject's discomfort can now be tracked effectively and affordably because of the spread of wearable sensors. However, surface or implanted electrodes are still needed to evaluate a person's neurological condition. Such EEG-based monitoring is at best, obtrusive and uncomfortable. Generally, wearable devices make use of technology to monitor the physiological response of humans in their daily lives. The wearable devices have sensors embedded in them which when in contact with the skin have the possibility to gather physiological data such as heart rate (Schüll, 2016). In order to gather these physiological data in real time, wearable devices are often paired with application which provide real time information to the user (Li, et al., 2016). Currently, wearable technology may be found in the form of watches and garments, that can link to smartphones through Bluetooth and are frequently used in conjunction with more complex applications that gather, process, and integrate physiological inputs (Rodrigues, et al., 2022). Two frequently used sensors to capture physiological data from research participants when placed under stress conditions are Affectiva Q Sensor and Nonin 3150 Wireless WristOx2 Oximeter devices.

- Affectiva Q Sensor

  The Affectiva wearable device was initially created for effective communication with autistic people (Vitug, et al., 2014), and it is now utilized to identify changes in arousal, tension, or excitement in environments other than a lab or a computer, which are typically found when the participant is being exposed to stressful conditions (Picard, 2009). In practice, the device transmits the captured data to the software Q live where it is read. The Sensor is quiet easy to install and it works in connection with a laptop via Bluetooth connection. Just like a wristwatch is worn so is the sensor also worn. It is made to be small and comfortable thereby reducing the chances of distraction during the physiological data collection process (Kristina, 2022). The Sensor is often used to capture Physiological signals such as Electro dermal activity (EDA), temperature and acceleration (ACC). Below in Figure 7 is an image of the Affective Q Sensor.

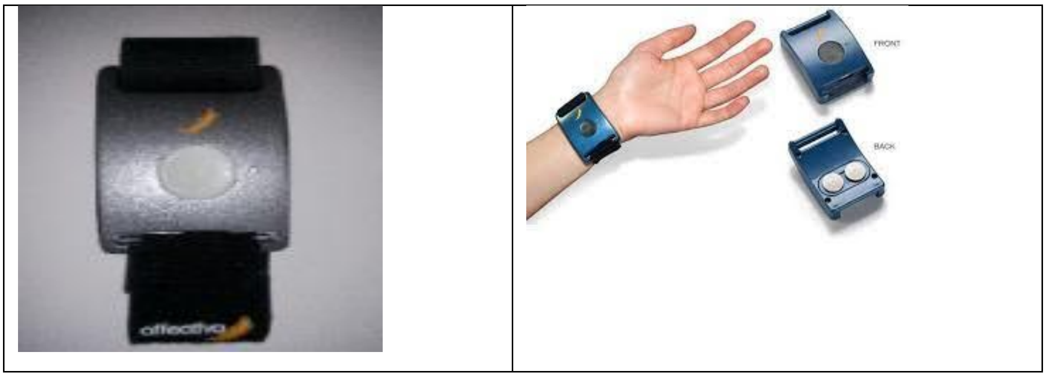


Figure 8. Affectiva Q Sensor Unit.

- <u>Nonin 3150 Wireless Wrist0x2</u>

  The Nonin 3150 Wireless wrist0x2 is a compact, lightweight pulse oximeter clip device frequently used to measure heart rate (HR) and peripheral capillary oxygen (Sp02). With the option to be turned off, this device may do discrete and continuous measurements while still maintaining Bluetooth communication and battery life (Masera & Tavakoli, 2020). It is often used to capture physiological signals such as heart rate (HR) and Sp02

Figure 9.The Nonin 3150 wireless wrist0x2

To effectively ascertain the neurological status of a research subject, the following physiological signals are measured with the use of wrist worn biosensors.

- Heart rate(HR) and Heart rate variability (HRV)
- Acceleration
- Electro dermal activities (EDA)
- Arterial Oxygen level(Sp02)
- Skin Temperature

The section of this thesis below will further expatiate on the viability of determining a person's neurological condition based on resulting physiological changes that may be efficiently and comfortably observed over time.

2.5.1 Heart Rate (HR) and Heart Rate Variability (HRV)

Heart rate variability is defined as the ability of the heart to respond to a varying degree of physiological and environmental stimuli (Suri, 2006), or in other words, the minor variations in the time between your heartbeats. Heart rate is defined as the rate of occurrence of cardiac beats in a specific period of time, typically expressed in beats per minute (Kranjec, Begu, Gerak, & Drnovek, 2014). The interval between beats is just slightly altered, changing by a fraction of a second. The normal physiological state of an individual consist of a variable heart rate and this parameter can be influenced by the following factors age, gender, psychological and physiological conditions (Sztajzel J. , 2004) (Reed, Robertson, & Addison, 2005). As a result the parameter's value can differ greatly for specific group of patients (Kranjec, Beguš, Geršak, & Drnovšek, 2014). It is crucial to note that the electrocardiogram (ECG) is the gold standard for determining HRV (McDuff, Gontarek, & Picard, 2014).

The ECG captures the electrical activity the heart produces on the body's surface (Berbari, 2000).Due to the fact that it immediately reflects cardiac activity, which is unmistakably impacted by Autonomic nervous system (ANS) changes, it is one of the most often employed signals in stress detection research (Asai, 2008).

Understanding the circulatory response to bodily stress and measuring and monitoring both physiological and mental stressors are both made possible by HRV (Hinde, White, & Armstrong, 2021).

2.5.2 Acceleration

An accelerometer is a tool that precisely measures acceleration. (Tinder & Richard, 2007) The proper acceleration of a body is its own instantaneous rest frame's acceleration (rate of change of velocity). For the extended monitoring of daily stress, a wearable electrocardiograph (ECG) with a tri-axis accelerometer (x, y, and z-axis) was created. It is made up of an amplifier, a microcomputer with an AD converter, a memory card, and a tri-axial accelerometer (Okada, Yoto, Suzuki, Sakuragawa, & Sugiura, 2013).

2.5.3 Electro dermal Activity (EDA)

According to a person's psychological state, changes in the electrical characteristics of the skin are measured as electro dermal activity (Geršak, 2020). In other words, it is defined as a measure of human sweating. EDA is considered one of the most promising noninvasive method that has been widely used for detecting stress and emotions (Rahma, et al., 2022).

EDA is a relatively straight forward, low cost and non-complex method used for stress studies. In its most simplified version, EDA is typically a measure of an individual's sweating. Sweating or sweat is a function of either a person's psychological condition or the body's thermal regulation mechanisms (Gerak, 2020). Enhanced sweating results in altered skin properties, such as increased conductance and decreased resistance, when there is emotional arousal, an increase in mental effort, or physical exertion (Peternel, Pogačnik, Tavčar, & Kos, 2012), (Asai, 2008). An attempt to understand the relationship between stress and EDA conducted by (Sanders & Cairns, 2010) depicts that there is a relationship between the two of them. This study shows that high levels of stress are associated with high EDA values and a high level of stress is an indication of high level of challenge, frustration and excitement (Vitug, Toplis, & Geli, 2016).

Physical exercise and stress are both known to have an impact on general perspiration, which can be detected by electro dermal activity (Parent, et al., 2020). Perspiration or sweating as it is commonly called emanates from our body thermal control processes or from out a person’s current psychological state. When an individual is psychologically aroused, excited his/her EDA increases (Geršak, 2020).

According to Boucsein, EDA can be measured in several forms (Boucsein, 2012). It can for example be expressed in skin conductance(SC) units if the electrical current is constant and this can be sub divided in to two aspects which are tonic phenomena i.e. Skin Conductance Level (SCL) and phasic phenomena which is skin conductance responses (SCR) (Eikenhout, 2017). Aside the electro dermal level (amount of sweat on the skin, EDA can be further described in details by analyzing the electro dermal responses. These reactions involve short-lived "peaks" of sweat that happen in response to a stimulus (Parent, et al., 2020). The stimulus can be specific which has a connection to an event or nonspecific.

The measurement of electro dermal activity is usually done at locations with high density of sweat glands such as palm, finger (Choi et al., 2012) or feet (Healey & Picard., 2005). Some other alternative locations used for the measurement of EDA signals are the wrist or the torso of the research subject (Schmidt, et al., 2018). As regard the monitoring of EDA, it can be done in a controlled laboratory or an environment outside a laboratory. In the case of a stable laboratory setting, the research subjects are required to carry out task while being stationed (e.g. in front of a Computer) while still under a monitored environmental condition. EDA is frequently employed in clinical settings and applied psychology for research on stress, pain, and human emotions (Poh, Swenson, & Picard, 2010) (Dindo & Fowles, 2008) (Benedik and Kaernbach, 2010).

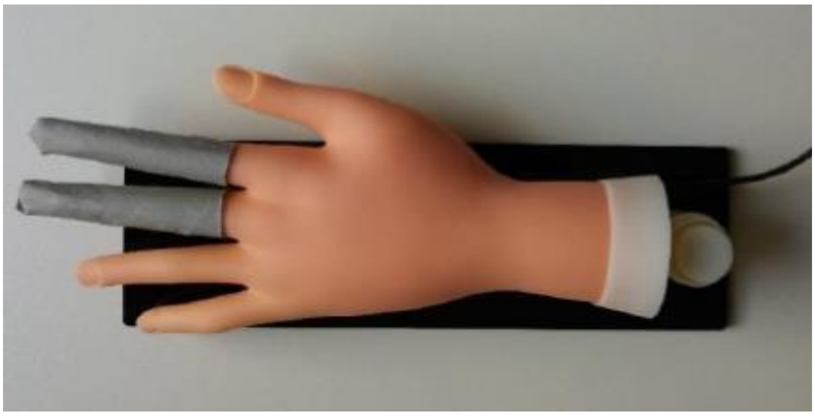


Figure 10. EDA simulator in human hand form with active electrodes on index and middle finger. Taken from (Geršak, 2020)

2.5.4 Oxygen Saturation (Sp02)

The proportion of the maximum amount of oxygen that the blood can carry is known as oxygen saturation. Peripheral capillary oxygen saturation is referred to as SpO2, and the Sp02 reading represents the expected amount of oxygen-saturated blood (Christian, 2018). The typical SpO2 range for a healthy person is 96% to 99%. What is deemed typical for a particular person may change due to high altitudes and other circumstances. The blood can be examined in a variety of ways to make sure its oxygen levels are normal. The most typical method is to measure the blood's SpO2 levels using a pulse oximeter. The usage of pulse oximeters is generally simple, and they are widely used in homes and healthcare settings. Light sensors are used by pulse oximeters to record the quantity of blood that carries oxygen and the quantity that doesn't carry. Due to the fact that physical exercise can significantly alter oxygen consumption and breathing rate, SpO2 readings can also be used to distinguish between psychological stress and physical activity (Christian, 2018).In preliminary study conducted by Wibawa (Wibawa & Purnomo, 2016) to fully understand the physiological pattern of human emotion (sad, happy, angry) with the use of Sp02 sensor which monitors the level of oxygen in the blood. Thirty (30) healthy participants took part in the investigation, and they underwent the baseline, video stimulation, and recuperation stages of the trial. While obtaining the behavior of oxygen level during video stimulation during these phases with the Sp02 sensor, it was discovered that 78% of the volunteers showed conspicuous change in lowering oxygen level during sad emotion. This shows that during sad emotions have a greater impact on lowering the level of oxygen in comparison to happy and angry emotion. The direct correlation between changes in blood oxygen levels and the length of video stimulation constitutes a study constraint. The longer the sad video stimulation is used to elicit emotional stress in the participants, the lower the level of oxygen that will be obtained.

2.5.5 Skin Temperature

Skin temperature is the measurement of the temperature of the body's outermost layer of skin, and it typically ranges between 33.5 and 36.9 ^{o}C (Arsalan, Anwar, & Majid, 2022). The ability of an individual to have a hot or cold sensation is dependent on the level of energy from or to the skin and the temperature of the skin is also dependent on the temperature of the air and also the time an individual spent in the environment. McFarland in (McFarland, 1985) discovered that skin temperature is connected to stressful and anxious conditions. Measurement of skin temperature can be done at different parts of the body such as the face, arm, finger and armpits and the measurement of temperature at these parts of the body gives different results under stress because the temperature on some parts of the body increases where on some other parts of the body the temperature decreases (Arsalan, Anwar, & Majid,

2022) . (Zhai & Barreto, 2006) developed a method for human stress measurement in relationship with skin temperature and it was discovered that there exist a negative correlation between skin temperature and human stress (Zhai & Barreto, 2006). The Trier Social Stress Test (TSST) or a less stressful control task were administered to male and female participants by Vinkers et al., (2013) in order to further study the impact of acute psychosocial stress on peripheral body temperature in humans using several readout measurements. Core temperature (intestinal and temporal artery) and peripheral temperature (skin on the face and torso) were assessed throughout the experiment. In comparison to the controlled environment, ot was discovered that exposure to stress decreased intestinal temperature but didn’t affect temporal artery temperature. This further explains that the effect of stress results in steady change in temperature. With this discovery, the study supports the addition of body temperature as a physiological parameter used for further research in stress studies.

3. Methodology

This chapter will document the methodology implemented to answer the research question. The design case study framework created by is the foundation for this study's methodological approach (Wulf, et al., 2011). The idea behind design case studies is to get a clear understanding of already existing practices and then in an iterative and participatory manner design application tailored to meet the need of the users while also evaluating the final outcome of the design with the users. This approach works well for designing IT artifacts since design is an interactive process that is influenced by the social norms of a specific application setting (Castelli, 2020). In the section that follow, the basic idea of the design case study and the empirical methods used in this research work are covered in more detail. See below in Figure 11 the schematic display of Design case study.

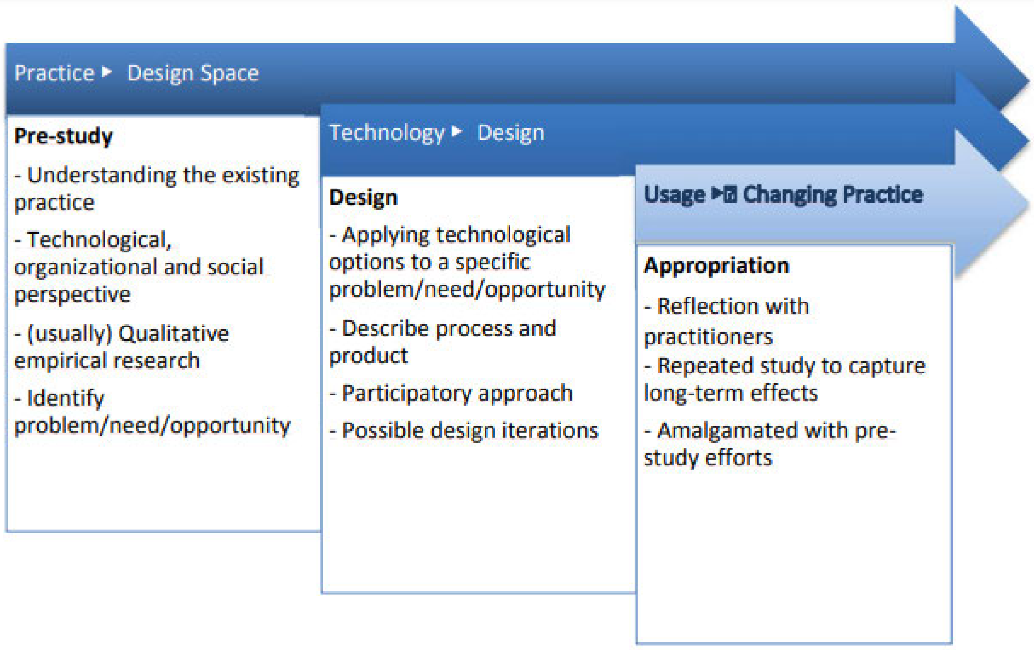


Figure 11. Case Study Design Structure (Wulf, et al., 2015 p.10)

3.1 Design case study

A multi-stage action research model (Wulf, et al., 2011) is used in design case studies, which combines methodologies from design research (Hayes G. R, 2011) (Wulf, et al., 2011) with those associated with the scientific description of the culture of a specific society or social activity. The purpose of Design case study (DCS) is to develop cutting edge ICT artefacts in an iterative way which supports a social practice or proffer solutions in order to solve the problems that exist in a particular organizational or social context. In the course of using DCS, research is also carried out over an extended period of time to understand how practices can adapt to a newly developed ICT artefact and vice versa. According to (Stevens, et al., 2018) DCS represents an on the ground approach to design by addressing users' requirements

Before iteratively and collaboratively building application-oriented solutions and co-evaluating them with the user, the core purpose of design case studies is to properly understand present practices. The social norms of a particular application area provide inspiration for design, which is therefore seen as an inclusive process. (Wulf, et al., 2015) divides design case study into three phases which are:

- The empirical Pre-Study
- The technology design phase
- The evaluation phase

It is best to conduct the empirical pre-study before beginning any interventions establishing the foundation for further action (Wulf, et al., 2015). The researcher gains a complete understanding of social practices throughout this stage of the study, including both formal and informal uses of the current tools and media. The creation of a "particular problem" or "need statement" that should serve as the foundation for a subsequent response is facilitated by this understanding of social behaviors (Wulf, et al., 2015) . As part of the technology design process, the design space is mapped by addressing the observed behaviors and issues, artifacts, or actions. According to (Wulf, et al., 2015), the design case study must outline "the specific design process, the engaged stakeholders, the applicable design approaches, and the emergent design concepts."

To assess the appropriation in its field of application, the intervention should be implemented in real-world settings (Wulf, et al., 2015). New specifications for redesign ought to result from this. Design case studies are therefore especially appropriate for dealing with an intense context to determine the needs for IT artifacts and examine how their use affects societal customs.

3.2. Methodology of Pre-Study

Diverse qualitative empirical techniques were used in each stage of the design for the case studies in this thesis. In research, qualitative methods are employed to have an understanding of the social practice that exists in the target group. Such methods include semi-structured interviews, focus groups, cultural probes, field observation and participatory design etc. The aim is to understand the individual contexts, record current practices, develop requirements and concepts, and test and evaluate prototypes. The section that follows below will give a description of the individual methods and how they were deployed in this work.

3.2.1. Interviews

Interviewing end users is one strategy for reaching the target population and identifying issues and needs. Structured interviews are utilized in the design case study approach to gather information about users and the context. It is also considered an effective method used to comprehend situations, gather wants, and perform evaluations (Castelli, 2020). In research, semi-structured interviews can be used independently or alongside other (qualitative) techniques (Longhurst, 2010). The formulation of the questions used in semi structured interviews is one of the main obstacles. They ought to be open-ended so that the interviewer gets a response other than "yes" or "no," which helps to establish an extensive and information rich conversation. Additionally, the leading questions should be able to be asked in a different order so that the interview's flow is flexible. An interview is not intended to provide results that are widely accepted (it is the same with most other qualitative methods). Understanding how people perceive, experience, or interact with circumstances or technology is a major focus of this study (Longhurst, 2010).

In the first phase of this work, semi structured interviews were conducted with three researchers in the digital health space. Two of them are researchers in the Faculty of life science at the University of Siegen while one is a researcher at the Nano science department at University of Kassel. It was important to interview researchers in these areas of research because they are the target users of the App we intend to design. Valuable and insider information as regard their practice was gathered during the interview which forms the base for the design of the App. Shown in table 1 is a description of the interview participants and some valuable information about them.

Participant	Age	Gender	Profession	Technology Knowledge
P1	35	F	Researcher(Life Science)	Medium
P2	29	F	Researcher(Life Science)	Medium
P3	28	F	Researcher(Nano science)	Medium

Table 1. Pre-study phase Participants

Two of the interviews were conducted online via zoom while the last interview was a face to face interview. The interviews lasted between 30-60 minutes with an average time of 45minutes. To start the interview, introductory questions that borders around their personality and research interests were asked. Thereafter open ended questions about their profession, research activities they conduct and get engaged in were also presented to the interviewees Since one of the goals of this work is to design a technological tool, questions about their knowledge of technology and the kind of technological goals they used for conducting research in a laboratory or controlled setting were also asked.

The interview session ended with questions as regard their expectations of how a Bench Mark App for stress elicitation will look like, the features and functionalities they will expect to have in such software tool and how they think this will help to make their research work easier.

3.2.1.1. Transcription of Interview and Thematic Analysis with MAXQDA

For the transcription of the recorded Interviews the tool Otter.ai was used. It is often used to translate speeches to text and also some selected para-linguistic conversations. After the translation of the speech to text, the six phase approach to thematic analysis developed by (Clarke, Braun, & Hayfield, 2015) was implemented. Familiarization with the data obtained from the interview was the first step taken and this was done by reading through all the interview text with the intention of identifying some common themes from the interview. Thereafter, the data was imported to MAXQDA which is a tool used for the analysis of qualitative data. In order to obtain codes from the qualitative data at hand, an exploration of the transcripts was done using the inbuilt functionalities of the MAXQDA tool. The codes extracted from the transcript were grouped and then further separated into themes. Some information from the interview that didn’t fit into any of the themes were discarded.

3.2.2. Focus Groups

Focus groups are utilized in the design case studies to gather user feedback on the prototypes created to start an internal dialogue and idea exchange.

They are a particular type of semi-structured interview that is conducted in a group setting to elicit as many unique insights as possible and to encourage participant inspiration (Bader & Rossi, 2002) (Breen, 2006). A moderator oversees the entire discussion and is responsible for maintaining an open environment where the participants can freely express themselves. It is important to consider critical similarities in the selection of participants to minimize contradictory viewpoints (Morgan & Krueger, When to use focus groups and why, 1993). A focus group, according to Ritchie (2003), comprises interviewing four to ten people who get together to discuss the research issue in a group setting. This format promotes data collection through the interaction of the participants while they are sharing ideas. Because of the diversity present, participants in a focus group have the potential to influence one another based on what they say (Mack, 2011). The frequent employment of this approach in qualitative social research to supplement other data-gathering techniques to provide researchers with more insight is vital (Morgan, 1996).

Due to distance in location of the researchers, the focus group session was conducted online via zoom. The session commenced with an introduction of the idea behind the App to the researchers. A brief description of the research gap and how this work intends to fill that gap was also presented to the researchers before the commencement of the focus group discussion. After about 20 minutes of preliminary discussion, the moderator of the session began asking introductory and personal questions to the participants so as to lighten up the group and get them mentally prepared for the session. Thereafter, questions that borders around how research is conducted in the life science space is conducted. The questions presented the researchers were centered on the following topics:

- The data collection methods and processes for researchers in a laboratory or controlled setting.
- Criteria used by researchers in recruiting participants for stress elicitation research activities.
- Connection that exist between technological tools, stress elicitation and neurological status assessment research activities.
- The role of wearable devices in Physiological signals capturing and stress studies
- Researchers' perception about the use of software tools for their research activities. While also outlining the pros and cons that exists.
- What the future of research in Life science will look like in connection with ICT tools such as Data Analysis software, Mobile Apps, Virtual reality, augmented reality and machine learning.

During the session we had four researchers in total and the researchers were encouraged to share their ideas without any form of restrain. All ideas where welcomed without any form of objection or criticism from the other researchers present in the group. The interviewer offered the researchers space to hold an open conversation at the end of the session so they could ask questions regarding the App, the features it will have, and the value proposition of the concept. To wrap up the session the researchers gave some expert recommendation, tips and suggestions on what they expect to have on such a stress elicitation App to make their research work easier. Some participants in the previously conducted interview session also participated in the focus group session. Show in table 2 is a description of the focus group participants with further information of their demographics.

Participant	Age	Gender	Profession	Technology Knowledge
P1	35	F	Researcher(Life Science)	Medium
P2	29	F	Researcher(Life Science)	Medium
P3	28	M	Researcher(Nano Science)	Medium
P4	45	M	Researcher (Digital health)	High

Table 2. List of focus group participants.

4. Idea Conceptualization and User workshop deliverables

Before investing time and resources in creating prototypes for the project, it was important to have a project conception and co-design phase which played an important role in ensuring the created design is tailored to specifically meet the users' needs which leads to project success. The conception and co-design phase of this project gave actors from various disciplines the opportunity to meet and share knowledge about not only the design process to be implemented but also the design content for the product so as to have a common or shared understanding on both aspects which then transcends into achieving the common objective of the project to be designed. The co-design process was also instrumental in identifying problems and then finding solutions while also exploring future alternatives through specific approaches (Cottam, Burns, Vanstone, & Winhall, 2006). As a first step, Initial requirements capturing, verification, feasibility analyses, and proof of concept prototypes were some of the activities executed in this phase of the work. Facilitating participation has thus become one of the pillars of co-design, as practitioners from various research and design domains have realized the value of incorporating varied user groups in the development phase of unique artifacts, goods, and services (Brandt, Johansson, & Messeter, 2005). At each stage of the design development process, co-design is possible. However, early (and frequent) involvement of individuals with varied viewpoints might aid in identifying the true problem space thereby producing better results. A guided strategy of discovery, ideation, and development can provide end users and those who serve them a voice, even though integrating several stakeholders at various phases of the process might be challenging. By doing this, informants become contributors, participants, and collaborators, and this can have a significant impact on the results. In the co-design process implemented in this work, a first step was to gather insights from the potential users through interview sessions so as to understand their practice while also identifying problems, needs and opportunities in the research area. Thereafter, we proceeded into exploring concepts with the use of personas, user journey map and user stories. This was a good way to put the users at the center of the design by creating a model of a user while drafting some of the pain points, needs, and motivation we envisage such model user will have. An ideal way to create a convergence for all of the ideas gathered in the previous steps is to do a co-creation with the users. This session entails giving the users the opportunity to design alongside the designer. The users have the chance to pen down some ideas or features they will like to see in such an app. In this endeavor, the co-creation session was essential for understanding the users' mental processes. Creating prototype is the last phase of this flow and it is in this phase the low fidelity and high fidelity prototypes are created. Figure 12 and 13 below gives a pictorial description of the co-design process flow that was implemented in the course of this project.

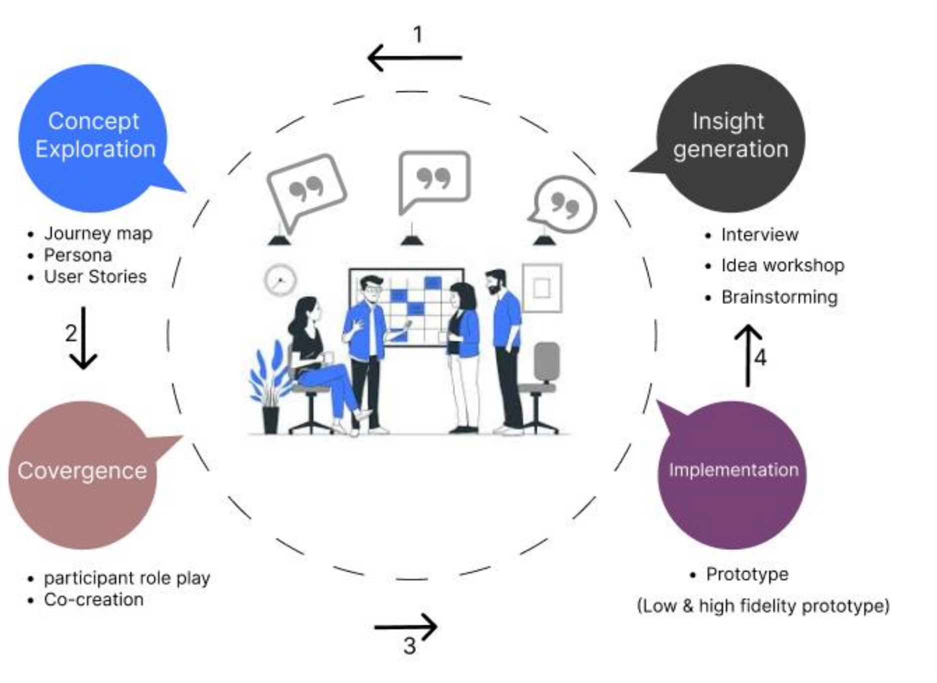


Figure 12. Co-design process flow (Koo & Ahn, 2018 p. 289).

		Co - design process
Co-design process	Insight Generation	Define Relations with stakeholders and design scope while also defining problem and potential users of your product
	Concept Exploration	Exploring concepts and generating ideas through User journey map. User stories, affinity map.
	Convergence	Communicating with stakeholders via co-creation.
	Implementation	Implementing the design via low fidelity and high fidelity prototypes.

Figure 13. The four step co-design process (Koo & Ahn, 2018, p. 234).

4.1 Insight generation by defining the problem and identifying users.

Generation of insights by defining the problem was an important activity to engage in during the participant workshop since it helps to sharpen the remaining part of the design process and give clarity on how success at the end of the project will be measured. To generate these insights and have a better understanding of the target group for which the benchmark App is designed for, a focus group workshop was conducted and the sole purpose of conducting this workshop was to collate information and insights by generating conversation around the topic of designing an app for stress elicitation with focus on researchers in the digital health research sphere. It was an online session which means the exchange of ideas and opinions were completely done online with a moderator observing and moderating the session. During the session, fascinating insights around themes such as motivation of the users, their emotions, opinions, ideas attitude, potential behavior and expectations were generated. Thereafter, a review of the extracted data gotten from the focus group workshop was analyzed, reflected upon and conclusion was made which played a key role in defining the direction for the design so as to meet the users' needs and help them achieve their goals. As can be seen from figure 14 below some of the key points, topics, areas of discussions and questions raised in the workshop are shown on the left side while some of the responses gotten from the participants in the course of the workshop are shown on the right side. An extensive array of ideas and concepts were generated during the insight generation phase but some of them which are not desirable for this project or do not fall within the scope of this work were discarded so as to avoid ambiguity or discontinuity in the process flow of this project.

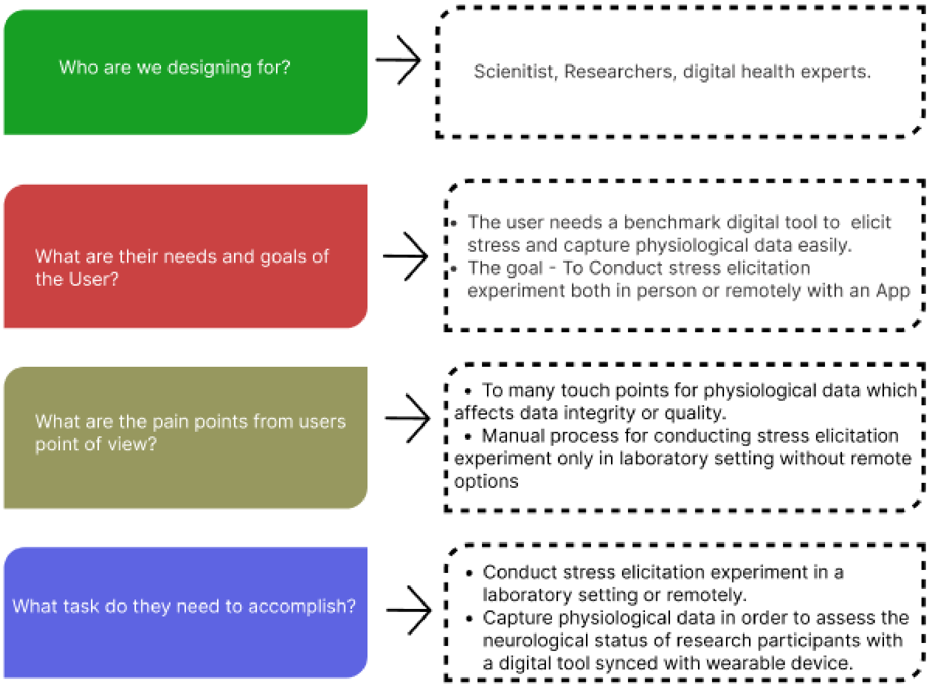


Figure 14. Questions asked to define the Problem and identify the users.

4.2.1. User Stories

While conducting the participant workshop user stories which are created by, with, or for users or clients for the purpose of creating a structure for how the system is functionally designed was implemented. The user story template card seen in figure 15 below was drafted in order to have an in-depth understanding of what the users desire to have in such a benchmark App for stress elicitation. During the workshop, the participants were given user story cards on which they had to write what their needs are and what they intend to achieve in their research activities with such a benchmark App. A straightforward template like "As a (user type), I want to (desired feature user wants), so that (benefit it gives)" was used in the user story card and the users had to follow this format in giving a response to the question.

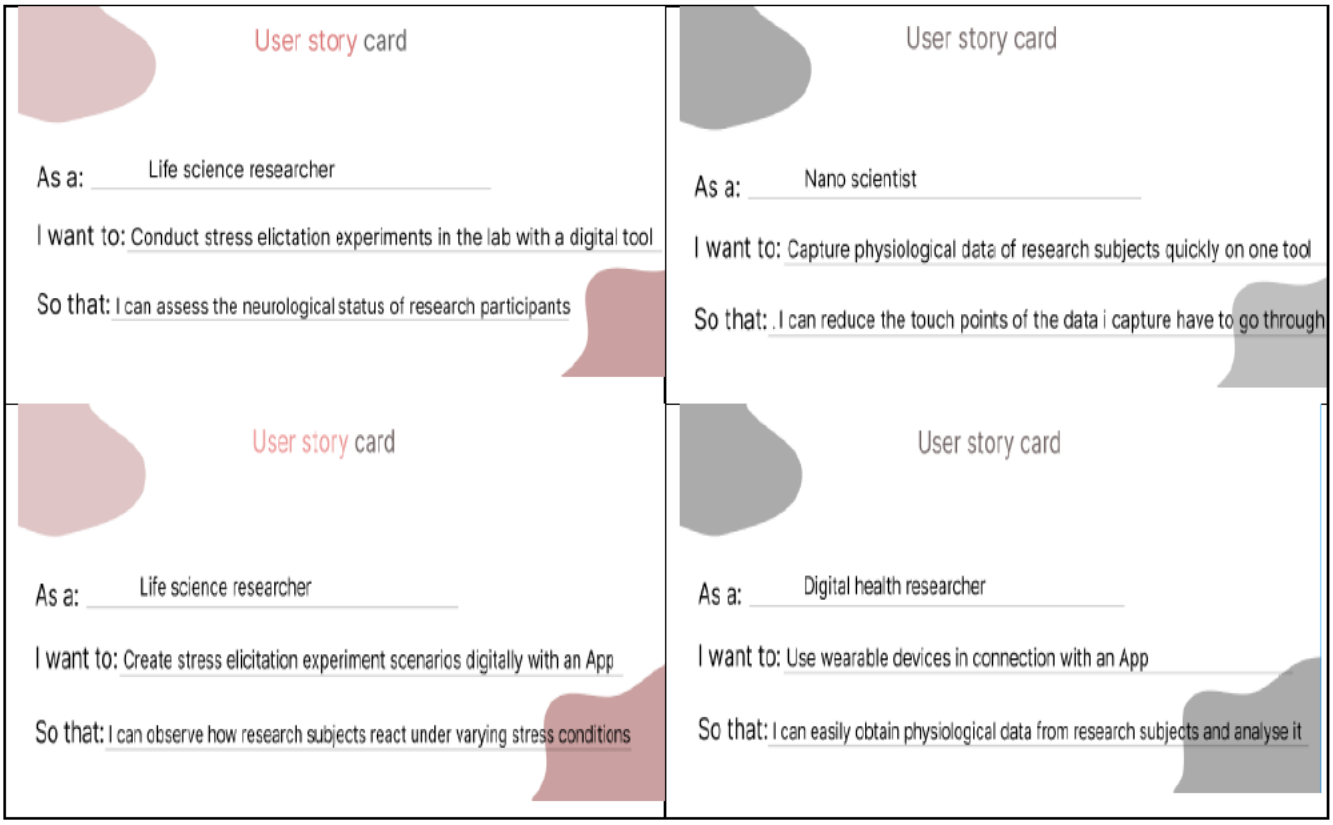


Figure 15. User stories obtained from workshop participants

The potential users of the benchmark App are looking forward to having a tool that will enable them to accomplish the following points below, as also seen from the user story cards in Figure 15 above:

- Obtain physiological data from research subjects with an App synchronized with wearable devices.
- A tool that will reduce the touch points physiological data from participants will go through before analysis.
- Observe the reaction of research subjects under varying stress conditions.
- Create stress elicitation experiment scenarios digitally with an App.

The aforementioned points were instrumental in taking informed decision as regard the design of the App. At every point in the design phase, each of the points were taken into cognizance so as to ensure the final prototype delivers value back to potential users. Outlined in figure 8 is the user story workshop guide which outlines the overview of the workshop, the participants and some of the tools used during the session.

USER STORY WORKSHOP GUIDE

Overview	This is going to be a hands-on workshop where we intend to identify user needs and generate insights from potential users so as to guide the design decision and process for the App
Goals	• Conceptualize the User interface features of the product • Prioritize the right tasks before embarking on the design. • Allow users communicate their needs and goals to the designers • Create a roadmap and structure for the App design
Duration	1 hour
Participants	Life Science researchers, Nano scientist User experience designer Note taker
Tools needed	User story card Sticky notes Pen Miro Software

Figure 16. User story workshop guide

4.2.2. User Persona

For the comprehension of the target audience, the concept of a user persona was implemented for generating ideas. In this case, the qualitative research earlier conducted in the early phase of this work was useful. The psychological advantage of the user persona in this context stems from the designer being able to identify with the personas, which allows them to develop an empathic knowledge of users by immersing themselves in the genuine experiences of others (Miaskiewicz, Sumner, & Kozar, 2008). This capability was important to be implemented in this work because it provides the needed knowledge to envisage how the users will react in different situations. Fig 17 below shows the User persona used in this project. It serves as a fictional character of our users based on the information gathered while interviewing the users. It outlines some important information about the typical user, the goals and pain points of the user and some of the goals they intend to achieve when making use of an App to conduct stress elicitation experiments.

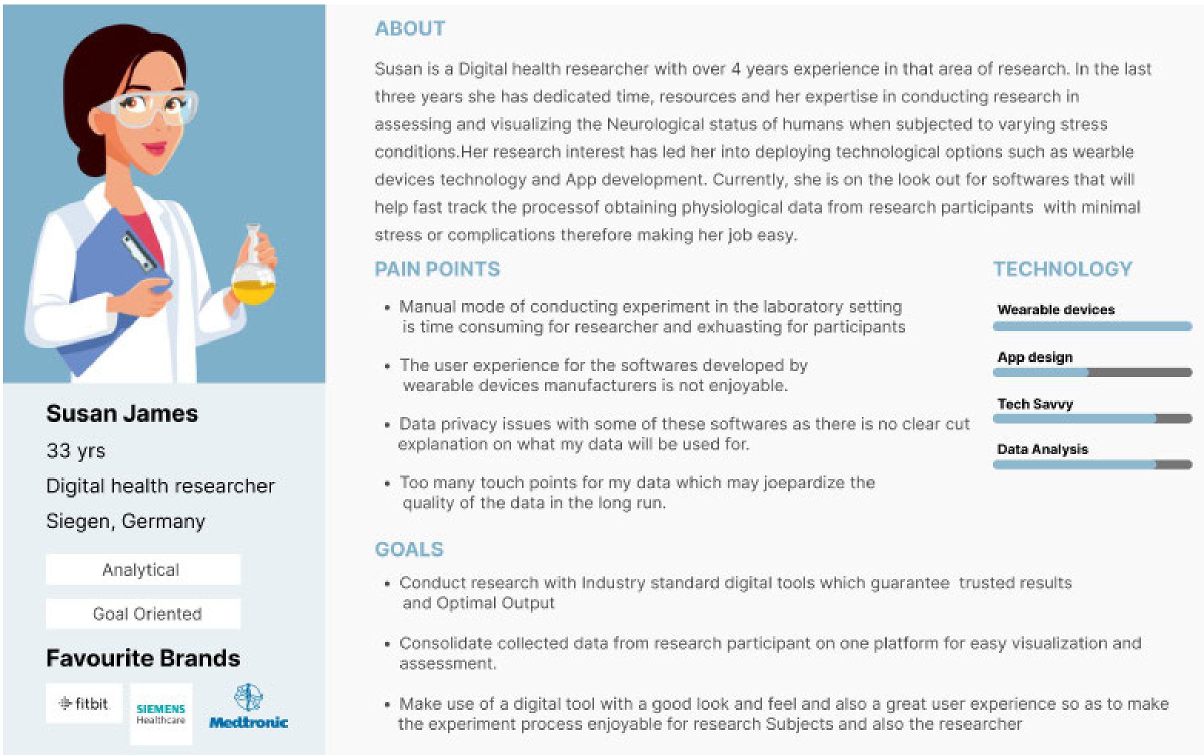


Figure 17. The User Persona

4.3. Requirement Gathering for the App Design

Before embarking on sketching design ideas for the user interface (UI) of the app, it was pertinent to gather requirements from users. Research has shown that effectively obtained requirements play a crucial role in ensuring that software design/development proceeds without hiccups (Malik, Chaudhry, & Malik, 2013). Typically, requirement collecting is done to acquire data from a variety of diverse perspectives or points of view. These views play a huge role in determining what will be designed or built. According to (Hickey & Davis, 2003) the success or failure of a system design is greatly influenced by the quality of the gathered requirement. It can be seen from the user stories in figure 7 above that our users in this context have varying needs. The requirement gathering session conducted was helpful to ensure that the product is designed with a high possibility of satisfying those needs. For this work, the method used to generate ideas for the benchmark App and gather requirement from the users is the Affinity mapping method. Section 4.3.1 below gives an illustration of the Affinity map method and how it was implemented in the project.

4.3.1. Affinity Map

Affinity Map method was used in the workshop session to brainstorm, gather ideas, requirements or subjective data about the research problem and then further categorized. During the mapping session, Ideas or elements which share characteristics or coexist in a group were categorized as belonging to that group. Thereafter defining the problem, generating ideas, grouping them into logical categories, creating clear headings for each category, and ranking the ideas inside each group were then implemented. The workshop participants were required to make use of sticky notes to write down their observation or ideas in few words. The ideas written on the notes were then clustered with regard to their affinity i.e. their similarity or importance to a shared theme. Thereafter groups were created and labelled from these themes and then clustered again. This process was repeated until the highest level had only few groups and the previously unmatched items were then organized from the bottom to the top. The ideas, themes and groups from the described affinity mapping process is shown in figure 18. With the points gathered from the process, it became clear which ideas is of top most priority for the users. This also helped in determining which ideas should be pursued for the design of the interface.

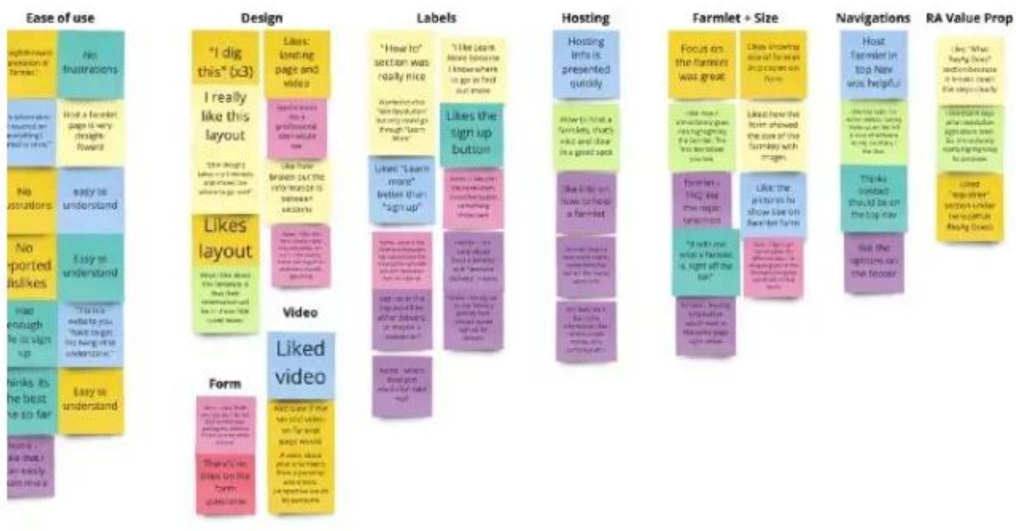


Figure 18. Affinity map from User workshop

Figure 19. Design features required by users.

4.3.2. User Journey Map

The use of user journey map during the workshop served the purpose of identifying both positive and negative aspects in the design and then helped to identify critical points in which improvements or redesign will be required (Javier, 2013) . The different points or stages in the map represents the user's perception at that point. While the negative features represent dissatisfactions that the design team must address in order to enhance the overall user experience of the product, the positive points or expressions made by the user in the map show points that cause some satisfaction in the user. In this research work, the map was used to ensure the priority was put on the user during the design process. The stages which the user of the benchmark app will go through at every point of interacting with the design was captured in the map. Figure 20 below shows the user journey map for this project which outlines the touch points of the user while interacting with the designed benchmark app.

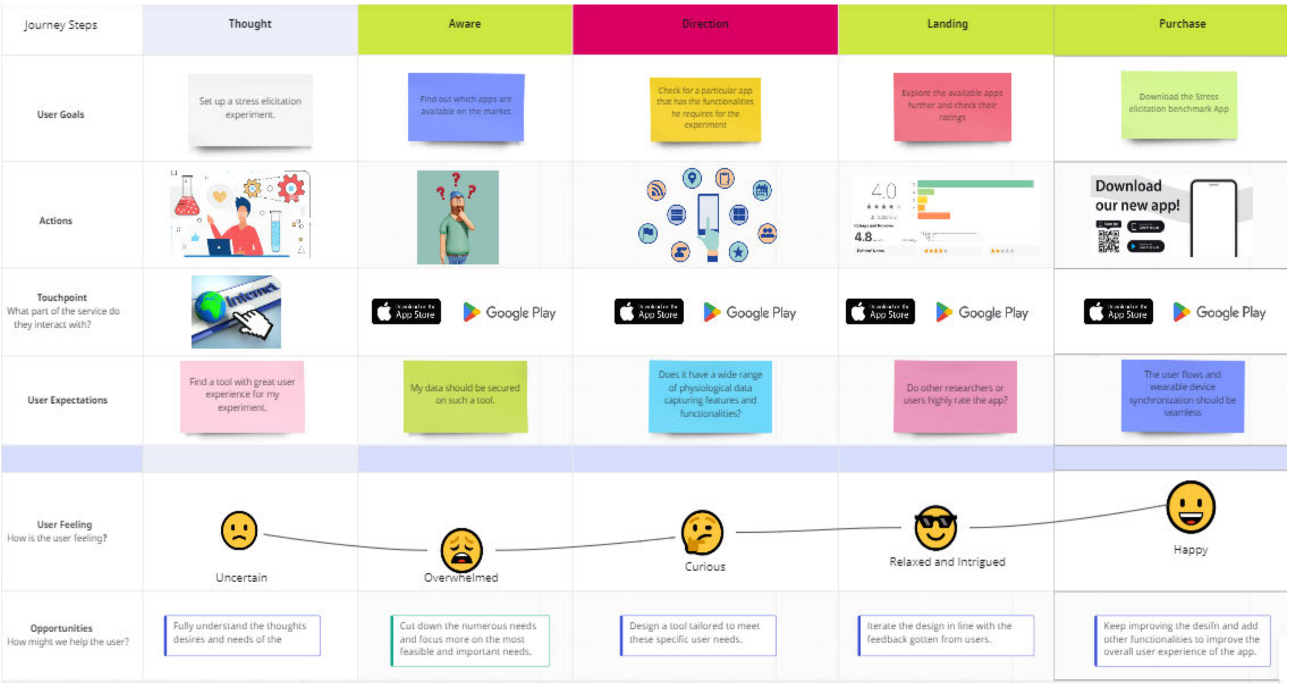


Figure 20. User Journey map

5. Implementation

In this section of the thesis, the steps taken for the implementation of the design will be outlined in detail. Starting with the early stages of the process which is the generations of design ideas and the sketching of low fidelity prototype and then moving to the high fidelity prototype. Insights gotten from the usability testing workshop and appropriation study conducted with users which was instrumental in defining the look and feel of the final design will also be discussed in this section. Below in figure 21 is an over of the entire project timeline with some of the milestones achieved at every point.

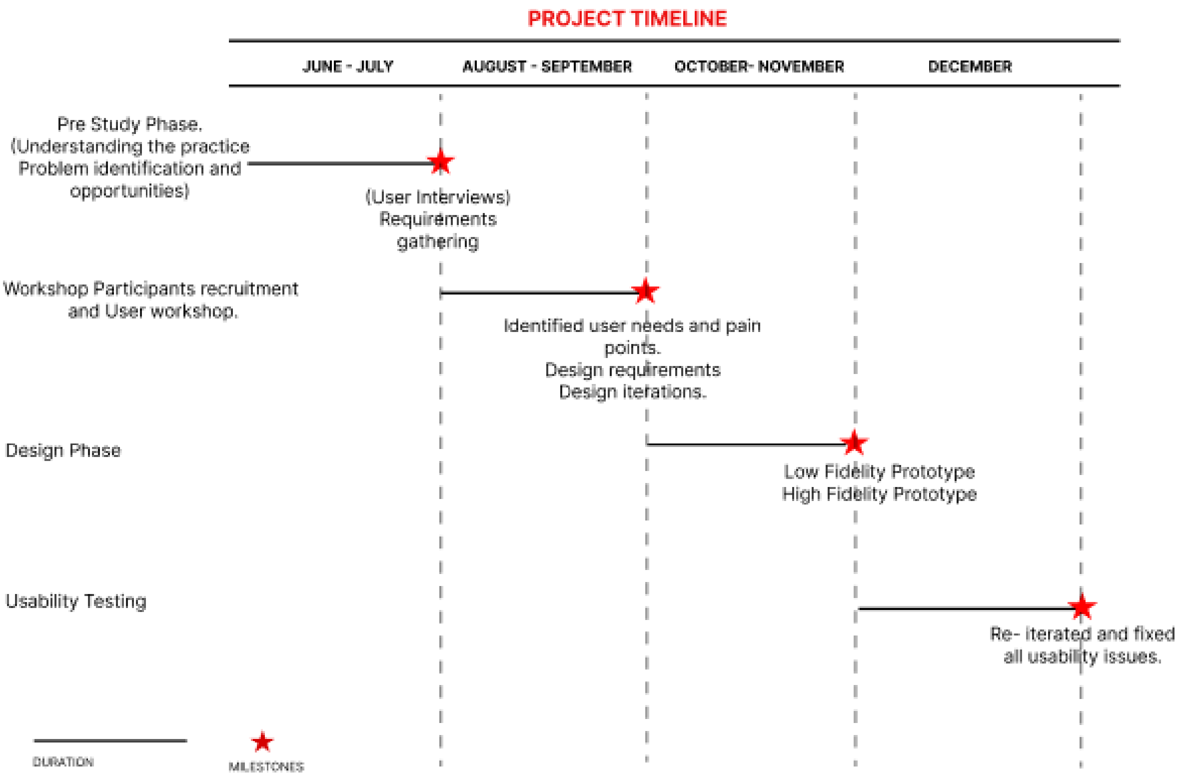


Figure 21. Project Timeline

5.1 Design and Prototyping

At this stage of the work, design and prototyping method offered the users of the App and the designer the opportunity to examine or discover some design problems and also evaluate some possible solutions to these problems. Firstly, the design phase of this project was majorly based on the empirical analysis of the previously conducted pre-study. For innovation sake early prototyping is encouraged so as to enable the users have a feel of the actual design before final deployment. After the generation of design ideas from the workshop conducted with researchers, sketches in the form of low fidelity prototypes designed on paper was created so as to give the design a structure. While creating these low fidelity prototypes, the user flows which defines how users will navigate through the interface of the app was also drafted. The following sub sections in this chapter will give a description of the various prototyping phases that the design was made to go through up until the final design.

5.1.1 Low fidelity prototype (Lo-fi).

Implementation of Low fidelity prototypes at this stage of the thesis was a rapid and effective way to explore initial ideas and revisions. With the use of a pen and paper, initial ideas for the interface of the App were first drawn in order to quickly have a baseline structure for the App and work though some aspects of the user experience. This method involved creating a set of mock-ups of the target interface that imitate specific elements of a system, generally of its user interfaces and interaction solutions, in a more affordable and adaptable manner. After the creation of the Lo-fi on paper, a tool called Balsamiq was then used to sketch out the Interface of the App with a mobile phone mock. See below in Figure 21 the created lo-fi prototype.

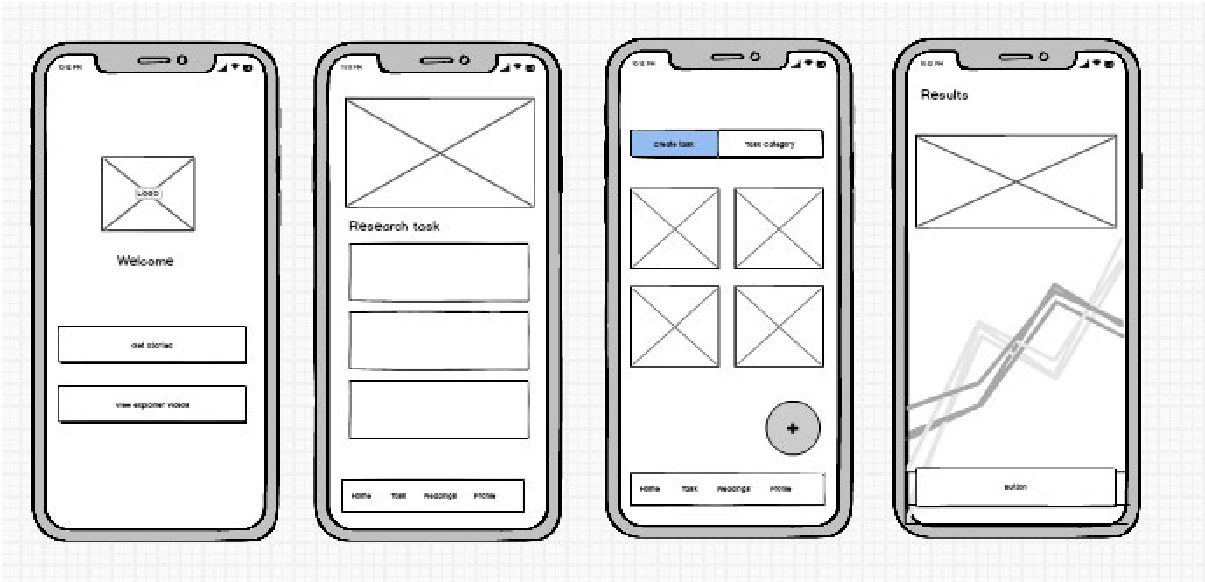


Figure 22. Low fidelity Prototypes

5.1.2 High Fidelity Prototype (Hi-fi)

High fidelity Prototype (Hi-fi) is more functional than a low fidelity prototype, and it provides a basis for the users to do a thorough evaluation of the design (Rudd, Stern, & Isensee, 1996). With the use of the software tool Figma, high fidelity prototypes were created while taking into consideration all design rules. Some of the rules adhered to during the design are shown below:

- Rules for Minimalist design with less aesthetics
- Typography, typeface and fonts.
- Recognition instead of recall
- Color scheme rules with appropriate contrast.
- Visual Hierarchy and alignment.
- System status visibility.
- Consistency and standard design rules.

The Hi-fi prototype created was then presented to the users by using the Usability testing workshop method to gather feedback which will help to structure and fine tune the final design. With the Hi-fi prototype it was possible to observe the users as they carried out tasks that are characteristic of the Bench mark app intended use. In this way, it becomes it was faster to get end user feedback and incorporate it quickly into the design. It also promoted user acceptance since the users can immediately see the recommendations they gave implemented in the design. Below in figure 22 are some selected interfaces of the High fidelity prototype.

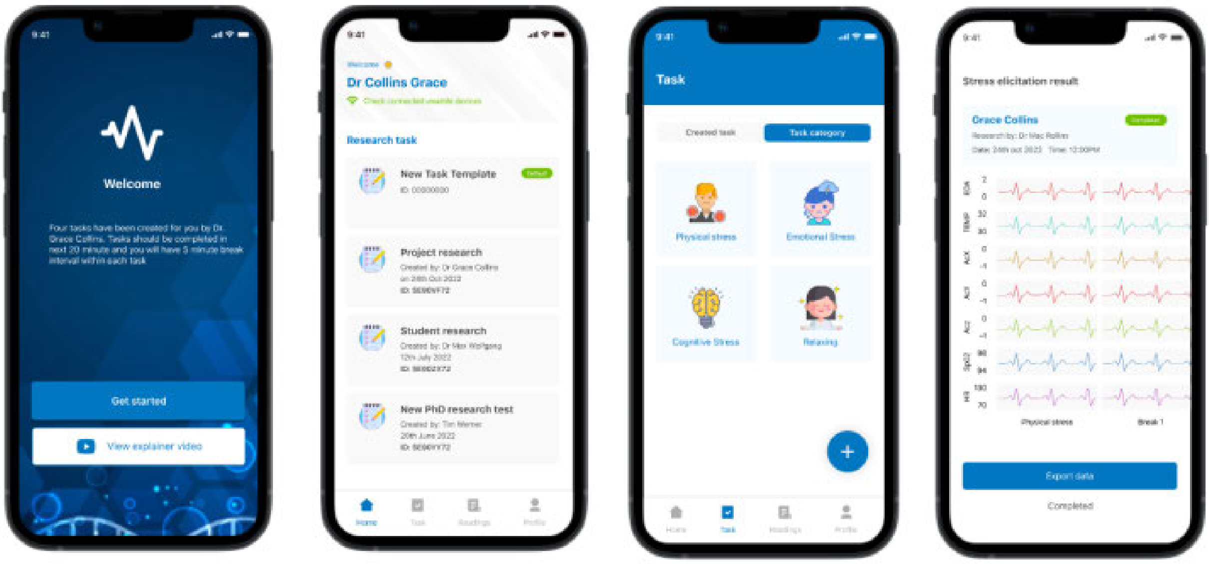


Figure 23. High Fidelity Prototype

6. Evaluation and Results

For the evaluation of the designed prototype in this project, the thinking aloud method was used. While conducting the thinking aloud method, it was required for a representative user to perform one or more task while verbalizing their intentions under observation by the moderator. This method provided the opportunity to visualize how the actual users of the prototype will interact with the prototype. Another use case for the method in this work was to evaluate the prototype's value and efficiency. By engaging the target users of the proposed system with tasks to accomplish with the design, the designer gets to obtain information about the users' cognitive process and obtain valuable feedback which will be useful for the iteration of the design. The thinking aloud method was also useful for the identification of usability issues in the design. It is considered one of the most effective ways in design to study different ways individuals perform similar tasks. It assesses higher level of thinking by obtaining verbal information from working memory (Olson, Duffy, & Mack, 2018). The motivation for conducting a thinking aloud method in this thesis is to ensure the design gets to reach its minimum level of usability, to identify potential usability issues in the design and lastly get feedback from users on how well the objective for the design was met. Going by Nielsen's insights on the conduct of usability testing, which states that 80% of the usability problems can be discovered with five users (Nielsen J. , 1994).

In order to select participants for the thinking aloud workshop, criteria such as level of education, proficiency with computers and technology, experience with the task domain and zero familiarity with the design were taken into consideration. Table 3 below shows the demographics of the thinking aloud workshop participants for this thesis.

Participants	Occupation	Gender	Age Range	Tech experience
P5	Student	M	20-25	High
P6	UX designer	M	25-30	High
P7	Life science researcher	F	30-35	Medium
P8	Fronted developer	M	25-30	High
P9	Nano Scientist	F	25-30	Medium
P10	Student	M	30-35	High

Table 3. Thinking Aloud Workshop Participants.

6.1 Usability Testing and Interface evaluation.

First and foremost, an introduction to the purpose of the workshop in general terms, the usability testing workshop guidelines and rules were explained to the participants. Listed below is the order of activities during the workshop session.

1. Concise statement and an outline of the project's goals and objectives
2. Explanation of the thinking aloud method with a short practical simulation of how to think aloud.
3. Issuance of statement to the effect that attendees are free to leave the meeting at any moment. Signing of consent form before the commencement of the session.
4. Description of the tasks and explaining to the users that I might not be able to provide help during the test.
5. Unveiling of the App interface to the users and commence the session while observing the users and taking down notes of usability issues the users encounter in the course of the session.

6. Conclude the session and ask for users' feedback as regard their experience while interacting with the product.
7. Ask the users if they have further questions
8. Conclude the session and implement some of the required changes in order to improve the overall user experience of the design.

Participants of the thinking aloud workshop also participated in the Usability workshop. In this manner, the participants had the opportunity to explain in detail some of the usability issues they encountered while making use of the App to execute task. The next section documents in details some of the usability issues gathered from users.

6.1.1 Usability issues and their categorization.

Prior to the commencement of the usability workshop, it was important to review and define some usability concepts. The definition of these concepts will aid in categorizing the usability issues from user errors which are often mistaken for each other. Usability according to ISO 9241-11 is described as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use (Alonso-Ríos, Vázquez-García, Mosqueira-Rey, & Moret-Bonillo, 2009). Usability Issues, on the other hand, are a collection of undesirable events, such as a user's incapacity to accomplish a task, an ineffective interaction, and/or a user's unhappiness, that are brought on by a confluence of usage context and user interface design variables. (Manakhov & Ivanov). While conducting the usability workshop with the participants and taking note of the usability issues encountered by users, the following key points were taken into consideration

- Effectiveness: An interaction is considered to be under this category if it is terminated by the user due to lack of clarity or terminated by the system or design in this case without the user accomplishing the task or reaching his/her goal. If the goal is a complex one which may be due to the existence of sub goals, we then assess the user's ability to reach the goals separately. On the other hand, a user's inability to perform a task will be categorized under the class of negative phenomena in the efficiency group.
- Efficiency: An interruption to the interaction or flow of the user while using the interface in a sequence of action is categorized under the efficiency group. Issues which give rise to hesitation before users' take decisions or interactions that lead to the users getting unexpected or undesirable results are all issues which questions the efficiency of the design.
- Satisfaction: Feeling of pain, irritation, discomfort, negative comments or sensations, attitude with negative interpretation toward the design are categorized under the

satisfaction category. While the users were making use of the product, it was important to observe their body language, and take note of some of the comments they made consciously or unconsciously.

6.1.2 Users Feedback and Proposed Improvement.

During and after the usability testing session useful feedback as regard the users' experience while interacting with the design interface was obtained. The feedbacks were analyzed, reviewed and streamlined in order to ensure the changes that will eventually be implemented or made to the design do not disrupt or change the initial objective or purpose of the benchmark App. The experience of the workshop participants while testing the App varied and was fascinating to see the different thought process of each user while navigating around the App. The insights obtained from each of the session was instrumental in ensuring that the final design of the App meets the expectations of the users and helps them achieve their goals with minimal efforts.

Participant 5(P5):

P5 is a student of HCI and has a strong technical background. He has an understanding design principles and usability which made it possible for him to critic and give some valuable feedback as regard the design and overall user experience of the prototype. He was able to follow through with the tasks without much difficulty and his thoughts and anticipations while navigating through the interface were also clearly communicated during the session.

> *P5. Generally, I find the interaction and user flows of the App very intuitive but I think in order to reduce the number of clicks a user makes when creating a tasks, it will be nice to make a default task which a user can easily edit and use for an experiment. The created task and categories should be on the same page for easy navigation.*

Participant 6 (P6):

P6 is a UX designer with over 3 years' experience of designing digital product. The usability testing session with P2 was very insightful. The feedback obtained during this session exposed some usability issues that were previously ignored while designing the interface. During this session, feedback was obtained from both the technical and user perspectives. The incorporation of this feedback will greatly enhance how the App is used by its users.

> *P6. I like the idea and overall concept of this app. From a technical perspective, I think some improvement can be made to the look and feel of the interface. The visual hierarchy and spacing of the text should be adjusted in order to improve readability*

and interaction on the App. The color scheme should be well balanced following the WCAG contrast and color requirements/ principles in order to allow for inclusion and accessibility in cases where you might have users with color blindness. The size of the buttons should also be increased a little bit since users have varying sizes of thumbs.

Participant 7(P7):

P7 has a life science background and it was important to have a participant with such background so as to get feedback from the perspective of a user in the field for which the design is being curated for. During the session, feedback in line with how a typical experiment session was gotten and this was very helpful in ensuring the design was created to suit a typical laboratory or medical experiment setting

P7. The entire idea behind this concept fascinates me and I think other researchers in my field will also be interested in making use of such App. The design also looks engaging and easy to interact with. I suggest you include the possibility to add multiple wearable devices since we make use of quiet a lot of devices for our experiments in capturing data from participants. There should also be a distinction between the researcher mode and user mode in the App for better usability.

Participant 8(P8):

P8 is in the software development space and has valuable experience as a front end developer. From a technical point of view, he gave some insights which were important to help users avoid cognitive overload when making use of the App

P8. Icons used in the App should be frequently used icons so it is easy for users to recognize them. Items on a list in the Menu tab should be reduced to (7±2) since the human brain can only process information up to (7±2) chunks in their short memory according to Miller's rule (Lu, 2011)*. White space in the App and the typography used in the design should be well balanced for better readability*

Participant 9 (P9):

P9 has a degree in Nanoscience and some wealth of experience in carrying out experiment in the Lab with research participants but medium tech knowledge.

P9. The interface looks good and the structure and layout of the tasks is well positioned. It will be nice if each created task has a task ID for easy identification. All typographical errors in the design should be corrected

Participant 10 (P10):

P10 is a student with a background in engineering. He has taken part in a usability testing workshop and has some experience as regard how it is conducted. He was able to navigate through the interface with less difficulty and also gave some valuable feedback.

P10. The aesthetics of the design is ok and I like the look and feel. It will be nice if you can include some short explainer videos for new users so they know how to navigate around the App while executing the tasks. Users should also get some feedback from the App informing them on where they are at the moment and the next steps to take. You can also explore the possibility of having the users execute the tasks from the comfort of their homes without needing to visit the laboratory.

6.2 Re-iteration and Final Appropriation

After the completion of the usability testing workshop, an important step to take in order to increase user acceptance of the design was to do an analysis of the results gotten from the workshop and then re-iterate the design to ensure it is "appropriable". If a user can quickly understand how a system's features can be used as resources for action, then the system is appropriate (Salovaara & Tamminen, 2009). Designing with an understanding that the user is an active agent who has the ability to adapt the product to serve their personal goal, it was important to take the topic of appropriation into consideration so as to ensure the design meets the needs of the user. Appropriation is the way in which technologies are adopted, adapted, and incorporated into working practice. This could entail customisation in the conventional sense (i.e., the intentional reconfiguration of the technology to meet local needs), but it could also just mean using the technology for purposes other than those for which it was intended or to further new objectives (Dourish, 2003). (Salovaara & Tamminen, 2009). In designing for appropriation, it was important to develop the designs in a way that empowers the user with the required functionalities in the App so as to enable them accomplish their task. Additionally, it was crucial to add interactions and functionalities when creating the app so that users may adapt it to their own work environment and encourage appropriation. The following section below which has figure 23- 26 will show some of the notable function included in the design and the redesign done after the usability testing workshop in response to the usability issues we noticed the users had while making us of the App. The re-iterations were done so as to ensure the final prototype is a user-centred design (UCD) with features that makes it possible for the users to easily achieve their goals.

- ## Usability Issue 1:

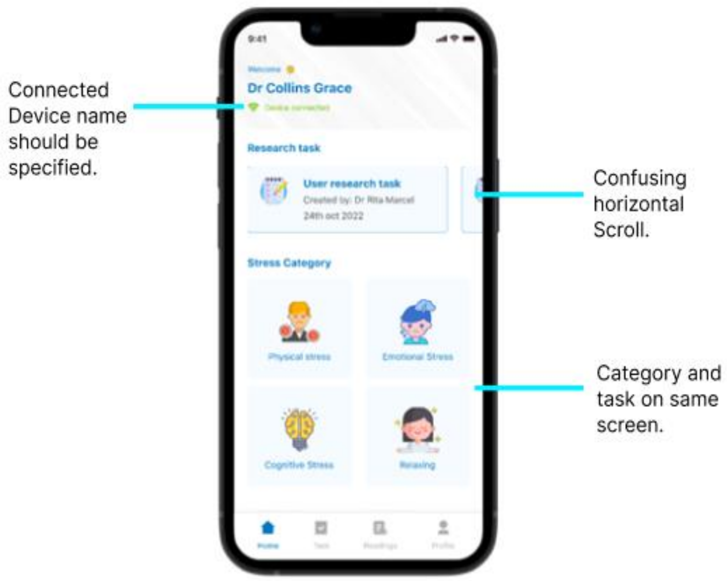


Figure 24. Stress Category and connected device indication

- ## Re-iteration 1:

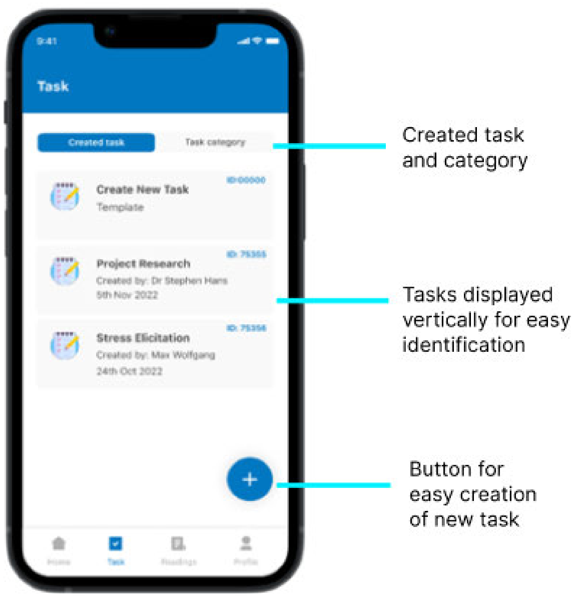


Figure 25. Re-iterated features to solve usability Issue 1.

During the usability testing workshop, users had a tough time making use of the horizontal scrolling feature seen in figure 23. This made it a bit difficult for them to see the previously conducted task and select the one they needed. Another discovered issue during the testing workshop was the layout of the task category and created task. Users needed a default task template which they could make simply add some few information to in order to create their own task. The issue with the connected device notification on the interface in figure 23 was also discussed. It was recommended to have the specific name of the wearable device written on the screen. As seen in Figure 24, care was taken to restructure the interface and take into consideration all the usability issues identified by the users. The horizontal scroll functionality was taken out and replaced by a vertical display of the task with a default task template at the top. The created task and all stress categories were place side by side for easy navigation and reach. It was also taken into consideration that since users will be creating multiple new task, it will be a good practice to have the create new task plus (+) button at the bottom of the interface for ease of clicking and reach since most users agreed to using their thumb finger for clicking the app features.

- <u>Usability Issue 2.</u>

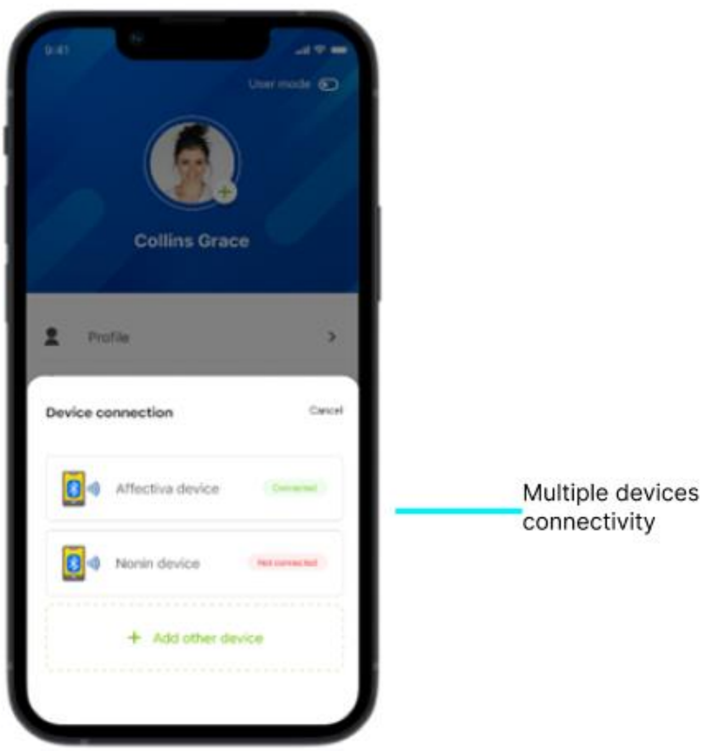


Figure 26. Multiple wearable devices functionality.

From figure 24 above, it can be seen that the functionality to have multiple wearable devices connected at the same time has been included which was previously not on the interface presented to the workshop participants. The connection status of the device is also shown at the right side. This makes it visible clear for the user to know the device that are connected and can be used for stress elicitation experiments.

- <u>Usability Issue 3.</u>

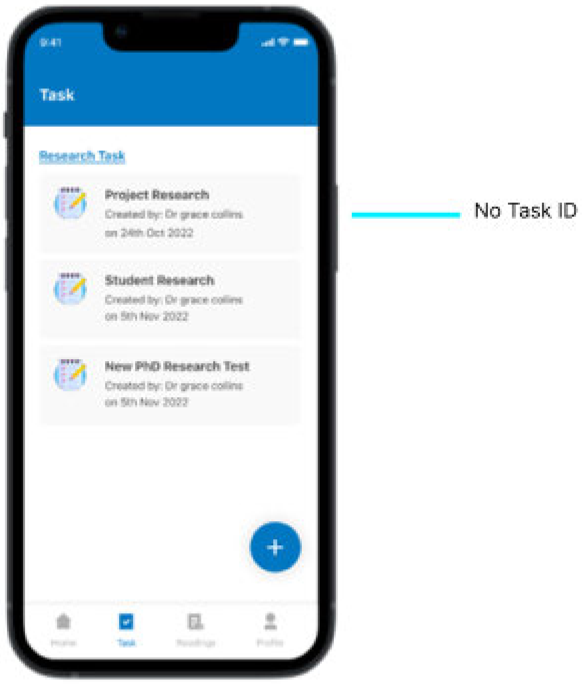


Figure 27. Task ID usability issue

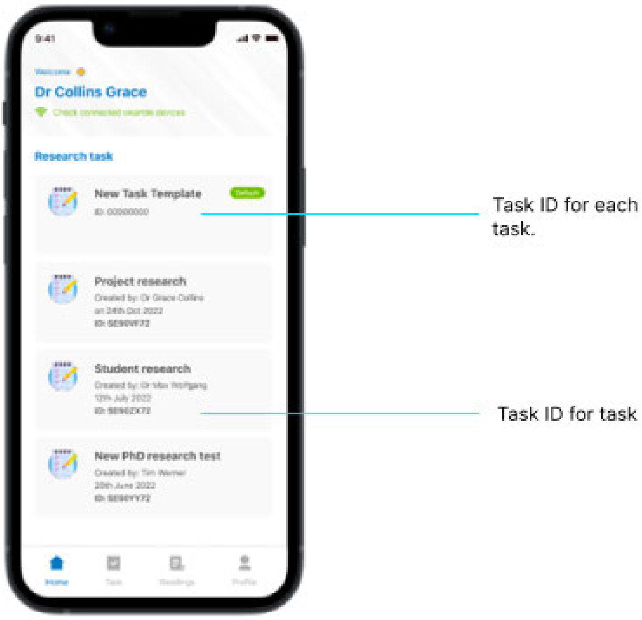


Figure 28. Task ID inclusion

In order to easily identify a task and distinguish it from others, task ID for each task was included in the re-iterated Interface as shown in figure 26. The previous interface had no task ID which made it difficult for users to distinguish between tasks.

- Usability Issue 4:

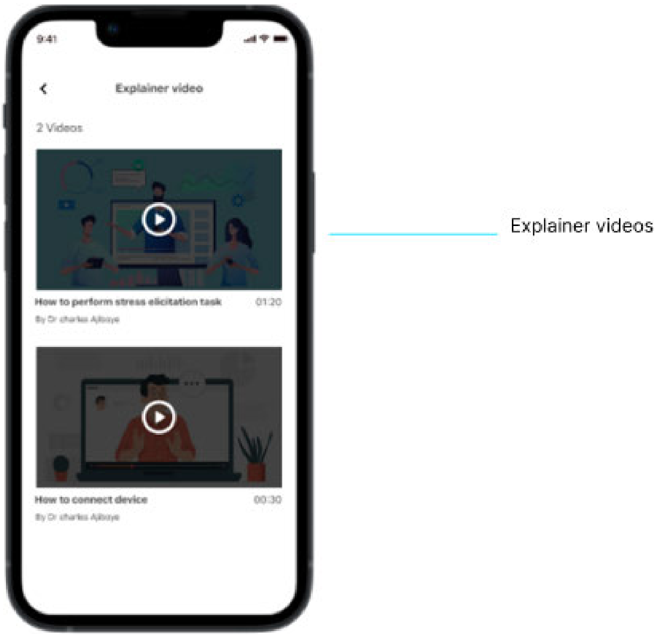


Figure 29. Explainer video features.

During the testing workshop, some users who aren't tech savvy had a hard time understanding some of the features and functionalities in the App. For ease of onboarding, it was decided that users needed to get a short glimpse of how to navigate around the app in order to execute the task handed over to them by the researcher. The inclusion of an explainer video helps in this case to highlight some important aspects of the App to the users and make them have a first impression of what to do at every point in time on the App.

- <u>Usability Issue 5:</u>

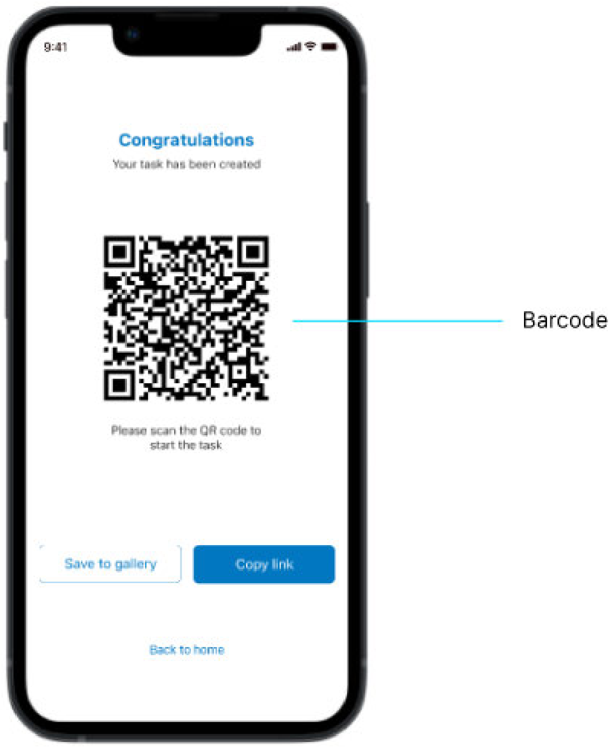


Figure 30. Barcode functionality

Some users raised up some questions as regard performing the experiments remotely without having to visit the laboratory. With the addition of the Barcode feature, the researcher has the opportunity to create task for users and simply have the users scan the barcode and then execute the task. This also creates the possibility for the task to be completed remotely by the users. With the barcode, the can complete the task from the comfort of their homes without having to be in a laboratory setting. The researcher also automatically gets the results of the experiment on his device as soon as the users have completed all of the tasks assigned to them. This helps in fast tracking the experiment process and makes it a comfortable one for both parties.

After the entire evaluation of the interactive design, some other minor suggestions received from the participants which aren't shown here were incorporated into the design to ensure all usability issues discovered in the workshop were eliminated.

7. Conclusion and Outlook

Through the application of Design Case study framework, this thesis attempted to answer the research question

How can a benchmark app be used for the elicitation of stress and capture of physiological data in the field of digital health and life sciences?

In order to understand the work context and practices, the first step taken to address the aforementioned question was to involve users throughout the design and iteration processes. This was accomplished using qualitative research techniques such as in-depth interviews, focus groups, and co-design workshops. The insights gathered at the early stages of the qualitative research process were instrumental in defining the direction in which the project will proceed and also ensure that the final design meets users' needs and help them achieve their goals. A study gap was found in the methods used to conduct experiments focused on stress elicitations, assessment, and visualization of the neurological status. It was discovered that when such experiments are conducted, a large repository of multimodal dataset of physiological signal data are made available in the public domain however there was no benchmark platform where these huge data can be consolidated. This context inspired the idea to create a benchmark app that can be used to set up and carry out stress elicitation studies in synchronization with wearable devices to collect physiological data from research subjects both at a distance and in a lab setting. In this manner, the huge amount of data generated in the course of conducted such experiments will be consolidated on one platform from which researchers can then draw from for further analysis on the software of their choice. This offers the advantage of reducing the collected data footprint and also fast tracks the entire stress elicitation experiment process for researchers.

In this thesis, a design which is expected to serve as a framework for the benchmark App has been created. With the observation and application of design principles and guidelines in the field of Human Computer Interaction, the deliverable from this work is a foundation for the further development of the Benchmark App for stress elicitation. It is important to note that the goal of this thesis was not to offer a design solution that would work for all situations, but rather to identify the potentials and opportunities that exist in this area of digital health research in order to conceptualize and implement design concepts that could then be further developed into fully functional software applications.

7.1 Limitation and future work

While executing this project, some limitations and hurdles were encountered which slowed down the process of the work and also had some impact on the final outcome and results of the thesis. At the research participant recruitment phase of the thesis, it was a bit challenging to get participants that meet our requirements since the project was focused on designing a digital tool for use specifically in the field of digital health and life science. We required participants with extensive knowledge and experience in conducting research in the laboratory with the use of wearable devices and various software tools for the capture of physiological data from research subjects.

Secondly, at the time of writing this thesis, the idea of designing a bench mark App for stress elicitation was the first of its kind and this meant it was a bit challenging to find sufficient research work or resources that could serve as a reference or guide for this work.

The future of developing Apps for use in experimental and laboratory settings looks promising and has so many possibilities and potentials. For the further development of this work, there exist the possibility of building upon the deliverables of this thesis and develop the already designed app to have the functionality of capturing physiological data without the need to be synchronized with a wearable device. With the use of algorithms and application programming interface (API) developers can explore the possibility of developing the app and incorporating physiological data capturing features or functionalities so as to eliminate the obtrusive discomfort wearable devices create for research participants.

References

Adam, R., & Cogan, N. (2019). he personal meanings and experiences of resilience amongst elite badminton athletes in the build up to competition. *Journal of Qualitative Research in Sports Studies*, 61-84.

Ahmed, S., Noushad, S., & Azher, S. Z. (2014). Assessment of major physical stressors and its psychophysiology; a comprehensive review. *Annals of Psychophysiology*, 3-8.

Alonso-Ríos, D., Vázquez-García, A., Mosqueira-Rey, E., & Moret-Bonillo, V. (2009). Usability: a critical analysis and a taxonomy. *International journal of human-computer interaction, 26(1),*, 53-74.

Anusha, A. S., Joy, J., Preejith, S. P., Joseph, J., & Sivaprakasam, M. (2017). Differential effects of physical and psychological stressors on electrodermal activity. *In 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 4549-4552.

Arsalan, A., Anwar, S. M., & Majid, M. (2022). Mental stress detection using data from wearable and non-wearable sensors: A review.

Asai, K. (2008). *The Role of Head-Up Display in Computer-Assisted Instruction, Human Computer Interaction.* doi:10.5772/85

Ashayeri, H., Shahsavarani, A. M., Lotfian, M., & Sattari, K. (2013). The effects of stress on visual selective attention: The moderating role of personality factors. *Journal of American Science*, 9(6s), 1-16.

Azher, S. Z., Noushad, S., & Ahmed, S. (2014). Assessment of major physical stressors and its psychophysiology; A comprehensive review. *Annals of Psychophysiology*, 3-8.

Bader, G. E., & Rossi, C. A. (2002). Focus Groups: A Step-By-Step Guide. *The Bader group*.

Bakker, J., Pechenizkiy, M., & Sidorova, N. (2011). What's your current stress level? Detection of stress patterns from GSR sensor data. *Proceedings – IEEE International Conference on Data Mining, ICDM, IEEE*, 573-580.

Barreto, A., Zhai, J., & Adjouadi, M. (2007). Non-intrusive physiological monitoring for automated stress detection in human-computer interaction. *In International Workshop on Human-Computer Interaction* (pp. 29-38). Berlin, Heidelberg: Springer.

Bascur, C., Rusu, C., & Quiñones, D. (2018). User as customer: touchpoints and journey map. *In International Conference on Human Systems Engineering and Design: Future Trends and Applications* (pp. 117-122). Springer Cham.

Beale, R., & Peter, C. (2008). The role of affect and emotion in HCI. *Affect and Emotion in Human-Computer Interaction Springer, Heidelberg*, 1-11.

(2015). *Behavior.* Retrieved from http://www.merriam-webster.com/dictionary/behavior

Behnoudi, Z. (2005). Health and occupational stress. Tehran, Iran: Boshra-Tohfeh Publications.

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of neuroscience methods,*, 80-91.

Benyon, D. (2013). Designing interactive systems: A comprehensive guide to HCI, UX and interaction design.

Berbari, E. J. (2000). Principles of electrocardiography. *J.D. Bronzino (Ed.), The Biomedical Engineering Handbook*.

Bernardi, L., Valle, F., Coco, M., Calciati, A., Sleight, P., & Coco, M. (1996). Physical activity influences heart rate variability and very-low-frequency components in Holter electrocardiograms. 234-237.

Bickford, M. (2005). *Stress in the workplace: a general overview of the cases, the effects and the solutions.* Canadian Mental Health Association.

Birjandtalab et al. (2016). A Non-EEG Biosignals Dataset for Assessment and Visualization of Neurological Status. *IEEE Workshop on Signal Processing Systems*.

Bjerknes, G., & Bratteteig, T. (1995). User Participation and Democracy: A Discussion of Scandinavian Research on Systems Development . *Scand. J. Inf. Syst. 7, 1 (Apr. 1995),*, 73-98.

Bjorntorp, P. (2001). Do stress reactions cause abdominal obesity and comorbidities? *Obesity reviews, 2(2)*, 73-86.

Blascovich, J., Loomis , J., Beall, A. C., Swinth, K. R., Bailenson, J. N., & Hoyt, C. L. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological inquiry*, 103-124.

Blumenthal, J. A., Sherwood, A., Babyak, M. A., Watkins, L. L., Waugh, R., & Georgiades, A. (2005). Effects of exercise and stress management training on markers of cardiovascular risk in patients with ischemic heart disease: A randomized controlled trial. *Journal of the American Medical Association, 293*, 1626–1634.

Bodker, K., Kensing, F., & Simonsen, J. (2009). Participatory IT Design: Designing for Business and Workplace Realities. *MIT Press Ltd*.

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

Brandt, E., Johansson, M., & Messeter, J. (2005). "The design lab: re-thinking what to design and how to design. *Design Spaces. EDITA IT Press,*, 34-43.

Braun, M., Weiser, S., & Pfleging, B. (2018). A comparison of Emotion Elicitation Methods for Affective Driving Studies. *Association of Computing Machinery, New York*.

Breen, R. L. (2006). A Practical Guide to Focus-Group Research. *Journal of Geography in Higher Education*, 463-475.

Brotman, D. J., Golden, S. H., & Wittstein, I. S. (2007). *The cardiovascular toll of stress.* Lancet, 370.

Bullinger, A. H., Hemmeter, U. M., Stefani, O., Angehrn, I., Mueller-Spahn, F., Bekiaris, E., . . . Mager, R. (2005). Stimulation of cortisol during mental task performance in a provocative virtual. *Psychophysiol. Biofeedback*, 205–216.

Calvo, M. G., & Gutiérrez-García, A. (2016). Cognition and stress, stress concepts. *Cogn Emot Behav Handb Stress*.

Carneiro et al. (2012). Multimodal behavioral analysis for non-invasive stress detection. 13376-13389.

Carroll E., I. (2010). Emotion Theory and Research: Highlights, Unanswered Questions, and Emerging Issues. *National Library of Medicine*. Retrieved from 10.1146/annurev.psych.60.110707.163539

Carroll, D., Turner, J. R., & Rogers, S. (1987). Heart rate and oxygen consumption during mental arithmetic, a video game, and graded static exercise. Psychophysiology 24,. 112-118.

Castelli, N. (2020). Designing human-centered systems for the Internet of Things. 42-44.

Chen, C. M., Anastasova, S., Zhang, K., Rosa, B. G., Assender, H. E., Yang, G. Z., & Lo, B. P. (2019). Towards wearable and flexible sensors and circuits integration for stress monitoring. *IEEE Journal of Biomedical and Health Informatics,*, 2208-2215.

Chirico, A., Yaden, D. B., Riva, G., & Gaggioli, A. (2016). The potential of virtual reality for the investigation of awe. *Front. Psychol. 7(2016)*.

Cho, D., Ham, J., Oh, J., Park, J., Kim, S., Lee, N. K., & Lee, B. (2017). *Detection of stress levels from biosignals measured in virtual reality environments using a kernel-based extreme learning machine.*

Choi , J., Ahmed, B., & Gutierrez-Osuna., R. (2012). *Development and Evaluation of an Ambulatory Stress Monitor Based on wearable sensor.*

Choi, J., Lee, K. K., & Choi, J. (2019). Determinants of user satisfaction with mobile vr headsets: the human factors approach by the user reviews analysis and product lab testing. *Int J Cont.*

Christian, A. G. (2018). Analysis of physiological signals to measure physiological stress.

Cinaz, B., Arnrich, B., La Marca, R., & Tröster, G. (2013). Monitoring of mental workload levels during an everyday life office-work scenario. *Personal Ubiquitous Comput*, 229-239.

Cipresso, P., Giglioli, I. A., Raya, i., & Riva, G. (2011). The Past, Present, and Future of Virtual and Augmented Reality Research: A Network and Cluster Analysis of the Literature. *Frontiers in Psychology*.

Clarke, V., Braun, V., & Hayfield, N. (2015). Thematic Analysis Qualitative psychology: A practical guide to research methods. 248.

Colligan, T., & Higgins, E. (2006). Workplace stress. *Journal of Workplace Behavioral Health* .

Cosemans, B., Marlen, C., Robert, G., Danny, F., & Van, K. (2014). Calculating the cost of work-related stress and psychosocial risks. *European Agency for Safety and Health at Work*. doi:http://dx.doi.org/10.2802/20493.

Cottam, H., Burns, C., Vanstone, C., & Winhall, J. (2006). RED PAPER 02: transformation design. *Transformation design. RED Paper 02; Design.*

Crabtree, A., Hemmings, T., Rodden, T., Mariani, J., Rodden, T., & Hemmings, T. (2003). Informing the Development of Calendar Systems for Domestic Use. 119-138.

Crifaci, G., Tartarisco, G., Billeci, L., Pioggia, G., Gaggioli, A., & Billeci, L. (2012). International Journal of Psychophysiology. *Innovative technologies and methodologies based on integration of virtual reality and wearable systems for psychological stress treatment*, 402.

Crosswell, A. D., & Lockwood, K. G. (2020). Best practices for stress measurement: How to measure psychological stress in health research. *Health psychology open*.

Czekaj, L., Daniszewski, M., & Domaszewicz, J. (2019). Validation of the Aidlab solution for measuring Heart Rate Variability. *Aidlab*.

Davis, M., Eshelman, E. R., & McKay, M. (2008). The relaxation and stress reduction workbook. *New Harbinger Publications.*

De Santos Sierra, A., Sanchez Avila, C., Bailador del Pozo, G., & Guerra Casanova, J. (2011). Stress detection by means of stress physiological template. *2011 Third World Congress on Nature and Biologically Inspired Computing*, 131–136.

DiDomenico, A., & Nussbaum, M. A. (2008). Interactive effect of physical and mental workload on subjective workload assessment. Int. J. Indus. Ergon. 38,.

Dindler, C., & Iversen, O. S. (2007). Fictional Inquiry-Design collaboration in a shared narrative space. *CoDesign. 3, 4 (Dec. 2007)*, 213-234.

Dindo, L., & Fowles, D. C. (2008). The skin conductance orienting response to semantic stimuli: Significance can be independent of arousa. 111-118.

Dohrenwend, B. P. (2000). The role of adversity and stress in psychopathology: Some evidence and its implications for theory ad research. *Journal of Health and Social Behavior 41*, 1–19.

Dourish, P. (2003). The appropriation of interactive technologies: Some lessons from placeless documents. *Computer Supported Cooperative Work (CSCW), 12(4),*, 465-490.

Dugosh, K. L., Paulus, P. B., Roland, E. J., & Yan, H. C. (2000). Cognitive stimulation in brainstorming. *Journal of personality and social psychology*, 722.

Eikenhout, L. M. (2017). Electrodermal activity in relation to self-reported stress and emotion intensity in individuals: an exploration of intra-subject relations through experience sampling and wearable sensors.

Estupinan, S., Rebelo, F., Noriega, P., Ferreira, C., Duarte, E., & Ferrreira, C. (2014). Can virtual reality increase emotional responses (Arousal and Valence)? A pilot study.

Empatica, i. (2015, september 15). *e4 wristband*. Retrieved from empatica: www.empatica.com

Fernand. (2007, September 2021). Retrieved from Centre for Studies on Human Stress. How to measure stress in humans. Quebec: http://www.stresshumain.ca/documents/pdf/Mesures%20physiologiques/

Flaxman, P. E., & Bond, F. W. (2010). Worksite stress management training: Moderated effects and clinical significance. *Journal of Occupational Health Psychology*, 347-358. Retrieved from https://doi.org/10.1037/a0020522

Gendron, M. (2010). Defining emotion: A brief history. *Emotion Rev 2(4)*, 371–372.

Geršak, G. (2020). Electrodermal activity-a beginner's guide. Elektrotehniski vestnik,. 175-182.

Giakoumis, D., Tzovaras, D., & Hassapis, G. (2013). Subject-dependent biosignal features for increased accuracy in psychological stress detection. *Int. J. Human–Comput. Stud*, 425-439.

Gjoreski, M., Gjoreski, H., Luštrek, M., & Gams, M. (2016). Continuous stress detection using a wrist device: in laboratory and real life. *In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct (Heidelberg: ACM),*.

Goode, L. (2019). Get Ready to Hear a Lot More About 'XR'. *wired*.

Gruen, D., Rauch, T., Redpath, S., & Ruettinger, S. (2002). The use of stories in user experience design. *International Journal of Human-Computer Interaction,*, 503-534.

Guitard, T., Bouchard , S., Bernier, F., Dumoulin , S., Laforest, M., & Bernier, F. (2010). The potential of stress indiced with 3-D videogames for teaching stress management skills to soldiers," in Proceedings 44th Annual Conference of the Association for Behavioral and Cognitive Therapy. 18-21.

Halskov, K., & Dalsgaard, P. (2006). Inspiration Card Workshops. *Proceedings of the 6th Conference on Designing Interactive Systems*, 2-11.

Harvard. (2021, February 15). *Harvard health publishing*. Retrieved from Health Harvard education: https://www.health.harvard.edu/mind-and-mood/protect-your-brain-from-stress

Hayashi et al. (2012). An fMRI study of brain processing related to stress states. *World Automation Congress (WAC)*, 1-6.

Hayes. (2018). Design, Action and Practice: Two Branches of the Same Tree. *A Practice-based Perspective on the Design and Use of IT Artifacts* (pp. 303-318). Oxford: Oxford, UK: Oxford University Press.

Hayes, G. R. (2011). The Relationship of Action Research to Human computer Interaction. *ACM Trans. Comput.-Hum. Interact. 18, 3 (Aug.2011*, 15:1–15:20.

Healey, J. A., & Picard, R. W. (2005). Detecting stress during real-world driving tasks using physiological sensors. *IEEE Trans. Intell. Transport. Syst.*, 156-166.

Hernandez, H., Paredes, P., & Roseway, C. A. (2014). Under pressure: sensing stress of computer users. *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*.

Hickey, A. M., & Davis, A. M. (2003). Elicitation technique selection: how do experts do it? *In Proceedings. 11th IEEE International Requirements Engineering Conference*, (pp. 169-178).

Hinchey, P. H. (2008). Action Research. New York, NY: Lang.

Hinde, K., White, G., & Armstrong, N. (2021). Wearable devices suitable for monitoring twenty four hour heart rate variability in military populations.

Holtzblatt, K., & Beyer, H. (1997). Contextual design: defining customer-centered systems. *Elsevier.*

Houde, S., & Hill, C. (1997). What do prototypes prototype? *In Handbook of human-computer interaction*, 367-381.

Hovsepian, K., al'Absi, M., Ertin, E., Kamarck, T., Nakajima, M., & Kumar, S. (2015). Stress: towards a gold standard for continuous stress assessment in the mobile environment. *in Proceedings of the 2015 ACM International Joint*, 493-504.

Howard, T. (2014). Journey mapping: A brief overview. Communication Design Quarterly Review. 10-13.

Humphrey, J. (2002). *Teenagers Will be Teenagers: Teenagers Under Stress.* New York: Nova Science Publishe.

Jallais, C., & Gilet, A. L. (2010). : Inducing changes in arousal and valence: comparison of two mood induction procedures. 318-325.

Jamal, M. (2007). Job stress and job performance controversy revisited: an empirical examination in two countries. *Int. J. Stress Manage. 14:175 doi: 10.1037/1072-5245.14.2.175*.

Javier, M. (2013, April 02). *A map of the customer experience: Customer journey map*. Retrieved from Javiermegias: https://javiermegias.com/blog/2013/04/customer-journey-map-mapa-experiencia-cliente/

Jones, M. K., Latreille , P. L., & Sloane, P. J. (2016). Job anxiety, work-related psychological illness and workplace performance. *Brit. J. Indus. Relat. 54, doi: 10.1111/bjir.12159*, 742–767.

Jönsson, P., Wallergård, M., Österberg, K., Hansen, A. M., Johansson, G., & Karlson, B. (2010). Cardiovascular and cortisol reactivity and habituation to a virtual reality version of the Trier Social Stress Test: a pilot study. 1397-1403.

Jørgensen, A. (1984). "On the psychology of prototyping," in Approaches to Prototyping. *Springer Verlag*, 278-289.

Jungk, R., & Mullert, N. (1996). Future Workshops: How to Create Desirable Features. *Institute for Social Inventions.*

Kaklauskas, A., Zavadskas, E. K., Seniut, M., Dzemyda, G., Stankevic, V., Simkevičius, C., . . . Gribniak, V. (2011). Web-based biometric computer mouse advisory system to analyze a user's emotions and work productivity. *Eng. Appl. Artif. Intell*, 928-945. doi:10.1109/CBMS.2013.6627790

Kambhatla, Nandakishore, & Leen, &. K. (1997, October). Dimension Reduction by Local Principal Component Analysis. *Neural Computation 9, no.7*, 1493-516. Retrieved from http://dx.doi.org/10.1162/neco.1997.9.7.1493.

Karim, N., Hasan, J. A., & Ali, S. S. (2011). Heart rate variability: a review. *J. Basic Appl. Sci*, 71-77.

Karthikeyan, P., Murugappan, M., & Yaacob, S. (2011). A review on stress inducement stimuli for assessing human stress using physiological signals. *In 2011 IEEE 7th International Colloquium on Signal Processing and its Applications* , 420-425.

Keles, H. O. (2014, April 26). Multimodality mapping approach for evolving functional brain connectivity patterns: A fNIRS-EEG study. Optica. Retrieved from https://opg.optica.org/abstract.cfm?uri=biomed-2014-BT5B.2

Kelley, T. (2005). The ten faces of innovation: IDEO's strategies for defeating the devil's advocate and driving creativity throughout your organization. .

Kensing, F., & Blomberg, J. (1998). Participatory Design: Issues and Concerns. *Computer Supported Cooperative Work, 7(3–4),*, 167-185.

Kim, H., Kim, D., Kim, S., Chung, W., Park, K., Kim, J., & Jeon, H. (2021). *Effect of virtual reality on stress reduction and change of physiological parameters including heart rate variability in people with high stress: an open randomized crossover trial*.

Kim, J., Park, J., & Park, J. (2020). Development of a statistical model to classify driving stress levels using galvanic skin responses. *Hum. Factors Ergon. Manuf. Serv. Ind*.

Kocielnik et al. (2013). Smart technologies for long-term stress monitoring at work. *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems*.

Koo, Y., & Ahn, H. (2018). Analysis on the utilization of co-design practices for developing consumer-oriented public service and policy focusing on the comparison with western countries and South Korea. *In ServDes2018. Service Design Proof of Concept, Proceedings of the ServDes. 2018 Conference* (pp. 281-297). Italy: Linköping University Electronic Press.

Kranjec, J., Beguš, S., Geršak, G., & Drnovšek, J. (2014). Non-contact heart rate and heart rate variability measurements. *A review. Biomedical signal processing and control, 13,*, 102-112.

Kristina, G. (2022, November 06). *Sensor Detects Emotions through the Skin*. Retrieved from MIT Technology Review: https://www.technologyreview.com/news/421316/sensor-detects-emotions-through-the-skin/

Kurniawan, H., Maslov, A. V., & Pechenizkiy, M. (2013). Stress detection from speech and galvanic skin response signals. *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems*, 209–214.

Kusserow, M., Amft, O., & Tröster, G. (2012). Monitoring stress arousal in the wild. . *IEEE Pervasive Computing*, 28-37.

Lazarus, R., & Folkman, S. (1984). *Stress, Appraisal, and Coping.* New York, USA: Springer Publishing Company.

Lewrick , M.; Link, P.; Leifer, L. (2018). The design thinking playbook: Mindful digital transformation of teams, products, services, businesses and ecosystems.

Li, H., Wu, J., Gao, Y., Shi, Y., Wu, J., & Li, H. (2016). . Examining individuals' adoption of healthcare wearable devices: An empirical study from privacy calculus perspective. 8-17.

Liao et al. (2005). A real-time human stress monitoring system using dynamic bayesian network. *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 70–70.

Longhurst, R. (2010). Semi-structured Interviewing and Focus groups. *Key methods in geography. N.J. Clifford, S. French, and G. Valentine,eds, SAGE*.

Lu, M. (2011). Magic Number 7: A Behavioral Economic Analysis of Miller's Rule. *Journal of Chongqing Technology and Business University, Social Science Edition*.

Lupien, S. J., & Seguin, F. (2013). How to Measure Stress in Humans. *Centre for Studies on Human Stress*.

Mack, N. (2011). Qualitative Research Methods: A Data Collector's Field Guide. *North Carolina: Family Health International.*

Makela, V. e. (2020). Virtual field studies: conducting studies on public displays in virtual reality. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI 2020).* New York: ACM.

Mäkelä, V., Prange, S., Rodriguez, S. D., Rivu, R., Piening, R., Zhou, Y., & Alt, F. (2021). Remote VR Studies--A Framework for Running Virtual Reality Studies Remotely Via Participant-Owned HMDs. arXiv preprint arXiv:2102.11207.

Malik, M. U., Chaudhry, N. M., & Malik, K. S. (2013). Evaluation of efficient requirement engineering techniques in agile software development. *International Journal of Computer Applications*, 82(3).

Manakhov, P., & Ivanov, V. D. (n.d.). Defining usability problems. *In Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, (pp. 3144-3151).

Mardini, W., Ali, G. A., Magdady, E., Al-momani, S., Ali, A. G., & Mardini. (2018). Detecting Human Emotions using electroencephalography (EEG) using dynamic programming approach.

Masera, G., & Tavakoli, M. (2020). Design and development of a wearable device for continuous monitoring of vital signs. 24-25.

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: a review. Cognition Emotion. 209–237.

McDuff et al. (2012). an intelligent system for emotional memory. *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems*.

McDuff, D., Gontarek, S., & Picard, R. (2014). Remote measurement of cognitive stress via heart rate variability. *In 2014 36th annual international conference of the IEEE engineering in medicine and biology society* , (pp. 2957-2960).

McEwen, B. (2000). The neurobiology of stress: from serendipity to clinical relevance. *Brain Research*.

McFarland, R. A. (1985). Relationship of skin temperature changes to the emotions accompanying music. Biofeedback and Self-regulation.

Miaskiewicz, T., Sumner, T., & Kozar, K. A. (2008). A latent semantic analysis methodology for the identification and creation of personas. *In Proceedings of the SIGCHI conference on human factors in computing systems*, (pp. 1501-1510).

Mika, S., Ratsch, G., Weston, J., Scholkopf, B., & Mullers, K. (1999). Fisher discriminant analysis with kernels, in: Neural Networks for Signal Processing IX. *Proceedings of the 1999 IEEE Signal Processing Society Workshop*.

Milczarek et al. (2009). *stress at work — facts and figures.* Luxembourg: Office for Official Publications of the European Communities.

Mishra, B., Mehta, S., Sinha, N., Shukla, S., Ahmed, N., & Kawatra, A. (2011). *Evaluation of work place stress in health university workers: a study from rural India.* Indian J. Commun. Med.: Official Publ. Indian Assoc. Preventive Soc. Med.

Mitchell, V., Ross, T., May, A., Sims, R., Park, C., & Ross, T. (2016). Empirical investigation of the impact of using co-design methods when generating proposals for sustainable travel solutions. 205-220.

Morgan, D. L. (1996). *Focus Groups as Qualitative Research SAGE publications*.

Morgan, D. L., & Krueger, R. A. (1993). When to use focus groups and why. *Successful focus groups: Advancing the state of the art. Sage*, 3-19.

Mulder, S., & Yaar, Z. (2006). *The user is always right: A practical guide to creating and using personas for the web.* New Riders.

Nelson, D. L., & Simmons, B. L. (2003). *Health psychology and work stress: A more positive approach.*

Newman , M. W., & Landay , J. A. (2000). Sitemaps, Storyboards, and Specifications: A Sketch of Web Site Design Practice. Proceedings of Designing Interactive Systems. 263-274.

Nielsen, J. (1994). Estimating the number of subjects needed for a thinking aloud test. . *International Journal of Human-Computer Studies,*, 385-397.p

Okada et al. (2013). Wearable ECG recorder with acceleration sensors for monitoring daily stress. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 4718–4721.

Okada, Y., Yoto, T. Y., Suzuki, T., Sakuragawa, S., & Sugiura, T. (2013). Wearable ECG recorder with acceleration sensors for monitoring daily stress: office work simulation study. *Annu Int Conf IEEE Eng Med*.

Ollander, S., Godin, C., Campagne, A., & Charbonnier, S. (2016). of wearable and stationary sensors for stress detection,. *IEEE International Conference on Systems, Man, and Cybernetics (SMC) (Budapest:IEEE),*.

Olson, G. M., Duffy, S. A., & Mack, R. L. (2018). Thinking-out-loud as a method for studying real-time comprehension processes. *In New methods in reading comprehension research* , 253-286.

Otto & Wood . (2001). Product Design: Techniques in Reverse Engineering and New Product Design. *Prentice Hall*.

Otto, B., Riepl, R., Klosterhalfen, S., & Enck, P. (2006). Endocrine correlates of acute nausea and vomiting. *Auton Neurosci*.

Otto, K. N. (2003). Product design: techniques in reverse engineering and new product development.

Oximetry, P. (2022). *NONIN*. Retrieved from Pulse Oximetry: https://www.nonin.com/technologies/pulse-oximetry/

Palanisamy, K., Murugappan, M., & Yaacob, S. (2013). Multiple physiological signal-based human stress identification using non-linear classifiers. *Electron. Electr. Eng*, 80-85.

Parak, J., & Korhonen, I. (2013). Accuracy of Firstbeat Bodyguard 2 Beat-to-Beat Heart Rate Monitor. *Firstbeat Technology Ltd.; Tampere,* .

Parent, M., Albuquerque, I., Tiwari, A., Cassani, R., Gagnon, J., Lafond, D., & Falk, T. (2020). Pass: a multimodal database of physical activity and stress for mobile passive body/brain-computer interface research. *Frontiers in Neuroscience*, 14, 542934.

Park, K. H., Kim, D.-h., Kim, S. K., Yi, Y. H., Jeong, J. H., Chae, J., . . . Roh, &. H. (2015, Sep 5). The relationships between empathy, stress and social support among medical students. *International Journal of Medical Education*. Retrieved from 10.5116/ijme.55e6.0d44

Peake, J. M., Kerr, G., & Sullivan, J. P. (2018). A critical review of consumer wearables, mobile applications, and equipment for providing biofeedback, monitoring stress, and sleep in physically active populations. *Frontiers in physiology, 9, 743.*

Pedersen, B. K., & Saltin, B. (2015). Exercise as medicine-evidence for prescribing exercise as therapy in 26 different chronic diseases. *Scand. J.Med. Sci. Sports 25*.

Peternel, K., Pogačnik, M., Tavčar, R., & Kos, A. (2012). A presence-based context-aware chronic stress recognition system. *Sensors*, 15888-15906.

Peuscher, J. (2012). Galvanic skin response (GSR). *Tech. rep.*

Pfaff, M. S. (2012). Negative affect reduces teamawareness: the effects of mood and stress on computer-mediated team communicat. *Hum. Factors 54 doi: 10.1249/01.MSS.0000142403.93205.35*, 560–571.

(2015). *Physiology.* Retrieved from http://www.oxforddictionaries.com/definition/english/physiology

Picard, R. W. (2009). Future affective technology for autism and emotion communication. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 3575-3584.

Poh, M. Z. (2011). continuous assessment pf epileptic seizures with wrist-worn biosensors. *ph.D. dissertation, Massachusetts institute of technology*.

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*, (pp. 1243-1252).

Porter, A. M., & Goolkasian, P. (2019). Video games and stress: how stress appraisals and game content affect cardiovascular and emotion outcomes.

Puccio, G. J. (2017). From the dawn of humanity to the 21st century: Creativity as an enduring survival skill. *The Journal of Creative Behavior*, 330-334.

Rahma, O. N., Putra, A. P., Rahmatillah, A., Putri, Y. S., Fajriaty, N. D., Ain, K., & Chai, R. (2022). Electrodermal activity for measuring cognitive and emotional stress level. Journal of Medical Signals and Sensors, .

Ramos, J. A., & Sharma, M. (2022). *Practical Stress Management: A comprehensive workbook.* Academic Press.

Ramos, J., Hong, J. H., & Dey, A. K. (2014). *Stress recognition – a step outside the lab.* SCITEPRESS – Science and Technology Publication. doi:10.5220/0004725701070118

Random Seed Games. (2015). Timeframe. Random Seed Games.

Ray, W. J., & Cole, H. W. (1985). EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes. *Science*, 750-752.

Red Barrels Games. (2013). Outlast Red Barrels Games.

Reed, M. J., Robertson, C. E., & Addison, P. S. (2005). Heart rate variability measurements and the prediction of ventricular arrhythmias. 87-95.

Reisman, S. (1997). Measurement of physiological stres. *In Proceedings of the IEEE 23rd Northeast Bioengineering Conference*, 21-23.

Renaud, P., & Blondin, J. P. (1997). The stress of Stroop performance: Physiological and emotional responses to color–word interference, task pacing, and pacing speed. *International Journal of Psychophysiology*, 27(2), 87-97.

RETTIG, M. (1994). Prototyping for tiny fingers. *Communications of the ACM*, 21-27.

Ritchie, J. (2003). The Applications of Qualitative Methods to Social Research. *In: J. Ritchie & J.Lewis, eds. Qualitative Research Practice: A Guide for Social Science Students and Researchers.*, 24-46.

Ritter, S. M., & Mostert, N. (2017). Enhancement of creative thinking skills using a cognitive-based creativity training. *Journal of Cognitive enhancement,*, 243-253.

Rivu, R., Jiang , R., Mäkelä, V., Hassib, M., Alt, F., & Alt, F. (2021). Emotion Elicitation Techniques in Virtual Reality. *IN IFIP Conference on Human-Computer Interaction*, 93-114.

Rodrigues, A. K., Armiliato, M. J., Gonçalves, T. R., De castro, E. K., Guerine , R. P., Zibetti, M. R., & Rigo, S. J. (2022). Dispositivos Vestíveis para monitoramento psicofisiológico do estresse: uma revisão narrativa. Revista Brasileira de Terapias Cognitivas. 60-68.

Roger, C. (2006, November November 12, 2022). *Smart Product decision*. Retrieved from http://blog.cauiri.org/2006/02/ieee_ definition_of_requirement.htm

Roh, T., Bong, K., Hong, S., Cho, H., Member, S., & Yoo, H. J. (2012). Wearable mental-health monitoring platform with independent component analysis and nonlinear chaotic analysis. *34th Annual International Conference of the IEEE EMBS, San Diego*, 4541–4544.

Rosenberg, W., Chanwimalueang, T., Adjei, T., Jaffer, U., Goverdovsky, V., & Mandic, D. (2017). *Resolving ambiguities in the lf/hf ratio: Lf-hf scatter plots for the categorization of mental and physical stress from HRV*.

Rothbaum , B. O., Hodges, L. F., Ready, D., Graap, K., Alarcon, R. D., & Ready, D. (2001). Virtual reality exposure therapy for Vietnam veterans with posttraumatic stress disorder. *Journal of Clinical Psychiatry*, 617-622.

Rudd, J., Stern, K., & Isensee, S. (1996). *Low vs. high-fidelity prototyping debate.*

Russell, L. (2019). The human stress response. *Nat Rev Endocrinol*. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8536782/#bib173

Salovaara, A., & Tamminen, S. (2009). Acceptance or appropriation? A design-oriented critique of technology acceptance models. *In Future interaction design II Springer London*, 157-172.

Sanders, T. A., & Cairns, P. (2010). Time perception, immersion and music in videogames.

Sano, A., & Picard, R. W. (2013). Stress recognition using wearable sensors and mobile phones. *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, 671–676.

Sauer, Franke , & Ruettinger. (2008). Designing interactive consumer products: utility of paper prototypes and effectiveness of enhanced control labelling. *Applied Ergonomics*, 71-85.

Scarpina, F., & Tagini, S. (2017). The stroop color and word test. *Frontiers in psychology*, 8, 557.

Schäfer, A., & Vagedes, J. (2013). A review on studies comparing photoplethysmographic technology with an electrocardiogram. *International Journal of Cardiology.*

Schmidt, P., Reiss, A., Duerichen, R., Marberger, C., Van Laerhoven, K., & Marberger, C. (2018). *In Proceedings of the 20th ACM International Conference on Multimodal Interaction.*

Schneiderman, N., Ironson, G., & Siegel, S. D. (2005). Stress and health: psychological, behavioral, and biological determinants. *Annual review of clinical psychology*, 1, 607.

Schrage, M. (1993). The culture(s) of prototyping. *Design Management Journal (FormerSeries)* , 55–65.

Schüll, N. D. (2016). Data for life: Wearable technology and the design of self-care. 317-333.

Sefelin, R., Tscheligi, M., & Giller, V. (2003). Paper prototyping - what is it good for? A comparison of paper- and computer-based prototyping. *Paper presented at the ACM Conference on Computer Human Interaction.*

Selye. (1956). The Stress of Life. *McGraw-Hil Edition, New York* .

Seo, S. H., & Lee, J. (2010). Stress and EEG. *Convergence and Hybrid Information Technologies*.

Shahsavarani, A. M., Azad Marz Abadi, E., & Hakimi Kalkhoran, M. (2015). Stress: Facts and theories through literature review. *International Journal of Medical Reviews*, 230-241.

Shapiro J P. (1993). "Screaming neon wheelchair," in No pity: People with disabilities forging a new civil rights movement. 211-236.

Sharma, N., & Gedeon, T. (2012). Objective measures, sensors and computational techniques for stress recognition and classification. *Comput. Methods Programs Biomed*, 1287-1301.

Sharma, N., & Gedeon, T. (2012). Objective measures, sensors and computational techniques for stress recognition and classification. *Comput. Methods Programs Biomed*.

Sharma, N., Dhall, A., Gedeon, T., & Goecke, R. (2014). Thermal spatio-temporal data for stress recognition. *EURASIP J. Image Video Process*.

Sheldom, c., Kamarck, T., & Mermelstein, R. (1983). A Global Measure of perceived stress. *Journal of Health and Social Behavior*.

Singh, N., Moneghetti, K. J., Christle, J. W., Hadley, D., Froelicher, V., & Plews, D. (2018). Heart rate variability: An old metric with new meaning in the era of using health technologies for health and exercise training guidance. *Arrhythm Electrophysiol*.

Smets, E., De Raedt, W., & Van Hoof, C. (2018). Into the wild: the challenges of physiological stress detection in laboratory and ambulatory settings. *IEEE journal of biomedical and health informatics*, 463-473.

Sourina, O., & Liu, Y. (2014). EEG-enabled affective human–computer interfaces. *Universal Access in Human–Computer Interaction: Design and Development Methods for Universal Access, Lecture Notes in Computer Science*, 536-547.

Stevens, G., Rohde, M., Korn, M., Wulf, V., Pipek, D., & Randall, M. (2018). Grounded Design: A Research Paradigm in Practice-Based Computing. *A practice based Perspective on the Design and Use of IT Artifacts*, 23-46.

Suresh, L. V., Chandrasekar, D., & Punita, D. P. (2016). “Mental.

Suri, J. S. (2006). Heart rate variability. *A review. Med. Bio. Eng.*

Sztajzel, J. (2004). Heart rate variability: a noninvasive electrocardiographic method to measure the autonomic nervous system. *Swiss medical weekly 134*, pp. 35-36.

Technology, N. p. (2016, february 10). *nonin*. Retrieved from nonin: www.nonin.com/puresat

Tinder, & Richard, F. (2007). *Relativistic Flight Mechanics and Space Travel: A Primer for Students, Engineers and Scientists.* Morgan & Claypool Publishers.

Vachiratamporn, V., Legaspi, R., Moriyama, K., & Numao, M. (2013). Towards the design of affective survival horror games: an investigation on player affect. *Humaine Association Conference on Affective Computing and Intelligent Interaction(IEEE)*, 576–581.

Vines, J., Clarke, R., Wright, P., McCarthy, J., Olivier, P., & Clarke, R. (2013). Configuring Participation: On How We Involve People in Design. *Proceedings of the SIGCHI Conference on Human Factors in Computing systems*, 429-438.

Vinkers, C. H., Penning, R., Hellhammer, J., Verster, J. C., Klaessens, J. H., Olivier, B., & Kalkman, C. J. (2013). The effect of stress on core and peripheral body temperature in humans. 520-530.

Vitug, G. J., Toplis, J. A., & Geli, R. I. (2016). Analysis of Player Stress Levels Using the Affectiva Q Sensor. *Philippine Information Technology Journal,*.

Vitug, G., Sabnis, N., Shi, Y., Di Pasqua, A., Bowman, W. P., & & Lacko, A. (2014). EVALUATION OF A.

Vizer, L. M., Zhou, L., & Sears, A. (2009). Automated stress detection using keystroke and linguistic features. *Int. J. Human–Comput. Stud.*, 870-886.

Walker, A. (2021, August 02). *esmartr*. Retrieved from esmartr.com: https://esmartr.com/blogs/articles/what-is-cognitive-stress#:~:text=Cognitive%20stress%20refers%20to%20any,feel%20stressed%20out%20or%20anxious.

Wang, Y., Wang, H., Xuan, J., Leung, D. Y., Xuan, J., & Wang, Y. (2020). Powering future body sensor network systems: A review of power sources. *Biosensors and Bioelectronics, 166, 112410.*

Wibawa, A. D., & Purnomo, M. H. (2016). Physiological Pattern of Basic Human Emotion State based on SPO2 Sensor. *Int. J. Syst. Signal Control Eng. Appl,*, 55-59.

Wijsman et al. (2013). Wearable physiological sensors reflect mental stress state in office-like situations. *Humaine Association Conference on Affective Computing and Intelligent Interaction, IEEE*, 600-605.

Wolf, O. T. (2009). Stress and memory in humans: twelve years of progress? *Brain research, 1293*, 142-154.

Wu, H.-K., Lee, S. W.-Y., Chang, H.-Y., & Liang, J.-C. (2013). Current status, opportunities and challenges of augmented reality in education...". *Computers & Education*.

Wulf, V., Müller, C., Pipek, V., Randall, D., Rohde, M., & Stevens, G. (2015). Practice-Based Computing: Empirically Grounded Conceptualizations Derived from Design Case Studies. *Designing Socially Embedded Technologies in the Real-World. Springer London.*, 111-150.

Wulf, V., Rohde, M., Pipek, V., Stevens, G., Pipek, V., & Rohde, M. (2011). Engaging with Practices: Design Case Studies As a Research Framework in CSCW. *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work*, (pp. 505-512). New York, NY, USA.

Xu, L. H., Koch, K. L., Summy-Long, J., Stern, R. M., Seaton, J. F., & Harrison, T. S. (1993). Hypothalamic and gastric myoelectrical responses during vection-induced nausea in healthy Chinese subjects. *Am J Physiol.*

Zhai, J., & Barreto, A. (2006). *Stress Recognition Using Non-invasive Technology In FLAIRS conference*, (pp. 395-401).

Zhai, J., & Barreto, A. (2006). Stress detection in computer users based on digital signal processing of noninvasive physiological variables. *IEEE Engineering in Medicine and Biology Society, Annual Conference*, 1355–1358.

Zijderveld, G. (2019). *Exploring Advanced Emotion AI with Gabi Zijderveld*. Retrieved from verypossible: https://www.verypossible.com/insights/exploring-advanced-emotion-ai-with-gabi-zijderveld

Zimmermann et al. (2003). Affective computing – a rationale for measuring mood with mouse and keyboard. *Int. J. Occupational Safety Ergonomics*, 539-551.

Appendix

- User Interview Questions

Interview Questions

1. Can you please introduce yourself (Name, your research background, current designation/research position, research group and any other personal thing about yourself you will like me to know).
2. What are your research interests and for how long have you been doing research in this area?
3. What is the motivating factor that drives you to research more in this area?
4. I read about your previous research work and I will like to know what was the motivation behind developing the Android App for measuring reward delay (Belohnungsaufschub)?
5. How did you recruit/select subjects for the testing of the android app idea?
6. Could you please tell me some of the research methods you applied when developing the android App.?
7. Since I'm working on a stress elicitation App concept, I will like to ask what data collection methods or techniques you think will be effective to get stress identification parameters like (heart rate, electro-dermal activities, acceleration from people with the use of an App or software.
8. Could you please tell some specific digital tools (Apps, website, software) you used for your work to obtain health related data from your subjects?
9. As a researcher, before using an App or software for research or data collection, what are some functionalities you expect the App to have before you decide to use it.
10. Could you please tell some of the technical, ethical or usability challenges you face when using these digital tools for your reward delay App research work?
11. How did you overcome these challenges or what alternatives did you use?
12. When testing some of the Apps you developed, what are some of the issues you think users experienced or what are some common feedbacks you got from the users when they tested your application?
13. What are somethings you will do differently in the future to avoid these issues users experienced when testing your application?
14. Generally speaking, what are some of the challenges you think researchers in the digital health & Life sciences research area face when trying to make use of digital applications to gather data from people for their research works.

15. When designing an application or software for researchers to carry out research, what factors do you think a designer must put into consideration in order to ensure the App meets the specific needs of the researchers?
16. How do you think digital technology will shape the field of Life science or digital health in the future?
17. In your opinion, do you think researchers prefer to stick to their traditional method of research or they are open to utilizing new tech tools (Apps, software) for their research works?
18. Do you have any questions or feedback?

Interview Questions

1. Can you please introduce yourself (Name, your research background, current designation/research position, research group and any other personal thing about yourself you will like me to know).
2. What are your research interests and for how long have you been doing research in this area?
3. What is the motivating factor that drives you to research more in this area?
4. I read about your previous research work and I will like to know what was the motivation behind the phone based approach avoidance bias training for smokers?
5. How did you recruit/select subjects for the testing of the android app idea?
6. Could you please tell me some of the research methods you applied when developing this phone approach for smokers?
7. Since I'm working on a stress elicitation App concept, I will like to ask what data collection methods or techniques you think will be effective to get stress identification parameters like (heart rate, electro-dermal activities, acceleration from people with the use of an App or software.
8. Could you please tell some specific digital tools (Apps, website, software) you used for your work to obtain health related data from your subjects?
9. As a researcher, before using an App or software for your research or data collection, what are some functionalities you expect the App to have before you decide to use it?
10. What were some of the technical, ethical or usability challenges you face when using these digital tools for your research work?
11. How did you overcome these challenges or what alternatives did you use?

12. When testing some of the apps or tech artefacts you developed, what are some of the issues you think users experienced or what are some common feedbacks you got from the users when they tested your applications?
13. What are somethings you will do differently in the future to avoid these issues users experienced when testing your application?
14. Generally speaking, what are some of the challenges you think researchers in the digital health & Life sciences research area face when trying to make use of digital applications to gather data from people for their research works.
15. When designing an application or software for researchers to carry out research, what factors do you think a designer must put into consideration in order to ensure the App meets the specific needs of the researchers?
16. How do you think digital technology will shape the field of Life science or digital health in the future?
17. In your opinion, do you think researchers prefer to stick to their traditional method of research or are they open to utilizing new tech tools (Apps, software) for their research works?
18. Do you have any questions or feedback?

Thinking Aloud Workshop Script

Thesis Topic: Designing a Benchmark App for Stress Elicitation

Introduction:

My name is **Charles Ajiboye** and I'm a Masters Student of Human Computer Interaction at the University of Siegen. For my master's thesis, I'm conducting research for the topic mentioned and a thinking aloud workshop is one of the research activities needed in order to design a digital product that not only meets the needs of the users but also has a great user experience.

Today, you will support us in executing some pre-defined task while interacting with the interface of the already designed App. The purpose of this workshop is to identify usability issues users might encounter when interacting with the App.

- Please remember, we are testing the App and not you or your tech savvy skills.

Thinking aloud method training

In this workshop, we are particularly interested in what is going through your mind or what you think about the tasks you will be asked to perform. To achieve this, it will be required of you to talk aloud as you work on the task. In other words, I want you to tell me everything you are thinking from the first point of seeing the statement of the task till you accomplish the task.

It is important for you to talk aloud constantly from the time I give you the task till you have completed it. You don't have to think too much or try to plan what you will say. Just say the words as it comes to your mind. It is only in this way that you can truly support us in highlighting all usability issues that might exist in the App. It is also important for me to mention that the things you will

search for while navigating through the app and things that you will see are as important for our observation as thoughts you are thinking in your mind so please verbalize everything. While you are doing the tasks, I might not be able to answer any questions so as not to interrupt the process. But if you happen to have any questions, please go ahead and ask them anyway. This will help us to learn more about what kinds of questions the Benchmark App for stress elicitation might bring up. I will definitely answer all your questions after the session. In cases where you forget to think aloud, I'll always remind you to "keep talking".

Task 1

You're a researcher who intends to carry out some stress elicitation task in your laboratory. With the use of the Stress Elicitation App. Create two tasks for your participants to engage in. Please note that the task can be under any stress category of your choice

Task 2.

As a researcher after creating some tasks, you realized you will like to edit the tasks in order to add some additional details to it. Please revisit a previously created task and edit it.

Task 3

Your stress elicitation App works in conjunction with some wearable devices e.g (Affectiva and Nonin. Check the connection status of these devices and try to connect the Nonin device if it is not connected initially.

Task 4.

After conducting some task with users, you realized you will like to add additional task to the Cognitive stress category. Please add a new task under the Cognitive stress category.

Task 5

Your participants have conducted all the tasks and as the researcher, you will like to view the readings captured during the tasks. Please open the readings of one of your participants and view the results.

Task 6

After viewing the results please send the readings of one of your participants to your email in CSV format.

Task 7

Log out of the App

ACKNOLWLEDGEMENT OF CONFIDENTIALITY OF RESEARCH DATA

I, **Charles Ajiboye**, Student ID:[redacted], herby agree to treat as confidential all information about the data provided to me for the purpose of my master thesis.

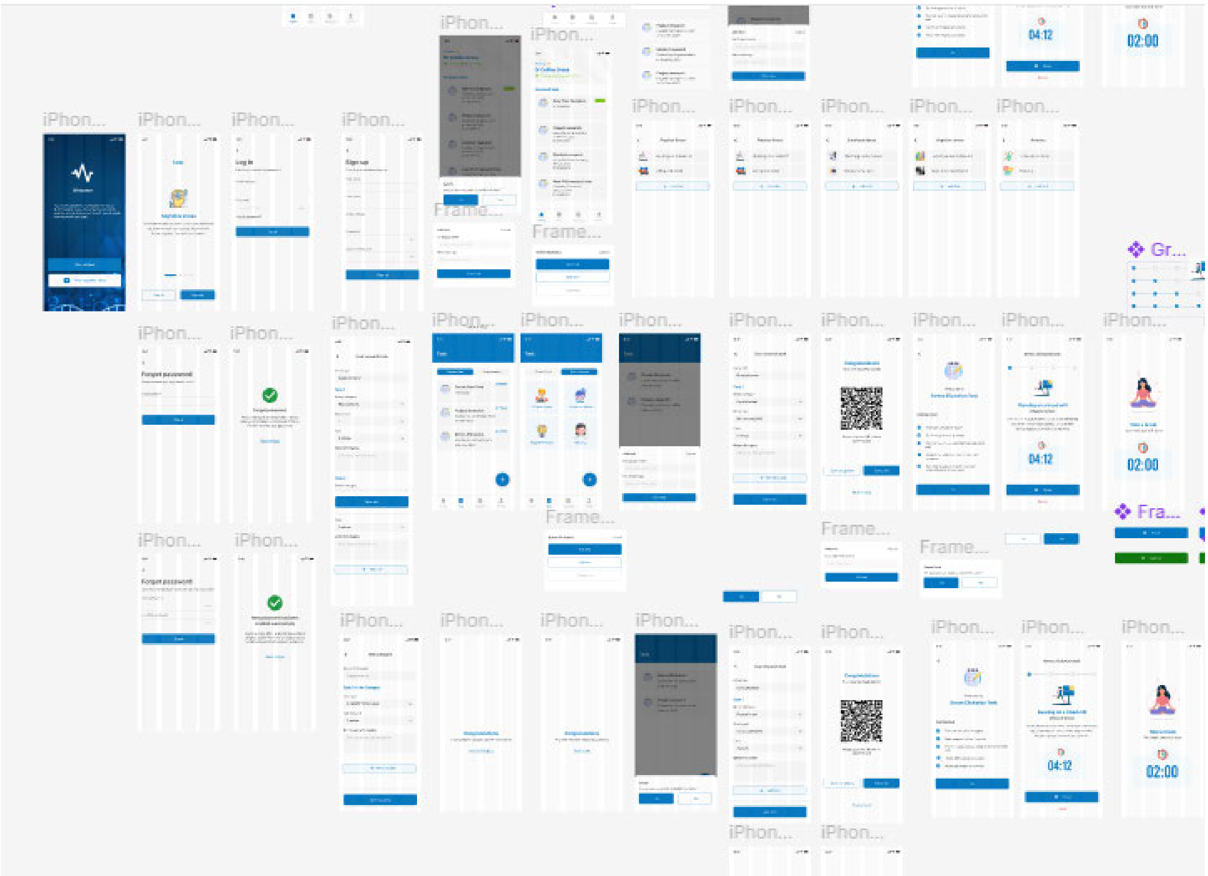
I understand that it would be a violation of policy to disclose such information to anyone without previous written agreement from the responsible persons. Failure to adhere to this policy will result in discipline, up to and including law pursuit.

[redacted]

Participant Signature:

Date: [redacted]

Snapshot of Figma Design Interface



Link to Figma File

https://www.figma.com/file/2MrimV1cgQ7Mv823n9XFt1/Stress-Elicitation-App-new-(1)?node-id=0%3A1&t=V54ltofuoLiq5b8Z-1