NOTTINGHAM TRENT UNIVERSITY  
SCHOOL OF SCIENCE AND TECHNOLOGY

An Investigation into Affective State in Virtual Reality for At-Home Depression and Anxiety Symptom Management

by

Hassaan Naveed

in

2022

Project report in part fulfilment  
of the requirements for the degree of  
Bachelor of Science with Honours

In

Computer Science

I hereby declare that I am the sole author of this report. I authorize the Nottingham Trent University to lend this report to other institutions or individuals for the purposes of scholarly research.

I also authorise Nottingham Trent University to reproduce this report by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purposes of scholarly research.

Hassaan Naveed

# Abstract

This report details the investigation into the possible uses of Virtual Reality in reducing stress in university students who may be suffering with symptoms of depression or anxiety, through the research of Affective State.

An experimental methodology was designed and carried out. Three participants between the ages of 19-24 were acquired and experienced the Nature Treks VR application, using the Oculus Quest 2. Levels of alpha power spectral density with the OpenBCI UltraCortex IV EEG, as well as subjective ratings on the SUDS scale were measured before and after the VR experience. Data was processed and analysed using MATLAB, EEGlab, and R.

It was found that all participants experienced a non-significant reduction in subjective levels of distress on the SUDS scale. Additionally, all participants experienced a non-significant increase in alpha power spectral density. This showed a non-significant, strong positive correlation between increase in alpha power spectral density and decrease in SUDS ratings.

# Acknowledgements

I would like to acknowledge and thank my supervisor, James Lewis for his continued support and helpful insights in this project.

Table of Contents

[Abstract II](#_Toc102998215)

[Acknowledgements III](#_Toc102998216)

[List of Figures VII](#_Toc102998217)

[List of Tables VIII](#_Toc102998218)

[Chapter 1 - Introduction 1](#_Toc102998219)

[1.1 – Introduction to the Project 1](#_Toc102998220)

[1.2 – Depression and Anxiety 2](#_Toc102998221)

[1.3 - Mental Health and COVID-19 3](#_Toc102998222)

[1.4 - Virtual Treatments 3](#_Toc102998223)

[1.5 - Affective State 5](#_Toc102998224)

[1.6 – Aims and Objectives 5](#_Toc102998225)

[1.7 – Intended Scope 6](#_Toc102998226)

[Chapter 2 – Context 7](#_Toc102998227)

[2.1 – Introduction 7](#_Toc102998228)

[2.2 – Virtual Reality 7](#_Toc102998229)

[2.2.1 – Background 7](#_Toc102998230)

[2.2.2 - Clinical Uses 8](#_Toc102998231)

[2.2.3 – Inducing Emotional States 10](#_Toc102998232)

[2.3 – Models of Emotion 10](#_Toc102998233)

[2.3.1 – Discrete Models 11](#_Toc102998234)

[2.3.2 – Dimensional Models 13](#_Toc102998235)

[2.4 – Estimating Affective States 15](#_Toc102998236)

[2.4.1 – Self Reports 15](#_Toc102998237)

[2.4.2 – Behavioural Measures 16](#_Toc102998238)

[2.4.3 - Electrophysiology 16](#_Toc102998239)

[2.5 – Similar Research 18](#_Toc102998240)

[2.5.1 – VR for Affective Research 18](#_Toc102998241)

[2.5.2 – VR for Stress Reduction 19](#_Toc102998242)

[Chapter 3 – New Ideas 21](#_Toc102998243)

[3.1 – Introduction 21](#_Toc102998244)

[3.2 – Planning the Methodology Development 21](#_Toc102998245)

[3.2.1 – Evaluating Stress 22](#_Toc102998246)

[3.2.2 – Emotion Models 23](#_Toc102998247)

[3.2.3 - Evaluating Affective State 24](#_Toc102998248)

[3.2.4 – Population of Interest 25](#_Toc102998249)

[3.2.5 – Hypothesis 26](#_Toc102998250)

[3.2.6 – Experimental Design 26](#_Toc102998251)

[3.2.7 – Apparatus 27](#_Toc102998252)

[3.3 – Evaluation 30](#_Toc102998253)

[3.3.1 – Correlation 30](#_Toc102998254)

[3.3.2 – Statistical Test 30](#_Toc102998255)

[3.3.3 – Success Criteria 31](#_Toc102998256)

[Chapter 4 – Investigation 33](#_Toc102998257)

[4.1 - Introduction 33](#_Toc102998258)

[4.2 – Timescale Planning 33](#_Toc102998259)

[4.2.1 – Stages 33](#_Toc102998260)

[4.2.2 – Timescale 33](#_Toc102998261)

[4.3 – Methodology Development 34](#_Toc102998262)

[4.3.1 – Dependant Variables 34](#_Toc102998263)

[4.3.2 – Hypothesis 35](#_Toc102998264)

[4.3.3 – Participant Gathering 35](#_Toc102998265)

[4.3.4 – Apparatus 36](#_Toc102998266)

[4.3.5 – Experimental Task 38](#_Toc102998267)

[4.3.6 – Processing 39](#_Toc102998268)

[4.3.7 – Analysis and Testing 39](#_Toc102998269)

[4.4 – Procedure 40](#_Toc102998270)

[4.5 – Data Processing 41](#_Toc102998271)

[4.5.1 – Pre-processing 42](#_Toc102998272)

[4.5.2 – Artifact Removal 43](#_Toc102998273)

[4.5.3 – Power Spectral Density 44](#_Toc102998274)

[4.5.4 – SUDS Ratings 45](#_Toc102998275)

[4.5.5 –Developing Statistical Analysis Functions 45](#_Toc102998276)

[4.6 - Testing 47](#_Toc102998277)

[Chapter 5 – Results 49](#_Toc102998278)

[5.1 – Introduction 49](#_Toc102998279)

[5.2 – Analysis of Results 49](#_Toc102998280)

[5.2.1 – Analysis of PSD 49](#_Toc102998281)

[5.2.2 – Analysis of SUDS Ratings 51](#_Toc102998282)

[5.2.3 – Analysis of Descriptive Statistics 53](#_Toc102998283)

[5.2.4 – Statistical Testing 57](#_Toc102998284)

[5.3 – Evaluation 58](#_Toc102998285)

[5.4 – Discussion 60](#_Toc102998286)

[Chapter 6 – Conclusion 61](#_Toc102998287)

[6.1 – Conclusions 61](#_Toc102998288)

[6.2 – Future Work 62](#_Toc102998289)

[6.3 – Legal, Social, Ethical, and Professional Issues 63](#_Toc102998290)

[6.3.1 - Legal Issues 63](#_Toc102998291)

[6.3.2 - Social Issues 63](#_Toc102998292)

[6.3.3 - Ethical Issues 64](#_Toc102998293)

[6.3.4 - Professional Issues 64](#_Toc102998294)

[6.4 – Synoptic Reflection 65](#_Toc102998295)

[References 67](#_Toc102998296)

# List of Figures

[Figure 1: Plutchik's Wheel of Emotions 15](file:///D:\Documents\Work\Uni%20Year%203\FYP\Report\Report%20Draft%20-%20Copy.docx#_Toc102940713)

[Figure 2: Circumplex Model of Emotion 16](file:///D:\Documents\Work\Uni%20Year%203\FYP\Report\Report%20Draft%20-%20Copy.docx#_Toc102940714)

[Figure 3: Oculus Quest 2 29](file:///D:\Documents\Work\Uni%20Year%203\FYP\Report\Report%20Draft%20-%20Copy.docx#_Toc102940715)

[Figure 4: International 10-20 System of Electrode Placement with 8 Electrodes 29](file:///D:\Documents\Work\Uni%20Year%203\FYP\Report\Report%20Draft%20-%20Copy.docx#_Toc102940716)

[Figure 5: OpenBCI UltraCortx IV 29](file:///D:\Documents\Work\Uni%20Year%203\FYP\Report\Report%20Draft%20-%20Copy.docx#_Toc102940717)

[Figure 6: Diagram of Experimental Task 30](file:///D:\Documents\Work\Uni%20Year%203\FYP\Report\Report%20Draft%20-%20Copy.docx#_Toc102940718)

[Figure 7: EEG Recording Before (Left) and After (Right) Bandpass Filter 32](file:///D:\Documents\Work\Uni%20Year%203\FYP\Report\Report%20Draft%20-%20Copy.docx#_Toc102940719)

[Figure 8: EEG Recording Visual Artifact Removal 33](file:///D:\Documents\Work\Uni%20Year%203\FYP\Report\Report%20Draft%20-%20Copy.docx#_Toc102940720)

[Figure 9: Lateral Eye Movement Artifact (Left) and Blinking Artifact (Right) 33](file:///D:\Documents\Work\Uni%20Year%203\FYP\Report\Report%20Draft%20-%20Copy.docx#_Toc102940721)

[Figure 10: Heatmaps of Alpha power of all participants (top to bottom) in baseline recording (left) and experimental recording (right) 34](file:///D:\Documents\Work\Uni%20Year%203\FYP\Report\Report%20Draft%20-%20Copy.docx#_Toc102940722)

[Figure 11: Periodogram of Experimental Recording 35](file:///D:\Documents\Work\Uni%20Year%203\FYP\Report\Report%20Draft%20-%20Copy.docx#_Toc102940723)

[Figure 12: Bar chart to Show Before and After PSD 37](file:///D:\Documents\Work\Uni%20Year%203\FYP\Report\Report%20Draft%20-%20Copy.docx#_Toc102940724)

[Figure 13: Bar Chart to Show Before and After SUDS Ratings 38](file:///D:\Documents\Work\Uni%20Year%203\FYP\Report\Report%20Draft%20-%20Copy.docx#_Toc102940725)

[Figure 14: Bar Chart to Compare Change in PSD vs Change in SUDS Ratings 39](file:///D:\Documents\Work\Uni%20Year%203\FYP\Report\Report%20Draft%20-%20Copy.docx#_Toc102940726)

[Figure 15: Box plots to show outliers in PSD data (left) and SUDS data (right) 40](file:///D:\Documents\Work\Uni%20Year%203\FYP\Report\Report%20Draft%20-%20Copy.docx#_Toc102940727)

[Figure 16: Scattergram to Show Relationship Between Change in PSD and Change in SUDS Ratings 41](file:///D:\Documents\Work\Uni%20Year%203\FYP\Report\Report%20Draft%20-%20Copy.docx#_Toc102940728)

[Figure 17: Scattergram to Show Density Distribution of PSD Before and After (Top Left), SUDS Before and After (Top Right), and Change in PSD vs Change in SUDS (Bottom Middle) 42](file:///D:\Documents\Work\Uni%20Year%203\FYP\Report\Report%20Draft%20-%20Copy.docx#_Toc102940729)

# List of Tables

[Table 1: Table of Results 38](#_Toc102940889)

[Table 2: Table of Success Criteria 44](#_Toc102940890)

# Chapter 1 - Introduction

## – Introduction to the Project

This project is an investigation into the possibility of utilising Virtual Reality for low-effort, at-home symptom management for individuals that may be experiencing depressive or anxious symptoms, through the research of Affective State. This chapter will provide a general overview of the topic and the associated background, as well as the intended scope, aims, and objectives of the project.

In Chapter 2, the in-depth context of the project will be explored, and how they influenced this project idea. This will involve a literature review into the theories and ideas associated with the project, such as the clinical uses of VR and theories of emotion, as well as possible means of estimating emotional states. Finally, this chapter will conclude with a discussion of the strengths and weaknesses of existing research into similar ideas, and how these ideas will influence this project.

Chapter 3 will follow on from the concepts discussed in Chapter 2, discussing the possible tools and technologies that could be used within the investigation, and developing the project idea. Different methodologies will be compared for the preparation of the development of the methodology used within this investigation. Furthermore, methods of assessing this will be created.

Chapter 4 will describe the process of planning the methodology, selecting the appropriate tools, technologies, and instruments to develop a successful experiment. The process of conducting the experiment and processing the data will be detailed, as well as any further design decisions that needed to be made, or any difficulties faced and how they were tackled.

Chapter 5 will discuss the results of the experiment, whether they appear to show success, and what insights and conclusions can be drawn from them. Any weaknesses in the methodology will be highlighted, and how these could have been mitigated will be discussed. The success of the methodology will be evaluated against the metrics designed in Chapter 3.

Finally, Chapter 6 will conclude the project, summarising the findings of the work, what they mean for the future, and how they could be useful. Limitations of the project will be discussed, and how these can lead into future ideas. Finally, a synoptic reflection will detail what I have learnt from undertaking this project, and how it has prepared me for future endeavours.

## 1.2 – Depression and Anxiety

Depression and anxiety are increasingly common mental conditions that affect a large population of the world. They often have severe consequences on an individual’s quality of life and ability to function (Malhi & Mann, 2018). They can be caused by biological or environmental factors, however most often it is a combination of the two, as suggested by the Stress-Diathesis model. This states that an external, environmental stressor is needed to trigger depression, however the required severity of the external stressor is dependent on their level of genetic predisposition (Colodro-Conde, et al., 2018).

Depression and anxiety can manifest in many ways, both psychological and physiological. The DSM-5 lists the two primary symptoms of depression as having a depressed, low mood and a diminished or loss of interest and pleasure in most activities, and other symptoms such as insomnia and feelings of worthlessness, while the main characteristics of anxiety are the presence of excessive worry, which is challenging to control and accompanies with any other symptoms such as restlessness, fatigue, or irritability (American Psychiatric Association, 2013). Depression and anxiety are very often experienced at the same time. Depression can lead to anxiety, and vice versa (American Psychiatric Association, 2013).

The recorded cases of clinical depression have been rising over the years, with an increase of approximately 13.47% recorded cases worldwide between 1990 and 2019 (Institute of Health Metrics Evaluation, 2021). The WHO reports that an estimated 5% of adults globally suffer from depression (World Health Organisation, 2021). It is suggested that university students between the ages of 18-24 are most at risk of developing depression and anxiety, with increased stress and pressure, and being away from home listed as the primary reasons (Mahmoud, et al., 2012).

## 1.3 - Mental Health and COVID-19

The recent COVID-19 pandemic has had a large impact on mental health, disrupting the daily lives of many individuals. Measures taken to reduce the impact of the pandemic, such as self-isolation, quarantining, lockdowns, and working from home have led to increased isolation, feelings of dread and pessimism, poor overall mental health (Kumar & Nayar, 2020). Many of these stressors can have a lasting effect and lead to depressive and anxious symptoms and episodes (Pfefferbaum & North, 2020).

Furthermore, the pandemic has also caused an increased amount of stress and demand on health services (Willan, et al., 2020). The increased demand front-line health workers can cause a great amount of undue stress, which is able to characterise itself as anxiety and depression (Zhang, et al., 2020). From a study conducted in China between January and February 2020, 54% of respondents rated a moderate to severe psychological impact from the COVID-19 outbreak. 29% reported moderate to severe anxiety symptoms, and 17% reported moderate to severe depressive symptoms (Cullen, et al., 2020). Additionally, many hospitals and care facilities are unable to offer adequate, timely mental health services to those already suffering from mental health issues due to logistical issues (Auerbach & Miller, 2020).

## 1.4 - Virtual Treatments

Virtual methods of treatment for a wide range of conditions have been gaining more popularity over recent years, and the COVID-19 pandemic has only accelerated the adoption of such techniques, as in-person methods were more difficult to access and provide (Wosik, et al., 2020).

The primary technology in this area is Virtual Reality. VR headsets have been used for treating many psychological and physiological issues for many years (Valmaggia, et al., 2016). The immersive properties of VR headsets can produce more authentic emotions in individuals than what could be experienced from a pre-recorded video (Quesnel & Riecke, 2018). They are a cost effective and convenient solution for placing an individua within a specific scenario and having a high degree of control over that scenario. Most notably, exposure therapy for phobias and anxiety disorders shows high effectiveness when utilising VR (Maples-Keller, et al., 2017). For certain, unorthodox stimuli, it may be more convenient to place the individual within a safe virtual environment designed for the exposure to that stimulus, as opposed to attempting to do so in a real-life environment, where the degree of control is much lower (Parsons & Rizzo, 2008). The virtual environment can be changed, adapted, or stopped quickly in response to how the individual acts, and therefore poses much less risk to than a physical environment.

The research into the possibilities of VR treatment methods for depression and anxiety are lacking, in comparison to the uses of VR in interventions for phobias, pain management, and PTSD. The small number of studies that have focused on VR as an intervention for depression and anxiety have found the delivery of therapies such as CBT to be highly effective (Lidner, et al., 2019). There is even less research into smaller scale, symptom management for people exhibiting symptoms of depression and anxiety, which can be crucial in aiding those on long waiting lists, or who may be prevented from seeking out such interventions.

Many symptoms caused by depression and anxiety, primarily low mood, loss of motivation, fatigue, and excessive amounts of worry, can turn seeking out psychotherapy into a daunting task, and due to the often long waiting times of mental health services, it can take many months for therapies to become available for those that have signed up (Reichert & Jacobs, 2018). During this time, the ability to engage in low cost and low effort techniques for managing symptoms of depression can be pertinent to an individuals’ quality of life. Many current techniques suggest meditation and exercise, however the potential of VR to change a person’s emotional state for depression and anxiety, reducing the number of distressful feelings they may be experiencing, has been untapped.

## 1.5 - Affective State

Affective State is a measure of an individual’s emotional state. There are many theories behind how people experience emotion, and how we can classify that for the objective measurement of emotions. Virtual Reality could be used to alter an individual’s affective state by presenting virtual environments that contain characteristics that may influence emotional state and recording their subjective and objective responses to those environments through self-report methods, behavioural analysis, or electrophysiology (Holzwarth, et al., 2021). Therefore, it could be possible to place an individual within a VR environment, and assess whether their affective state measurements suggest positive emotions have been induced within them, therefore lowering distressful or unpleasant emotions, and leading to relaxation, for the purposes of low-effort, at-home symptom management using VR

This research could be extremely beneficial for many young university students who either own or are interested in owning VR headsets and are struggling with mental health issues such as depression and anxiety, or those who are in stressful situations without any mental health conditions. The project could provide insights into ways these individuals can manage their mental health better at home, while waiting for more intensive psychotherapies.

## 1.6 – Aims and Objectives

The project aims to investigate the potential of commercially available, low-effort, at-home Virtual Reality techniques to induce positive emotions in an individual by altering their affective state, in order to identify the effectiveness of VR as a tool for reducing distress in university students who may be experiencing stress, or symptoms of anxiety and depression.

The objectives of this project will be defined as:

* A methodology will be developed for assessing an individual’s affective state and levels of distress
* Data will be extracted from the implementation of this methodology as a laboratory experiment
* Tools for processing and analysing the data will be developed
* Data will be fully processed and analysed using scientific statistical analysis techniques
* Results and conclusions will be drawn from the data
* Methodology will be assessed against success criteria

## 1.7 – Intended Scope

The scope of the project will involve developing a methodology for an investigation into whether commercially available VR experiences can be used to induce positive emotions and reduce feelings of stress in an individual. This methodology will be implemented as a laboratory experiment. The results will be processed and analysed using a variety of scientific statistical methods, to arrive at a conclusion of whether this methodology provides useful results and could be investigated further. The project will not involve the development of any new VR experiences, as time constraints and other academic responsibilities limit what can be spent, and the learning of an entirely new field of technology adds an unprecedented layer of complexity that would be extremely difficult to carry out at a high level of success. Therefore, the focus will be on pre-existing technologies, and the development within this project will involve the development of a methodology, as well as any programming that will be required for the processing and analysis of the data.

# Chapter 2 – Context

## 2.1 – Introduction

This chapter will concern the background research that led to the development of the research topic and question. Relevant literature and theories surrounding the primary topics, as well as specific studies that have influenced reading and the formation of the idea will be discussed. These studies will be critically evaluated, highlighting where they may benefit the research, and how their drawbacks may be addressed in this investigation.

## 2.2 – Virtual Reality

### 2.2.1 – Background

Virtual Reality (VR) has been a rapidly developing area of technology since its inception, with many practical applications outside of video gaming (Mazuryk & Gervautz, 1996). At its core, VR is a computer-generated experience that aims to be perceived as a real, physical environment by the user. This is accomplished through interactability and replacing many of the user’s perceptions of the real world with visual and auditory stimuli. More advanced VR applications may involve tactile stimuli as well. The primary aim of VR is to make a user feel immersed in a virtual environment (Sherman & Craig, 2003).

All VR headset often compromises of a head mounted display, that contains individual screens for each eye. Many commercially available VR headsets also involve controllers, as well as positional tracking using “Lighthouses” or “Inside-Out Tracking”. Enthusiast grade VR often requires a computer to render the virtual environment and process the user’s interactions in real time. Some headsets are “all-in-one” and have bespoke processing units to eliminate the need for a PC, while a few headsets are only comprised of a frame for a mobile device to fit inside.

Virtual Reality has seen rapid growth in recent years, with many major technology companies investing in its research and development. With this growth, VR has been steadily adopted and provided many benefits in clinical situations (Riva, 2009).

### 2.2.2 - Clinical Uses

One of the most common uses of VR is surgery training. Historically, the only method of developing surgical skills for junior doctors has been under supervision within an operating room (Aim, et al., 2016). This has led to many increased costs and ethical concerns, as the number of opportunities to develop these skills may not be able to keep up with demand (Stirling, et al., 2014). Additionally, as techniques have advanced and evolved within surgery, the acquisition of many complex skills has become increasingly difficult through observation alone (Li, et al., 2017).

VR is able to address many of the concerns and difficulties of surgical training practices. VR simulations can offer more realistic experiences than conventional methods of videos and e-learning (Aim, et al., 2016). Trainee surgeons may interact with all aspects of a realistic 3D anatomical structure, in order to learn and practice surgical techniques, and performance can be easily monitored through recording of the experience, allowing for higher quality training. This also leads to reduced overall cost as patients and supervisors are no longer required, and the training can take place in any location (Yiannakopoulo, et al., 2015). Research has found that while VR simulators still require development in their accuracy and efficacy, trainee surgeons have found to complete surgeries substantially faster and are less likely to cause damage when trained through VR simulators than conventional means (Li, et al., 2017).

VR can also be used in the diagnosis and assessments of various mental conditions, such as ADHD, Parkinson’s disease, Alzheimer’s disease, and paranoia (Pandita & Won, 2020). Due to the ability to manipulate stimuli within a virtual environment, as well as tracking the reactions of patients, research has shown VR to be a powerful diagnostic tool in comparison to traditional diagnostic materials such as clinical interviews and psychometric tests (Rizzo, et al., 2004). (Freeman, et al., 2014) was able to develop an assessment method for calculating the recurrence and severity of paranoia in individuals who were susceptible.

Research into therapeutic applications of Virtual Reality have also found high levels of success. Exposure therapy is one of the most effective methods of treating a specific phobic stimulus (Riva, 2009). However, this has many practical disadvantages. In-vivo therapy where the patient is physically exposed to the stimulus are highly effective, but can be inconvenient, have a high cost, and can sacrifice controlled environments. Meanwhile In-vitro techniques where the patient observes an image or video of their stimulus can be less effective as they are not experiencing the physical stimulus but be more convenient when the phobic stimulus cannot be presented physically (Banos, et al., 2002).

The ability to offer high levels of immersion while retaining the advantages of a controlled environment through adjusting the intensity, complexity, and realism of a phobic or distressing stimulus tackles the disadvantages of In-vivo therapy, while remaining equally effective (Banos, et al., 2002). A study comparing the effectiveness of VR therapy for spider phobia found no significant difference between traditional In-vivo methods and VR therapy (Miloff, et al., 2016), suggesting that VR therapy may be used in place of in-vivo methods at more convenience and a much lower cost.

While phobia treatment is one of the most common and well researched application of VR in a therapeutic setting, benefits have been found in the treatment of Post-Traumatic Stress Disorder (PTSD) for similar reasons, such as the ability to control stimuli and place patients in specific situations (Rizzo, et al., 2009).

Research has also shown that VR based therapies are able to reduce symptoms in individuals with depression and anxiety. A meta-analysis by (Baghaei, et al., 2021) found that the use of CBT in virtual environments to be effective in supporting the treatment of anxiety and depression. This was also supported by (Ioannou, et al., 2020), who found that VR interventions can be more effective than traditional approaches. However, many of the studies within these meta-analyses do not observe the difference between tethered and non-tethered VR headsets, which could provide some insights into whether the type of headset used can influence a patients’ immersion and subsequently the effectiveness of the therapy.

### 2.2.3 – Inducing Emotional States

While many studies that research the effectiveness of virtual therapies for anxiety and depression find promising results, these studies focus on high level therapies such as CBT, while disregarding the potential lower-level applications of at-home use. Symptoms of anxiety and depression often cause an individual to be more fatigued and have less motivation to seek out and continue with therapies. Research has shown that many factors can prevent individuals from accessing CBT, such as social stigma (Parcesepe & Cabassa, 2013), high costs of treatment (Webb, et al., 2017), and high commitment (Kehel-Forbes, et al., 2016). In these cases, low-cost virtual reality could be employed to offer low effort ways of managing symptoms of depression and anxiety, such as the use of commercially available VR experiences to reduce distress and induce calm emotions.

Some research has shown promise for the ability of VR to change an individual’s emotional state. For example, (Quesnel & Riecke, 2018) found that participants who experienced certain environments in VR displayed feelings of “awe”. (Pinilla, et al., 2021) suggests that VR systems could be used for studying an individual’s emotional state through the use of affective state detection techniques, based on underlying theories of human emotion.

## 2.3 – Models of Emotion

Throughout most of the history of emotional psychology, the two most widely accepted models are the Discrete model and the Dimensional model (Barrett & Russell, 2015). Discrete models are based on the idea that specific emotions can be categorically divided, while the Dimensional models suggest that emotion lies on a spectrum of continuous variables (Bestelmeyer, et al., 2017). These major schools of emotional theories are contradictory to one another; however, they are both utilised in affective computing (Leslie, et al., 2015).

### 2.3.1 – Discrete Models

The Evolutionary Theory of Emotion, proposed by (Darwin, 1872), suggests that emotions are a direct, evolutionary response to an external stimulus. Darwin observed that emotions are universally present in both humans and animals and are expressed through muscle movement, particularly in the face such as eyebrows. (Ekman & Friesen, 1971) expanded on this theory, providing a better foundation for the classification of emotions, stating several characteristics of basic emotion. These are that all humans are born with emotions; humans exhibit the same emotions in the same situations; these emotions are expressed similarly across humans; and humans show different physiological patterns when expressing the same emotions. Based on their experiment studying facial expressions in response to stories, evidence showed that there are at least 6 facial expressions that are universal regardless of culture. These are happiness, anger, sadness, disgust, surprise, and fear.

A similar style of model known as the Ortony, Clore, and Collins (OCC) theory of emotions was later proposed by (Ortony, et al., 1988), expanding on Ekman’s model with many more distinct categories of emotion. This is the idea that 22 different types of emotions can be distinguished and clearly differentiated from each other and ordered within a hierarchy. The model suggests that the strength of an emotion directly depends on the environment that an individual is within, and the objects and events within that environment. The first stage in in eliciting an emotional response is perception of the stimuli. The second stage is appraisal, where the environment and the situation are evaluated. Finally, the emotional response is elicited (Francis Jr, et al., 2009). Due to the categorical nature of this model, it may be compatible with many modern-day algorithms that aim to classify and categorise emotions. However, the theory places the primary factor of an emotional response on the environmental factors and does not discuss any physiological factors.

Similar to this theory, (Plutchik, 1982) proposed a wheel of emotions that divides them into 3 levels: primary, secondary, and tertiary. 8 basic emotional states are defined as ecstasy, adoration, terror, amazement, grief, loathing, rage, and vigilance. Any emotional state can be described as a combination of any of the 8 basic emotions. As with the aforementioned theories, it suggests that an emotional response is a reaction to a stimulus in the environment. The stimulus is appraised, the subjective experience of the emotion is experienced, and neural activity elicits behavioural responses.

A number of other major theories of emotion are the James-Lange Theory (James & Lange, 1922), the Cannon-Bard theory (Cannon, 1927), and the Schachter-Singer theory (Schachter & Singer, 1962). These theories are similar, in that they all state emotional response has 4 distinct components: appraisal, feeling, physiological change, and behaviour. All theories, however, suggest different orders of these components taking place.

(James & Lange, 1922) propose that emotional response is triggered due to a perceived internal physiological change within the human body, in response to external stimuli. This is in contrast with several other theories, such as the evolutionary model and the OCC, which place less emphasis on the physiological aspects of emotional response. The feelings of an emotional state occur after the behavioural and physiological responses, which occur after the appraisal of the situation. The Cannon-Bard theory opposes this, in favour of the idea that all components of an emotional response occur independently of each other, with no distinct order of operation (Cannon, 1927). However, much evidence has been placed against this theory, suggesting that the components of an emotional response do depend on each other (Dimberg, et al., 2000). Finally, the Schachter and Singer theory suggests that physiological responses occur first, and then the individual attempts to find a reason for those changes within their environment (Schachter & Singer, 1962). Depending on the reason identified, a cognitive label is given to these feelings, which measures the intensity of that emotion.

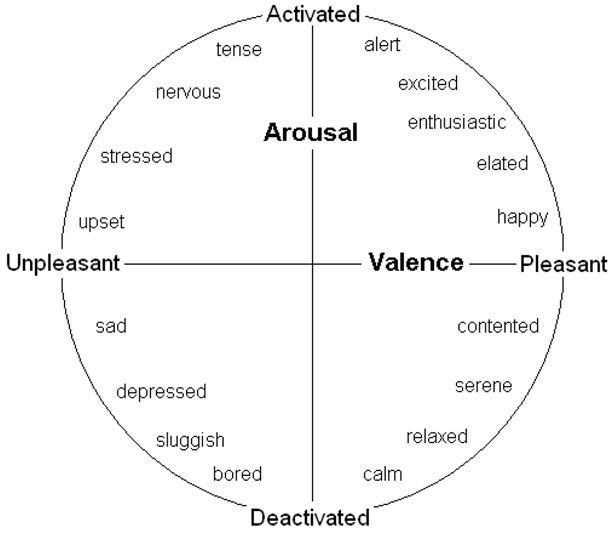
A picture containing pinwheel, outdoor object, vector graphics

Description automatically generatedDiscrete models of emotion have many uses in affective state research. Due to their categorical nature, they are compatible with many systems and algorithms that aim to classify emotion, as well as allowing for self-report methods of affective research. However, much evidence can contradict discrete theories of emotion. (Cacioppo, et al., 2000) argues that basic emotional states have not been found to be associated with specific facial expressions, as suggested by Ekman. Additionally, it has been suggested that many people find difficulty in assessing, discerning, and describing their emotional states (Saarni, 1999). This suggests that many individuals do not interpret emotions as individual, discrete events. This issue is addressed by the dimensional models of emotion.

**Figure 1: Plutchik's Wheel of Emotions**

### 2.3.2 – Dimensional Models

Dimensional models of emotion originated from psychologist Wilhelm Wundt, who proposed that emotional responses have three dimensions (Wundt, 1897). These are valence, arousal, and intensity. Valence is the range of whether an emotion is pleasant or unpleasant. Arousal is the range of whether an emotion is arousing or subduing. Intensity is whether an emotion is strenuous or relaxing.

Later research found that the arousal and intensity dimensions are overlapping with each other, and a 2-dimensional model may be sufficient to capture emotional states (Ekman, 1957). This led to the development of the Circumplex Model of Emotion by (Russel, 1980), who proposed that all emotions can be represented on a circle, with valence and arousal placed on an axis acting as two separate dimensions. The horizontal axis representing valence, and the vertical axis representing arousal. This is one of the most used theories when researching affective state (Posner, et al., 2005).

**Figure 2: Circumplex Model of Emotion**

Other researchers have proposed the Evaluative Space Model (Cacioppo, et al., 1997), which attempts to allow for the assessment of emotional responses that may have simultaneous negative and positive connotation, such as “bitter-sweet”. Negativity is represented as the horizontal axis, and positivity is represented as the vertical axis. A z-axis is used to measure aversion towards a stimulus.

While dimensional models of emotion offer a more ambiguous, subjective approach than discrete models, they may offer more use for assessing an individual’s emotional state through the use of physiology, as emotion can be studied as continuous variables in relation to the change in processes within the human body.

## 2.4 – Estimating Affective States

Following the discussion on the different theories of emotion, techniques to measure an individual’s emotional state can be assessed. There are three primary categories of techniques that are commonly used for estimating an individual’s affective state. These are self-reports, behavioural measures, and electrophysiology.

### 2.4.1 – Self Reports

Self-reports allow users to subjectively rate their own experience of emotion through answering a series of questions. These can be either open ended, asking for a descriptive experience, or closed, utilising specific answers. Self-report methods are relatively easy to use as they only require the questions to be displayed on a piece of paper or screen, and a method of response. As they are asking users directly about their own personal experience with their mental state, they can be a direct measure, unlike the other techniques (Tang & Tang, 2020). An issue with this, however, is that they can be susceptible to experimenter bias, as participants may adjust their responses to be in favour of supporting the research question (Rosenman, et al., 2011).

One of the most commonly used self-report method for discrete affective state measurement is the Pick a Mood (PAM) (Desmet, et al., 2016). The PAM uses pictures to assess an individual’s affective state. There are nine possible moods, and three characters for each mood. Participants can choose which moods they feel most strongly with.

The most used method for dimensional affective state measurement is the Self-Assessment Manekin (SAM) (Bradley & Lang, 1994). Similar to the PAM, the SAM uses pictures to assess affective state. It employs three scales: valence, arousal, and dominance. Each scale has five pictures. Participants can select which picture they feel most strongly on each of the three scales or use the blank space in between the pictures as an intermediary state. The SAM is one of the most well-known and established tools for assessing affective state.

### 2.4.2 – Behavioural Measures

Behavioural measures aim to assess affective states through empirical, observable means. This includes facial expressions, as previously mentioned, as well as vocal patterns and body movements.

It has been recently suggested that arousal can be interpreted from an individual’s body movements; faster movements can be associated with higher levels of arousal in an individual. An experiment found that negative emotions such as sadness can be associated with lowering the head, while boredom can be associated with leaning the head on an arm, through playing a series of recordings with various emotional content (Bull, 1978)

Classification algorithms can be used to classify a user’s emotional state based on their vocal patterns, using features such as volume, tempo, and pitch. (Vogt, et al., 2008). An individual who is talking at a faster and disjointed tempo, for example, may be more likely to be experiencing a stressful emotion.

Much research has been conducted on the association between facial expressions and affective state, as has been mentioned in the previous section. These facial expressions can be visually analysed, and an emotional state can be interpreted from them using the Facial Action Coding System (FACS) (Ekman & Friesen, 1971). This is a system that outlines possible movements of facial muscles and describes a facial expression as the combination of several movements.

Many behavioural measures may become difficult to accurately assess while using a VR headset. The head mounted display may get in the way of natural body movements and inhibit the analysis of facial expressions. Additionally, users may find it awkward or unnatural to speak during an experiment for vocal analysis.

### 2.4.3 - Electrophysiology

The final method that will be discussed is electrophysiology. This is the measurement of the electrical properties of different cells within the human body. The primary methods of electrophysiology are facial electromyography (fEMG) for facial muscle activity, electrocardiography (ECG) for heart activity, and electroencephalography (EEG) for brain activity.

Research has shown that ECGs are capable of assessing an individual’s arousal, the intensity of their emotional state. The intervals between each heartbeat can be used to calculate heart rate variability (HRV). Higher HRV is commonly associated with higher levels of arousal (Thayer, et al., 2009).

Additionally, valence can be calculated from fEMG and EEG signals. (Dimberg, 1982) argues that muscle activity above the eyebrows is associated with negative valence, and therefore negative emotional states. Additionally, muscle activity in the cheeks is associated with positive emotional states. This, however, has similar issues to the implementation of behavioural measures of facial expressions. The head mounted display can interfere with the instruments required for accurate fEMG.

EEG can measure the frequencies of electrical activity within the brain. The most studied frequencies in EEG based research are Delta (0Hz – 4Hz), Theta (4Hz – 8Hz), Alpha (8Hz – 12Hz), Beta (12Hz – 35Hz), and Gamma (above 35Hz). Delta frequencies in the brain are most commonly associated with individuals in a state of deep sleep. The theta band is most associated with awake individuals in a state of deep relaxation and drowsiness, as well as meditative concentration. The alpha band is primarily associated individuals who are awake and alert, in a relaxed mental state. The Beta band is associated with focused mental activity, and finally the Gamma band is associated with intense mental activity and concentration (Abhang, et al., 2016).

Studies have shown that states of relaxation and pleasant emotions are associated with increased levels of alpha power (Ulrich, 1981). Additionally, (Price & Budzynski, 2009) found that increased Beta activity and decreased Alpha activity have shown to be most associated with stress and anxiety which has been backed up by (Cahn & Polich, 2006) who concluded that increases in alpha power are associated with increased calmness and positive affective states, as the slower alpha frequencies inhibit faster frequencies in the brain, leading to reduced mental activity.

EEG may have many issues with VR based stimuli, as eye movements, blinking, and muscle movements can cause artifacts in the EEG signal (Klug & Gramann, 2020). As VR applications involve movement for immersion, and the processing of visual stimuli, it is difficult to reduce the number of artifacts created in the signal. It may be possible to remove these artifacts, however, using Independent Component Analysis (ICA) (Makeig, et al., 1997). This is an algorithm that separates out components of an EEG signal that may be caused by artifacts. Each of these components can therefore be inspected and removed.

## 2.5 – Similar Research

### 2.5.1 – VR for Affective Research

(Kumar, 2021) employs the use of EEG signals to detect the emotion of individuals after viewing several videos within a VR headset. The videos are selected to induce different combinations of valence and arousal in the participants, and the EEG results are correlated with subjective reports on the SAM. The purpose of this study is to develop machine learning based tools for the classification of affective states. This study provides many insights into the process of using EEG signals in conjunction with VR experiences to measure affective state. For example, the study details techniques for the acquisition of EEG signals and the scientific methods of processing them and extracting relevant information. Similarly, (Yu, et al., 2022) also measures valence in response to VR environments for emotion recognition algorithms, for the purposes of developing a comprehensive dataset of 3D VR videos for the purposes of affective research.

These studies have been a major influence in my research, as it concerns very similar theories and ideas to the aim of this research. However, they are limited in their influence, as the aim of these studies is to predict a variety of emotions, as opposed to inducing a specific emotional state. Despite this, the research is valuable in demonstrating that virtual reality can be used to measure affective state, and specifically valence. Using the outlined principles, one could study the relationships between specific affective states and reduced levels of distress.

Finally, (Quesnel & Riecke, 2018) attempt to measure the subjective experience of “Awe” through Virtual Reality and physiological response. Participants were shown several interactive VR environments, and the emotional state of awe was measured through a correlation of goose bumps and self-reported responses to a questionnaire. They found a positive correlation between the physiological response of goose bumps, and self-reported values. This suggests that specific emotions could be studied using a combination of physiological and self-report measuring tools. However, the use of goose bumps as a physiological response to an emotion may lack validity, as other variables such as differences in temperature may also cause goose bumps to appear.

### 2.5.2 – VR for Stress Reduction

(Kim, et al., 2021) studies the effects of VR on stress reduction through first inducing participants into a stressful state of discomfort through VR videos with a large amount of unnatural movement. Following this, participants were exposed to relaxing VR environments. Participants levels of stress were measured through HRV. It was found that VR relaxation techniques were effective in reducing levels of stress. This study shows that it is possible to use VR to reduce an individual’s physiological and perceived levels of stress. However, it manually induces participants into a state of stress through a potentially nauseating experience, which can lead to ethical issues. Additionally, it does not use HRV to measure affective states, therefore conclusions about a participant’s specific emotional state cannot be made. Finally, it utilises the Simulator Sickness Questionnaire for self-reporting subjective levels of distress. This may have issues as the instrument measures distress because of motion sickness (Sevinc & Ilker, 2020). A different self-report tool may need to be selected to measure distress from depressive and anxious symptoms.

(Bjorling, et al., 2022) measured the effects of commercially available VR games, such as Nature Treks VR, on the stress levels of a sample of students from universities and high schools. They employed the Perceived Stress Scale to measure levels of perceived stress over the course of the previous month. The study was conducted in several sessions in order to assess the long-term effects of VR interventions on stress. It was found that VR is highly effective at reducing perceived levels of stress within an individual and provides many insights into the possible tools and technologies that could be employed within this investigation, such as the PSS and Nature Treks VR. However, this study suffers from the issue of only using self-report methods to assess stress, and no objective measurements are taken in conjunction with self-report to validate the accuracy of the results. Additionally, the study places no focus on the possibility of studying the affective state of the participants.

# Chapter 3 – New Ideas

## 3.1 – Introduction

Following the discussion on similar ideas in Chapter 2, it was identified that while many studies that research affective state as a response to VR through EEG exist, there are none that specifically study the use of this research in techniques for symptom management in individuals who may suffer from symptoms of depression and anxiety. Furthermore, while meta-analyses have shown promising results in studies researching the effectiveness of VR as a means of CBT delivery for individuals suffering with symptoms of depression and anxiety, there is lacking research in the potential of VR as an ”at-home” symptom management tool for individuals who may be on a waiting list for CBT, find the idea of therapy too daunting, or do not exhibit enough symptoms to be given a diagnosis but feel as though stress management tools may assist them.

Therefore, this section will focus on further developing the idea of VR as a method of low effort, at-home symptom management through studying affective states. The wider theories discussed in the previous chapter will be analysed for specific ideas to be focused on to narrow the field of research and provide a clear, reproducible, and justified methodology.

## 3.2 – Planning the Methodology Development

Considering the theories and ideas discussed in Chapter 2, as well as the overall aims and objectives of the proposed project, several aspects of the development of the methodology will need to be planned and discussed. These will include tools for assessing the participants subjective level of distress, the underlying theory that will be used for assessing affective state as well as the tools that will be used to gather the data to assess the affective state. Additionally, various possible methodologies of the experimental design will be contrasted, as well as any other apparatus required.

### 3.2.1 – Evaluating Stress

There are a variety of methods for evaluating the severity of symptoms in an individual with depression or anxiety. The most common type of method are self-report questionnaires. While self-report is susceptible to experimenter bias, it has been chosen over physiological measurements as it is much easier and quicker to analyse, therefore leaving more room for thorough affective state analysis. Additionally, while physiological measures may have higher objectivity, the subjective nature of self-report may be advantageous within this research, as the aim is to assess whether subjective feelings of distress are reduced.

The most commonly used self-report questionnaires for anxiety and depression are the GAD-7 and the PHQ-9, respectively (Teymoori, et al., 2020). They both display multiple questions about several different symptoms of the condition and allow the user to circle a response between 0 and 3. The response is a measure of how frequently the individual has been experiencing those issues over the previous 2 weeks, with a higher number being a higher frequency. The sum of the circled numbers describes the severity of the condition. While these are widely used instruments for the diagnosis of depression and anxiety, they would not prove to be useful within this investigation. This is because the aim is to measure a direct change before and after the VR intervention, to assess whether the intervention influenced their levels of distress. These questionnaires are designed to measure change over a longer period, for the analysis of effectiveness of courses of therapy.

Another possible self-report tool is the Subjective Units of Distress Scale (SUDS) (Tanner, 2012). This involves a scale from 0-100 used to measure the user’s subjective, perceived intensity of feelings of distress, and is used in many cognitive behavioural therapies. The primary advantage of this scale is that it emphasises recording the intensity of distress as it is felt in the moment, which is more useful than the PHQ-9 or the GAD-7. The scale can be used to measure a participant’s level of distress directly before and after the intervention.

### 3.2.2 – Emotion Models

As previously mentioned in Chapter 2, most models of emotion fall within two categories: Discrete and Dimensional. Discrete models focus on assigning labels to specific emotions, while Dimensional models treat emotion as continuous variables. The primary advantage of discrete models is that they are compatible with modern day classification algorithms, as well as allowing for individuals to self-report specific emotional states that they are in. The advantages of dimensional models are that they may allow for a better analysis of an individual’s change in physiology.

Research has shown that both models of emotion have uses depending on their situation, and one is not objectively better than the other (Harmon-Jones, et al., 2017). For this study, however, the dimensional models of emotion will be used. This is because the investigation aims to lower an individual’s level of distress by inducing pleasant emotions as opposed to capturing a specific affective state. Therefore, the ability to measure change in a continuous variable is more suited to this experiment.

The main dimensional models of emotion that have been discussed in Chapter 2 are the Circumplex Model of Emotion and the Evaluative Space Model. The Circumplex Model measures affective state in 2 dimensions: valence and arousal. The Evaluative Space Model measures affective state in 3 dimensions of positivity, negativity, and aversion. The Circumplex Model may be more useful for this study as valence is a useful measure of whether an individual is in an unpleasant or pleasant affective state. The level of combined positivity, negativity, and aversion in the Evaluative Space Model may become overcomplicated for the nature of the experiment, and it would be more focused to observe a single axis. Additionally, the Circumplex Model is highly researched, and appears to be used in many of the similar investigations that have been discussed in Chapter 2.

### 3.2.3 - Evaluating Affective State

A variety of methods of capturing an individual’s affective state were discussed in Chapter 2. These will be contrasted to identify which method would be most useful within this study. The 3 primary areas outlined were self-reports, behavioural measures, and electrophysiology.

The primary method of self-report that could be utilised is the Self-Assessment Manekin, as the Pick a Mood questionnaire focuses on discrete models of emotion. This can capture an individual’s valence, arousal, and dominance through a series of pictures. This would be a useful method of assessing the user’s level of valence, however it may be susceptible to bias from demand characteristics (Orne, 1996), as individual’s may adjust their responses to align better with the expected findings of the study. As the method for assessing the individual’s subjective sense of stress is also a self-report method, this may lead to additional bias. Rating lower sense of stress may lead them to rate a higher level of valence, or vice versa.

Many behavioural measures such as facial expressions and body movements may not be useful for a VR based intervention. The head mounted display will block a large portion of the face, causing a reduction of data and meaning facial expressions may not be accurately read. Additionally, natural facial expressions may not be made by the participant while wearing the head mounted display. Body movements may also be unnatural, due to the controllers of the VR headset. The participant will be holding controllers and moving to interact with items, therefore natural body movements may not occur, leading to inaccurate results. Vocal pattern analysation is the most viable option for behavioural measurement. However, speaking in an experimental setting may be awkward and unnatural for the user, leading to inaccurate results. Finally, many behavioural measure techniques are primarily used for discrete models of emotion, and therefore would not be useful in this experiment.

Electrophysiological measures include fEMG, ECG, and EEG. fEMG suffers from the same issues as facial expression analysis. The head mounted display will block a portion of the face and therefore lead to reduced and inaccurate results. ECG and EEG are both viable options for measuring affective state, however they differ in which variables they are capable of measuring. Research has shown that ECG is more useful for measuring arousal through heart rate variability, while EEG is more useful for measuring valence through alpha frequency band power.

The two primary options that remain are the SAM or EEG. The self-report nature of the SAM may lead to inaccurate results because of bias, while the EEG may produce artifacts in its data collection. These artifacts, however, can be removed through data processing. The EEG may also lead to more accurate results, as it is not susceptible to bias.

### 3.2.4 – Population of Interest

The population that the experiment is interested in testing should be decided first. This the demographics of people that a sample will be selected from, to assess whether the results of the experiment are useful to a subset of people and can therefore be generalised to the wider population. The aim of the experiment is to assess whether low-intensity, at-home VR environments can lead to a positive affective state and a reduction of distress in an individual.

Research suggests that young adult (18-24 years old) students are one of the groups most at risk of symptoms of anxiety and depression (Mahmoud, et al., 2012). According to the CDC, the 18-29 age group was the highest among respondents in a national health survey who suffered from depressive symptoms (Villaroel & Terlizzi, 2020). Additionally, the research focuses on at-home management using a VR device, therefore would prove most beneficial to those who already own or are interested in owning a VR device. One study found that most people who own a VR headset are 18-29 years old (Mottelson, et al., 2021).

Therefore, it was decided that the target demographic would be university students within the 18-25 age range, who either own or are interested in VR, and suggest they are under any level of distress and therefore may benefit from VR intervention. Other demographics such as gender will not be considered, as this would not provide any useful information for the study.

### 3.2.5 – Hypothesis

One of the primary stages in the scientific method is the designing of the hypothesis. This is the theory that will be tested in the experiment, and should state the expected, operationalised dependant variables, how they will change in response to the independent variables, and the population of interest that this change is expected to occur in. Once the hypothesis is developed, it is split into two parts for statistical analysis: the null hypothesis and the alternative hypothesis. The null hypothesis suggests that the original hypothesis is not true, while the alternative hypothesis is the inverse of the null.

Therefore, the hypothesis should state the variables of ratings on the SUDS scale, and alpha frequency band power. It should state how these variables are expected to change in relation to the VR intervention, and the population of interest that the change should occur in.

### 3.2.6 – Experimental Design

Experimental design is the method in which groups will be allocated for the experiment. Some experiments will require control groups and experimental groups, while other may only need a single group. For example, an independent measures design will have an experimental group that performs the experimental task, and a control group to do a base task, to identify whether a change occurs between groups. Meanwhile a single-armed experiment will simply contain a single condition that all participants perform, usually comparing data from before and after the task. This is often referred to as a repeated measures design (Massaro, 2009).

A single armed experiment in this scenario would involve collecting baseline data from participants at the start, then placing them within the VR intervention, and then collecting data after, to measure whether a change occurred within the sample directly due to the VR intervention. Meanwhile, using multiple conditions would involve the experimental condition being placed in the VR intervention, and then having data collected. A control group would be placed in a basic, neutral VR experience. The data would then be compared between the two conditions, to assess if a difference exists between the samples.

While a single-armed experiment may incur bias from demand characteristics, as the participants will have a clearer idea of what response is expected from them, it is more feasible in situations with fewer participants, and the experiment will be easier to conduct and control. A single-armed experimental design may suite this experiment better, as measuring change within the same sample before and after the VR intervention will provide more accuracy on whether the change is due to the intervention itself, as opposed to the type of VR experience. Additionally, it reduces risk of differences between participants leading to extraneous variables that may be difficult to control.

### 3.2.7 – Apparatus

#### 3.2.7.1 - Virtual Reality Experience

As the research aims to focus on at-home VR based interventions, the possible experiences to use will be limited to commercially available, free, or low-cost applications that are aimed at towards the public, and available on major digital storefronts such as the Oculus store or Steam. The application must also be aimed at reducing stress. Additionally, studies have shown natural landscapes to be effective for stress reduction, relaxation, and mindfulness. Taking this into account, the potential VR applications will be limited to those focused on natural landscapes. The most popular VR games that fit the outlined criteria are “Nature Treks VR” and “TheBlu” (Evrim, 2021).

Nature Treks VR is a nature-based VR game that offers many different natural environments for the player to choose from, that are themed around colour and designed to induce specific emotional states. There is no objective for the game, instead the user is encouraged to explore and interact with these environments, and tools are offered such as the ability to grow plants to promote this. Each environment contains simulated wildlife and the ability to control the time of day and weather.

TheBlu is an immersive VR application that allows users to experience different habitats and wildlife of the ocean. It emphasises stress reduction through providing “awe inspiring” experiences, such as contact with a whale. Like Nature Treks VR, it a simulated, 3D rendered environment. However, it does not offer any interactability. The intent is for the user to remain in one location and simply observe the environment. Additionally, research has shown that some individuals may find the idea of being submerged underwater uncomfortable, nauseating, and distressing.

#### 3.2.7.2 - Virtual Reality Headset

Similar to the criteria for the VR experience, the VR headset should be chosen in accordance with the desired aim of the experiment. As an “at-home” solution is the goal, a low-cost, commercially available VR headset should be chosen to conduct the experiment. The headset will also need to be capable of running the Nature Treks VR software, so mobile VR headsets are not an option.

The most common commercially available VR headsets are the Oculus Quest 2, the Valve Index, and the HTC Vive, according to the Steam Hardware Survey (Valve, 2022). Out of these, the Oculus Quest 2 is owned by most users using the Steam platform, compromising of 48% of the VR headsets in responses. The Oculus Quest 2 is also the cheapest headset out of the three. This is due to both the lowest upfront cost of the headset itself, as well as the headset being an all-in-one unit, therefore not requiring a base-stations for tracking or a PC with hardware specifications that meet the minimum requirements for running the Nature Treks VR software. This also means that the Oculus Quest 2 does not require to be “tethered” to a PC and can be used wirelessly, which improves the overall immersion of the headset.

#### 3.2.7.3 - EEG Headset

A variety of different low-cost EEG headsets are available for the collection and processing of research-grade data for frequency power analysis. Popular EEG headsets include the InterAxon Muse headband, the OpenBCI Ultracortex IV, the Neurosky Mindwave, and the Emotiv Insight.

The Muse is a compact headband shaped EEG device, offering 5 sensors. It uses Bluetooth to send data to a mobile application. The headband uses dry electrodes at the FPz, AF7, AF8, TP9 and TP10 locations based on the international 10-20 system. The MindWave is the lowest cost headset out of the selection, but only offers a single channel dry electrode positioned in the middle of the forehead. Bluetooth is also used to transmit data to a bespoke mobile application. The OpenBCI is an open-source headset that can be 3D printed for use with an OpenBCI board such as the Cyton. It offers up to 16 channels depending on the board used and can transmit data to a desktop application through a dongle or over Wi-Fi. Finally, the Insight is another 5-channel headset that uses Bluetooth to transmit data to a smartphone application.

The compatibility of the EEG device with a VR headset should be considered first. Due to the “headphone” style design of the MindWave and the Insight, it would be difficult and impractical to fit them over the head mounted display or the head strap of a VR headset. Due to the fact that the Muse has a headband design, it would easily fit underneath a VR headset. The OpenBCI could also be fitted over a headset and offers additional electrodes for more accurate data recordings.

Due to being the most easily accessible EEG headset for data collection, the OpenBCI Ultracortex frame and Cyton board will be used to collect EEG data. The Cyton board allows for 8 electrode channels. These electrodes will be arranged in the standard international 10-20 system.

## 3.3 – Evaluation

### 3.3.1 – Correlation

A correlation test is a statistical tool to generate a descriptive statistic of the strength and direction of a correlation between two variables. The experiment aims to measure whether a correlation exists between affective state and subjective levels of distress. Both variables are likely to be quantitative and will result in interval data. Therefore, the Pearson’s R correlation test should be used to determine the strength and direction of a correlation. This is a parametric test, meaning that it is used if the data follows a normal distribution. In the case that the data does not fit the requirements for a parametric test, the Spearman’s Rho test is a non-parametric equivalent.

Both tests produce a normalised covariance value between -1 and 1. A positive value suggests a positive correlation, while a negative value suggests a negative correlation. The closeness of the number to 0 describes the strength of the correlation. This value will be useful in assessing whether the research is studying the correct variables, and if it has merit to be studied further.

### 3.3.2 – Statistical Test

After calculating the correlation coefficient, a statistical test can be applied to calculate the level of significance, referred to as the p value. This value described the likelihood of the null hypothesis being correct, and therefore whether it should be accepted or rejected. The p value is compared against an alpha value, which is the threshold of whether an experiment can be considered statistically significant. A smaller alpha value reduces the risk of a type 1 error, where the null hypothesis may be falsely rejected, however it also creates difficulty in achieving statistical significance. The most common alpha value used in 0.05, and this will be used within the experiment.

The most common statistical tests to employ when using a Pearson’s R or Spearman’s Rho are the t-test or the z-test, using Fisher’s method for transforming the distribution towards a normal distribution, in the case that it is not symmetrical. For this experiment, a t-test would be better suited as it is not expected to achieve a large sample size. As the experiment will have a single-arm, a paired t-test is mots applicable. Additionally, as there is an expected difference between the results, a one-tailed t-test would be applicable.

For the purposes of evaluating statistical significance of the data collected directly before and after the VR intervention, the paired t-test, or its non-parametric equivalent, the Wilcoxon Signed-Rank test can be applied. These are paired tests used for a repeated measures design, when data is collected from the same participants in both conditions.

### 3.3.3 – Success Criteria

Finally, a method of assessing the effectiveness of the methodology should also be designed. Using this, it would be possible to identify weaknesses in the methodology for future studies.

One possible method of evaluating the performance of the experiment is through interviewing the participants after they participate. This will give a subjective measure of the participants’ experiences while undergoing the experiment, which aspects of it they enjoyed, any difficulties they faced, and whether they felt as if the experiment had a positive effect on them.

Another method of evaluating the success of the experiment would be through analysing the results through descriptive statistics and statistical analyses, as described above. This will provide an objective measure of whether the experiments’ results are able to support the success of the experiment.

Using this outline, a success criteria can be defined. This contains a series of statements, which can be either true or false. The higher the number of criteria that have been met, the stronger and more successful the experiment. If a low number of criteria are met, then the experiment is considered weaker and unsuccessful.

* A majority of participants state that the experiment was overall, a positive experience
* A majority of participants state that the experiment helped them in inducing pleasant emotions and reducing distress
* No participants had any major difficulties that disrupted the experiment
* No participants expressed major discomfort during the experiment
* The results suggest a positive correlation between the variables
* The results suggest a strong positive correlation between the variables
* The results suggest statistical significance of the experiment

# Chapter 4 – Investigation

## 4.1 - Introduction

This chapter provides the details on the investigation of the project. This involves how the methodology was planned, developed, and implemented. The main aspects of this were the design process of the experiment, how it was conducted, and how the data was processed and analysed with the development of analytical scripts in MATLAB and R.

Many aspects of the experimental design have been detailed in Chapter 3, discussing the advantages and disadvantages of alternate solutions, before arriving at a decision. This Chapter will explain the chosen concepts and apparatus in detail and explain how they will interact with each other to form a complete design.

## 4.2 – Timescale Planning

### 4.2.1 – Stages

The investigation of this project will be split up into several stages. First, the development of the methodology will be outlined, following the discussions in chapters 2 and 3. Each aspect of the methodology will be contrasted with alternatives, and justification will be provided on the final choices. The next stage will be the gathering of participants following a specific sampling protocol. After this, the experiment will be conducted, and finally, the data will be processed and analysed for evaluation.

### 4.2.2 – Timescale

The development of the experimental design is estimated to take approximately four weeks. This will involve the identification of dependant variables, creation of a hypothesis, an investigation of participant gathering techniques, a review of possible apparatus and the acquisition of selected apparatus, and the designing of the experimental task. In addition to this, the data processing stages will be researched and planned, as well as the methods of analysing the data through descriptive and inferential statistics.

After the experiment has been fully designed and finalised, the participants for the experiment will be gathered. This stage will take approximately 1 week and will involve the designing of the materials used to gather and inform participants of the experiment, and the scheduling of dates for experiments to take place. Outside of this, the laboratory will be set up and the apparatus will be fully tested to ensure all aspects are compatible with each other.

Once the participants have been gathered, the experiments will take place over a course of 2 weeks. During this time, all participants will take part in the experiment, and data will be gathered. In addition to this, the technologies for the data processing and analysis will be tested, and familiarity will be gained with them

The final stage will concern the processing and analysis of the data and will take an estimated 3 weeks. In this stage, all data will be processed using the outlined tools and methodologies, including the development of any resources required. Additionally, tools for evaluating the data will be developed using a selected programming language. Finally, the processed data will be tested to evaluate the performance of the investigation.

## 4.3 – Methodology Development

### 4.3.1 – Dependant Variables

The possible methods of assessing an individual’s affective state were discussed in the previous chapter, and the results were that the Self-Assessment Manekin (SAM) and EEG were the most viable methods to assess levels of valence. The SAM is a self-report tool, and therefore may be susceptible to subjectivity and bias from the participant. EEG is a more invasive tool that is more difficult to utilise, however may receive more accurate, objective results that are not subject to bias. Therefore, EEG will be used to evaluate affective state. Valence will be assessed through studying the change in alpha frequency band power, as discussed in Chapter 2. The primary method of calculating band power is through Power Spectral Density Analysis, which is the average power of a specific frequency band within the brain.

Various self-report methods of assessing an individual's stress levels were also discussed in Chapter 3. Self-report was chosen as the medium over physiology to gain a wider variety of types of data, in addition to the subjective nature of self-report being advantageous for the aims of this experiment. The idea that a participant subjectively feels as though their stress has been reduced is important. A comparison of these tools showed that many assess symptoms over a course of time, while the SUDS scale measures feelings of distress in the moment. This is more useful for this experiment as feelings of distress should be captured as a direct response to the VR intervention.

### 4.3.2 – Hypothesis

As discussed in Chapter 3, the hypothesis is an integral aspect of the development of an experimental methodology. It should be clear, concise, and describe exactly what is expected of the experiment in order for an accurate scientific, and reproducible methodology to be designed. Based on prior research, the dependant variables that are being measured will be the alpha frequency band power for affective state, and ratings on the SUDS scale for feelings of distress. The population of interest, as discussed in chapter 3, will be university students between 18-25 years of age.

Therefore, the hypothesis for this experiment is that university students aged 18-25 will report lower levels of stress on the SUDS scale of measurement and have higher alpha frequency band power after experiencing VR based intervention than before. The null hypothesis is that there will be no difference in self-reported stress levels on the SUDS scale of measurement or in alpha frequency band power after experiencing VR based intervention than before.

### 4.3.3 – Participant Gathering

Sampling techniques are used to gather a selection of the population of interest for taking part in the experiment. The results can then be generalised to the rest of the population. The population of interest is 18–25-year-old university students, as this group may be most at risk of symptoms of depression and anxiety, as well as being the most likely to own or be interested in VR headsets.

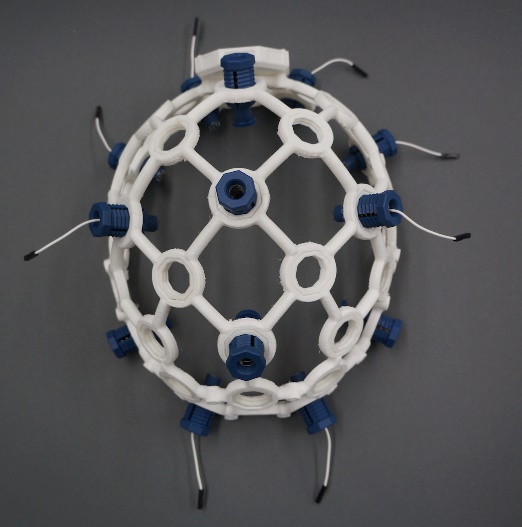
Sampling techniques to gather a specific group of university students could involve volunteer sampling, through advertising in relevant online spaces or around a campus. Another technique would be opportunity sampling, through asking people on a campus at a location close to the laboratory to take part in a study at that moment. However, this may lead to an unrepresentative sample as recruiting people from the same location may mean they all have common characteristics. Volunteer sampling avoids this issue as many different spaces could be advertised to. Additionally, volunteer sampling may lead to individuals who are interested in VR and who may be suffering from stress responding to the advertisement, which is the population that the study is interested in.

### 4.3.4 – Apparatus

Following the review of potential VR headsets for use within the experiment in Chapter 3, the Oculus Quest 2 was chosen to be used within this study. This is because in comparison to alternatives, it was the cheapest and most widely available headset, fitting with the requirements that the headset should be targeted towards the average consumer, for a low effort at home VR experience. Additionally, it can run most VR software internally, reducing cost and improving immersion, which could be beneficial for the results of the study. The headset would be running at its default, native resolution of 1832 x 1920 per eye, and a default refresh rate of 90Hz.

**Figure 3: Oculus Quest 2**

Diagram, venn diagram

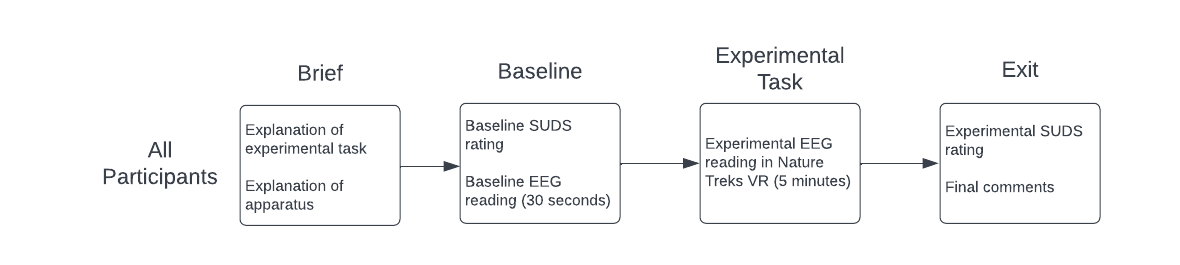
Description automatically generatedFollowing the review of low-cost EEG headsets, the OpenBCI UltraCortex IV with the Cyton board was chosen to collect the EEG data. This is due to having an advantage of more available electrode channels than alternative headsets, as well as a lower cost due to being open source and the ability to 3D print the frame, and increased customisability and options. The EEG data would be collected through the OpenBCI GUI interface, using the dongle rather than over Wi-Fi. The dongle was decided to be used as this would avoid any potential interruptions to the experiment due to situations such as the Wi-Fi signal within the laboratory going down or being too weak. The UltraCortex was tested to fit over the Oculus Quest 2, and it was able to do so without any of the primary, data collection electrodes being blocked. Eight electrodes were arranged in the international 10-20 system for reproducibility, and the sampling rate of the recording was set to 250Hz. This is to satisfy the Nyquist frequency, where the sampling rate should be at least twice as high as the highest frequency being measured (Wiergraber, et al., 2016).

**Figure 4: International 10-20 System of Electrode Placement with 8 Electrodes**

**Figure 5: OpenBCI UltraCortx IV**

The two primary nature-based VR applications focused on relaxation that were identified were Nature Treks VR and TheBlu. Out of these, Nature Treks VR was chosen to be the VR based intervention that the participants would experience during the experiment. This is because it offers a higher amount of interactability and choices, as well being easily available on the Oculus store for the Oculus Quest 2, unlike TheBlu, which was only available on the Steam Store. Additionally, some individuals may experience discomfort from underwater based environments. Nature Treks VR offers several different locations for participants to experience. However, to standardise the experiment between participants, a single environment should be chosen for all participants. This is to avoid the type of experience becoming a confounding variable that may affect the results. The Green Meadows environment was chosen, as it offers the most nature, which has shown to be effective for relaxation, and does not contain any potentially distressing stimuli.

### 4.3.5 – Experimental Task

Following the selection of all the dependant variables to be measured, the designing of the hypothesis, the acquisition of participants, and the selection of apparatus, the experimental procedure can be developed based on an experimental design. It was identified in Chapter 3 that a single-armed experiment may provide more accurate and focused results for this research, as the study should focus on changes within a sample, rather than the differences between two samples, which may create extraneous variables. The experimental procedure should state exactly what the experiment will entail, including details on what is expected of the participants, how the apparatus should be used, and how the data will be collected. Exact values should be used when discussing measurements. This will improve the reliability and the reproducibility of the study. Using this design, an experimental procedure can be developed.

**Figure 6: Diagram of Experimental Task**

At the start of the experiment, the participants will each answer on the SUDS scale. Then a baseline, control EEG recording will be taken with the participants seated and still, to avoid artifacts. The VR headset will then be fitted, and participants should be briefed on the controls of the VR experience and asked to remain seated and not make any extreme movements, to avoid disrupting the contact of the EEG. They will then undergo 5 minutes of uninterrupted time to explore and interact with the Green Meadows environment in Nature Treks VR, running natively on the Oculus Quest 2 at default graphical and audio settings, while the EEG headset collects data. After this, the equipment will be removed from the participant, and they will be asked to answer on the SUDS scale again in a different colour, to differentiate between the before and after.

### 4.3.6 – Processing

EEG recordings can have a variety of sources of interference that may affect the signal. This signal should be cleaned of any such artifacts before analysis can take place. To process the dataset, the EEGlab program, an extension for MATLAB, will be used. This is a widely used open-source program that can read a time series EEG recording, offer tools and visual graphs to aid in the artifact removal and data analysis process.

The data will first be bandpass filtered, to remove frequencies from the recording that will not be useful within the analysis. After this, the data will be visually inspected to remove any bad channels or easily identifiable artifacts. After this, Independent Component Analysis will be used to split the signal into several individual components which can be inspected for further artifacts.

### 4.3.7 – Analysis and Testing

After the data processing has completed, the data can be analysed using Power Spectral Density. This will identify the strength of various frequencies throughout the EEG recording. This can be independently analysed to identify which frequencies had the most activation throughout the experiment, as well as the overall change in frequency in comparison to a baseline test. A script should be developed in MATLAB using EEGlab libraries and functions to extract the mean PSD value for the alpha frequency in both the baseline and experimental results. The change between them can then be calculated. The SUDS scale ratings can also be analysed independently to identify whether the participants perceived a change in their internal feelings of distress or anxiety. Both of these variables can then be analysed together to identify whether a relationship exists between them.

The analysis of descriptive statistics and statistical analyses should then be performed in order to evaluate the results of the experiment. As discussed in Chapter 3, this will involve calculating a correlation coefficient using either Pearson’s R or Spearman’s Rho, depending on the normality of the data distribution. A t-test can then be applied to calculate the level of statistical significance. Development of tools in R should occur to facilitate the analysis of descriptive and inferential statistics.

## 4.4 – Procedure

A sample of participants was acquired through volunteer sampling, by advertising in relevant online university spaces such as course-specific Microsoft Teams servers and society Discord servers. Initially, six participants were recruited, however three participants were unable to continue with the study due to unforeseen circumstances. A higher number of participants would have been preferable to gain more accurate, generalisable, and statistically significant results.

After initial interviews, three healthy individuals of 19-21 years of age agreed to take part in the study. All participants were given a brief detailing the purpose of the experiment, as well as what their involvement would entail, and gave their written, informed consent to take part in the study. All participants stated they did not have a history of motion sickness or epilepsy, and did not have much previous experience with VR, although they were interested in it. Additionally, all participants stated that they were under some levels of distress due to university workload.

Participants signed an informed consent form, and then briefed that they would experience the VR environment for 5 minutes. They were shown the EEG headset, and how it worked and would fit over the VR headset was explained to them. They were given an opportunity to ask any questions, and then rated a number on the SUDS scale, outlining their current levels of distress. The participant was then seated to avoid excessive movement of the VR headset, and the VR headset was then placed on the participant and adjusted until comfortable. The participants were given the controllers and taught how to use them before beginning the experiment, to avoid disruptions.

The EEG headset was then fitted on the participant and connected to the OpenBCI GUI interface for collecting data. Before the VR headset was powered on, a data collection session was started on the GUI interface for 30 seconds in order to collect a baseline, for comparison purposes. During this period, the participants were asked to sit completely still, with their eyes closed. The VR headset was then turned on and a new data collection session started. During this time, the participants were free to explore and interact with the experience. An unexpected issue was encounter here, where one of the participants accidentally pressed a button that exited the game. The experiment had to be restarted in this case. In future investigations, consideration should be taken care to fully allow the participants to gain familiarity with the controllers before the experiment begins. After 5 minutes, the data collection session ended, and the apparatus was removed from the participant. The participant was then asked to mark their distress on the SUDS scale again, and the procedure was repeated for all participants.

## 4.5 – Data Processing

Artifacts in EEG signals can occur from several sources, such as voluntary and involuntary muscle movements, eye movements and blinking, and external sources such as ambient electrical waves from electrical equipment. It is important to minimise the number of artifacts present in EEG data for the data to be accurate and useful for analysis.

While care was taken to reduce the number of artifacts within the data, such as seating the participants for the duration of the experiment to avoid excessive movement, the participants were still able to freely move their head and torso in order to observe and interact with the VR environment.

### 4.5.1 – Pre-processing

Chart, histogram

Description automatically generatedA picture containing text, screenshot, document

Description automatically generatedBefore attempting to suppress artifacts within the dataset, the data was bandpass filtered between the frequencies of 4Hz – 47Hz, as delta frequencies in the brain below 4Hz are associated with sleep, while frequencies above 47Hz may come from external interference such as power line noise, which operates at 50Hz-60Hz (Shamlo-Bigdely, et al., 2015).

**Figure 7: EEG Recording Before (Left) and After (Right) Bandpass Filter**

Timeline

Description automatically generated with medium confidenceThe EEG data is then plotted using a stacked display and inspected visually for each participant. Here, any explicit artifacts can be identified and marked for removal. This included the removal of channels that did not collect enough data due to ill-fitting electrodes, or channels that were particularly noisy and had a much larger amplitude than others. Then, individual sections of the data can be marked for removal where artifacts can be seen, such as large spikes in amplitude. Upon visual inspection of the EEG data, it was extremely noisy. This was likely due to signal interference from the VR headset itself, as well as movement from the participants. This has the possibility of reducing the usefulness of the data. In future investigations, more passive VR experiences could be used to avoid this.

**Figure 8: EEG Recording Visual Artifact Removal**

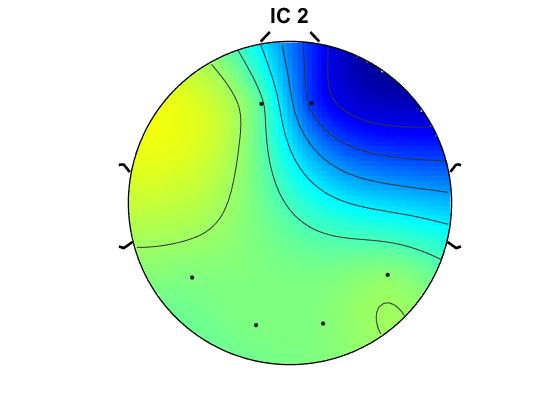
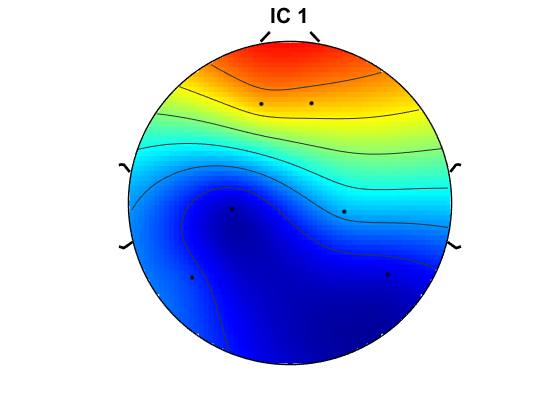
### 4.5.2 – Artifact Removal

When collecting data using EEG, a further source of artifacts can occur from various groups of neurons activating and interfering with the EEG signal. In this case, algorithms such as Independent Component Analysis (ICA) can be applied in order to clean the signal and remove artifacts caused by interference, through separating data into several individual components of the EEG signal that may not be produced by brain activity, which can then be visually inspected and removed from the dataset.

A variety of algorithms for implementing ICA are available. Second Order Blind Identification (SOBI) was chosen as it has shown to be highly effective for removing artifacts from neuronal interference, eye movements, blinking, and muscle movement (Tang, et al., 2005). SOBI ICA was applied to the EEG data for each participant individually.

The components generated by the SOBA ICA algorithm are then inspected visually to identify artifacts from eye movement, muscle movement, and blinking, and these components are selected to be removed from the dataset.

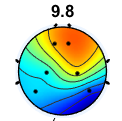
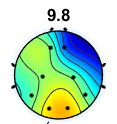
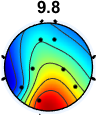
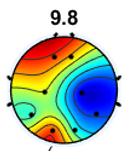
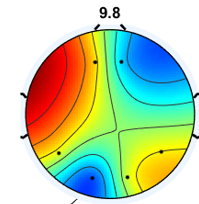
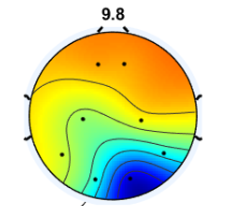
**Figure 9: Lateral Eye Movement Artifact (Left) and Blinking Artifact (Right)**



### 4.5.3 – Power Spectral Density

To identify which frequencies were the most powerful over the course of the EEG data capturing, the average Power Spectral Density can be calculated. This is a measure of a signals power over a specific frequency. Through calculating the mean PSD for the alpha frequency for the baseline EEG recordings, as well as for the experimental recordings, the change in power of the alpha frequency can be calculated. Welch’s method is a common approach for calculating PSD, as it can improve the accuracy through reducing noise by sacrificing the frequency resolution. This method of PSD calculation is preferable over others, as the collected data is noisier than is desired due to interference from the VR headset.

**Figure 10: Heatmaps of Alpha power of all participants (top to bottom) in baseline recording (left) and experimental recording (right)**



A MATLAB script was developed to calculate the PSD, using libraries available from EEGlab. The spectopo() function offered by EEGlab is able to calculate a matrix of average power spectra over ascending frequencies. The program should be able to take an EEGlab data source as an input variable and apply the spectopo function using Welch’s method to generate the matrix. The mean power of the alpha band should then be calculated from this matrix and displayed.

Chart

Description automatically generatedFollowing this guideline, the program was developed, using a window size of 256 samples and an overlap of 50% as the parameters for Welch’s method. This function was successfully able to calculate the mean PSD of the alpha frequency band in each EEGlab data source. The EEGlab GUI interface can also be used to visualise the PSD of various frequency bands through generating a periodogram.

**Figure 11: Periodogram of Experimental Recording**

### 4.5.4 – SUDS Ratings

The ratings on the SUDS scale of measurement were also processed for analysis. This was done through calculating the change before and after the VR intervention. A negative number would represent decreased subjective levels of distress after VR intervention, while a positive number would represent increased subjective levels of distress.

### 4.5.5 –Developing Statistical Analysis Functions

R was chosen to develop programs for the calculation and visualisation of descriptive statistics and statistical testing. It is an open-source scripting language designed for the purpose of statistical computing and was chosen over other scripting languages such as Python due to its primary use case being statistical testing, while Python provides a more general approach to data science.

The three primary areas that need to be developed are functions for data visualisation, functions for descriptive statistics, and functions for inferential statistics. For data visualisation, graphs such as bar graphs to visualise change, as well as histograms to identify data distribution for further testing will be required. For descriptive statistics, functions for calculating average values and correlation coefficients will be needed, as well as graphs to visualise correlations. Finally, a function to calculate a p value and report whether the data is statistically significant using a t-test will be required for statistical analysis.

An R script containing a series of functions for the statistical analysis of the data was created, and the dataset was converted to a comma separated value (CSV) file type for ease of analysis. The script would initialise any library dependencies, and then load the CSV files as global variables to avoid unnecessary repetition and increase efficiency. For the purposes of descriptive statistics, the standard R mean() and sd() functions were used to calculate the mean and standard deviation of the dataset.

Then the ggplot2 library was used for the visualisation of the data. This is a powerful open-source R package that provides many functions for the visualisation of many different types of graphs. This library provides an advantage over the base R functions for plotting as the functions are more verbose, allowing for ease of use. Additionally, it provides many customisation features to make the graphs more visually pleasing, so data can be inferred from them easily.

The first type of graph that was developed was a bar chart. Three separate functions were created – a bar chart for the visualisation of SUDS scores before and after VR intervention, a bar chart for the visualisation of PSD before and after intervention, and a stacked bar chart for the comparison of the change in SUDS scores and the change in PSD. These would be useful for descriptive statistics as it allows, immediately, to identify how the VR intervention has affected both dependent variables., and if those variables appear to be related. Next a boxplot function was developed to assess whether the data contained any outliers. This would be important for inferential testing as outliers may reduce the accuracy and validity of the data. A function to plot a scattergram for the visual identification of a correlation was also developed. This would assist in identifying whether a correlation existed before inferential testing, to provide an expected result of the correlation coefficient. Finally, a density distribution diagram would be developed for the visualisation of the distribution of all variables. This is an important stage in descriptive statistical analysis as the normality of the data will suggest what kind of statistical test to perform.

After this, the functions for conducting inferential testing can be developed. Firstly, to supplement the distribution graph and receive a more objective measure of the level of normality of the data, the Anderson-Darling statistical test was implemented to calculate the level of normality for each set of data. This would provide an objective measure of whether to use a parametric or non-parametric test, as opposed to subjectively interpreting which test to use based on visual analysis. Then, the statistical testing algorithms can be implemented. The standard R functions were used to implement the Pearson’s R test and the Spearman’s Rank test. Both functions automatically apply a single-tailed t-test to the generated r value and provide all the details of the outcome of the test. Additionally, the Paired t-test and Wilcoxon test were also implemented for the analysis of the base data.

## 4.6 - Testing

Once the development of statistical analysis tools has completed, the experimental methodology can then be tested to ensure that it produces results that are accurate, reliable, and produce results that suggest further study into the area will be useful.

The primary method of testing the methodology will be through passing the data through each function that was developed in R, to ensure that all the data is capable of statistical analysis. An issue here was encountered due to the lowest possible sample size of the Anderson Darling test in R being 7. The sample size of the data was 3, and therefore the normality tests required prior to a parametric statistical test were unable to be performed. As it is difficult to acquire a larger sample size at this stage in the investigation, the Spearman’s Rank correlational test would need to be used as it requires fewer assumptions about the data, despite the weaker testing power than its parametric counterpart. Every other function that had been developed produced accurate and useful results or graphs.

# Chapter 5 – Results

## 5.1 – Introduction

In this section, the results of the experiment will be analysed and discussed, and the developed experimental methodology will be evaluated against outlined testing criteria, in order to assess the success of the investigation, and whether future work into studying this field could be beneficial.

## 5.2 – Analysis of Results

A table of results is as shown below. The analysis of these results will be undertaken in this section, including the generation of graphs and statistics using the developed functions within R.

**Table 1: Table of Results**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Participant 1 | Participant 2 | Participant 3 |
| PSD Before | -11.8 | -13.4 | -17.6 |
| PSD After | -3 | -11.4 | -15.3 |
| PSD Change | 8.8 | 2 | 2.3 |
| SUDS Before | 25 | 20 | 30 |
| SUDS After | 15 | 5 | 20 |
| SUDS Change | -10 | -15 | -10 |

### 5.2.1 – Analysis of PSD

Studying the results of the Power Spectral Density, it can be identified that through all participants, the amount of alpha frequency band power showed an increase between the baseline readings and the experimental readings. As increase in alpha band power suggests an increase towards positive valence, it can be assumed that the VR experience had an effect of relaxation and in the participants and induced pleasant emotions within them, therefore changing their affective state.

Chart, waterfall chart

Description automatically generatedAdditionally, for all participants, the alpha or theta frequency bands were showed the highest average PSD after VR intervention. These frequency bands are most associated with increased levels of relaxation and the processing of positive stimuli, therefore the idea that participants exhibited a change towards positive valence and pleasant affective states can be reinforced.

**Figure 12: Bar chart to Show Before and After PSD**

While the results shown in the PSD analysis suggest that there was a change towards positive valence in the participants, the accuracy and validity of the results could be improved through utilising an EEG headset with a greater number of electrodes. This would allow for the acquisition of more data, which would lead to more precise measurements. It would also allow for the analysis of alpha band power in specific regions of the brain with increased accuracy, which would allow the experiment to draw meaningful conclusions about structures within the brain that are associated with positive valence and relaxation. As an EEG headset with a greater number of electrodes was difficult to acquire for this experiment, due to increased costs, the results can be considered limited in this regard. Despite this, however, the results do show promise, and indicate that VR methods of relaxation can increase alpha power, and lead to positive valence, which is successful for aims of the study.

### 5.2.2 – Analysis of SUDS Ratings

Chart, bar chart

Description automatically generatedThe findings from the SUDS Ratings show that all participants reported mild to moderate levels of distress before the experiment took place. After VR intervention, all participants reported lower levels of distress after undergoing VR intervention.

**Figure 13: Bar Chart to Show Before and After SUDS Ratings**

This suggests that the VR intervention was able to decrease the subjective feelings of distress within the participants, as all of them reported a decrease in distress levels. This is successful for the aims of the study, as it suggests that the experimental task was able to both lead the participants to a pleasant affective state, as well as reducing distress. This is promising for the capabilities of VR for at-home reduction of distressful feelings that may be caused by symptoms of anxiety and depression. A limitation in this regard, however, is that self-report can be subject to bias, as the participants may adjust their responses to align better with the expected outcomes of the study. Strategies to mitigate bias were taken, such as allowing the participants to answer on the SUDS scale in privacy, so they wouldn’t perceive pressure from the experimenter. Additionally, it was explained to the participants in the initial brief that honesty was of utmost importance when answering on the scale. However, additional strategies could have been employed to increase the validity of the results, such as decreasing demand characteristics by deceiving the participants about the aim of the study. This however may lead to ethical implications, the nullification informed consent, as the participant may feel as though they consented to the study that was explained to them, and not the study that took place. This would also require a thorough debriefing, leading to increased time required of the experiment. Another method to reduce bias would be using physiological measurements such as Heart Rate Variability in conjunction with self-report to confirm the results. This would lead to increased objectivity, however, would incur more time and resources required for the experiment. As mentioned in previous chapters, the subjective nature of self-report is advantageous in this situation, therefore the full replacement of it with objective measurements may reduce the effectiveness of Chart, bar chart

Description automatically generatedthe study.

**Figure 14: Bar Chart to Compare Change in PSD vs Change in SUDS Ratings**

### 5.2.3 – Analysis of Descriptive Statistics

Descriptive statistics are a useful tool in measuring the characteristics of the data, to make assumptions and prepare for statistical analysis.

From calculating the mean values of the alpha PSD, it can be identified that the average alpha PSD before VR intervention is -14.2. The average alpha PSD after intervention is -9.9. This suggests an average increase of 4.3 alpha PSD. Additionally, the standard deviation of the alpha PSD indicates that the baseline readings were much more clustered around the mean, with a standard deviation of 2.9. This suggests all participants had similar alpha frequency power before the VR intervention. After the intervention, the standard deviation is much higher, at 6.2. This suggests that the VR intervention had varying levels of impact on the alpha frequency power between participants. Some participants may have been more susceptible to positive valence and pleasant affective states through the environment and VR experience chosen. It can be inferred that either the selected experience of natural landscapes or the medium of VR may be more effective for some people than others in influencing a specific affective state. Therefore, an aspect that could be researched further is the influence of different types of media on affective state. Additionally, a higher standard deviation could suggest that the data may be less reliable. This is likely since the experiment had a small sample size.

The average rating on the SUDS scale before intervention was 25, which is a quarter of the possible total rating. This is characterised subjectively on the scale as between minimal and moderate levels of anxiety or distress, with no interference in functioning. The average rating after intervention was 13.3, which is characterised as being alert and awake, and concentrating well. This suggests that all participants experienced a mean decrease of 11.6 in their subjective feelings of distress. We can infer that all participants began the experiment with mild to moderate feelings of distress. After VR intervention, their subjective feelings of distress all lowered, therefore it is likely that the VR intervention had a positive effect on levels of anxiety and distress. There may also be a relationship between decreasing levels of subjective distress and increase alpha frequency band power, which will be assessed with statistical analysis. Additionally, it can be identified that the standard deviation is greater after the VR intervention (7.6) than before (5). This suggests that all participants began the VR experience with similar levels of subjective distress, and the intervention had varying levels of impact on them. However, the standard deviation of the average change was 2.8, suggesting that the change experienced through all the participants was similar. This may suggest that the higher standard deviation in the after values may be due to the varying levels of subjective distress before the intervention. The experiment affected the participants in similar amounts, however the participants that began with higher subjective distress Chart

Description automatically generatedChart

Description automatically generatedthan others also ended with higher subjective distress than others.

**Figure 15: Box plots to show outliers in PSD data (left) and SUDS data (right)**

A boxplot can be used for the detection of any outliers within the data. This is an important step in descriptive statistics as any outliers may lower the reliability and validity of the dataset, thus leading to the possibility of making incorrect inferences. Boxplots were plotted for all of the data, and it can be inferred that no outliers exist within the dataset, as there are no highlighted data points outside of the plots. This suggests that there are no outliers within the data, therefore increasing the validity of it and improving the accuracy of any inferences. Eventually, within statistical analysis, an algorithm can be used to gain an objective measure of outliers in conjunction with this graph.

Chart, line chart

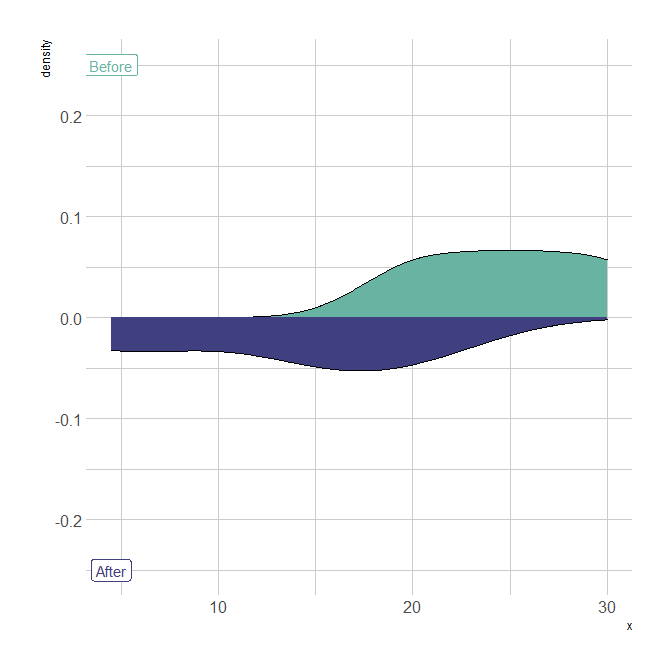
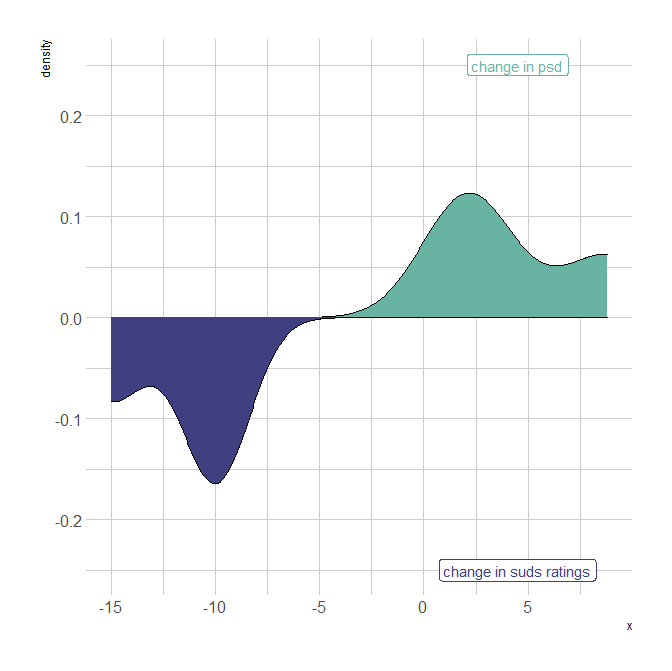
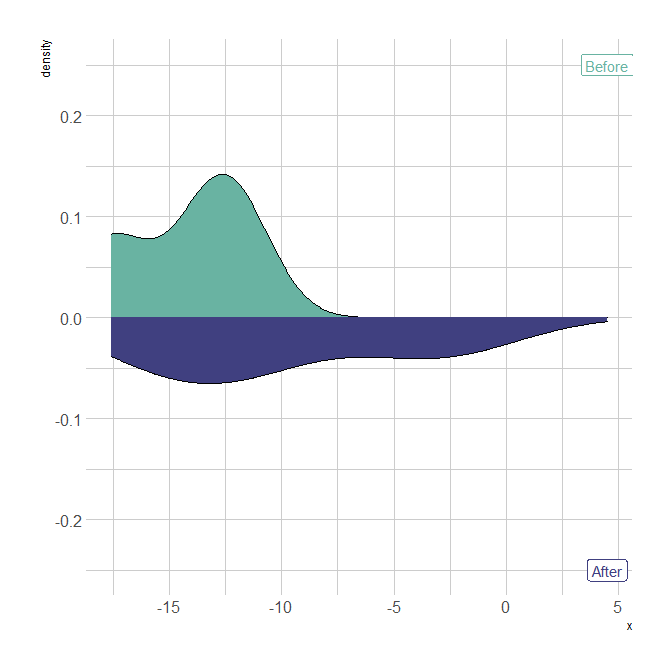
Description automatically generatedNext, a scattergram was used to visually assess whether a correlation existed between the change in SUDS ratings and change in alpha PSD. A line of best fit was automatically calculated within R to identify the direction of a correlation, if one exists. From the scattergram it can be identified that a positive correlation appears to exist between the two variables, which suggests that there is a possibility of a relationship between increasing valence and decreasing subjective feelings of distress. The correlation between the two variables appears to be strong, therefore it is likely to achieve a value close to 1 when assessing the correlation coefficient.

**Figure 16: Scattergram to Show Relationship Between Change in PSD and Change in SUDS Ratings**

Finally, for the purposes of deciding whether a parametric test or non-parametric test should be used to perform statistical analysis on the data, the distribution of the datasets should be identified. If all the data appears to be normally distributed, a parametric test can be used. Otherwise, a non-parametric test must be used. Parametric test can be more powerful than non-parametric; however, both are useful in making assumptions about the data.

After plotting density distribution diagrams for the entire dataset, it can be identified that the data does not appear to follow a normal distribution. Primarily, the PSD scores after VR intervention are widely distributed, which is in line with the standard deviation that was calculated for the PSD data. Additionally, the SUDS ratings from both before and after the intervention are widely distributed. Additionally, the Before scores for PSD and the change scores have a bimodal distribution.

**Figure 17: Scattergram to Show Density Distribution of PSD Before and After (Top Left), SUDS Before and After (Top Right), and Change in PSD vs Change in SUDS (Bottom Middle)**



### 5.2.4 – Statistical Testing

After an analysis of the descriptive statistics has been completed, the calculation and analysis of inferential statistics can commence. The first task is to objectively confirm the predictions made using the boxplots, density distribution diagrams, and scattergram.

The Anderson-Darling statistical test was implemented in R to assess the distribution of the data. This test produces a p-value which suggests the probability of the data not fitting a normal distribution. Unfortunately, due to the sample size of the dataset being 3 and the minimum required sample size of the Anderson-Darling test in R being 7, the test was not able to be performed on the data. The data should not be assumed to be normally distributed, as this may lead to incorrect statistical analysis being performed. Therefore, non-parametric test will be used for further analysis. While these are less powerful than parametric tests, they make less assumptions about the data, and therefore apply in this case.

Next, the Tukey’s Fences method will be used as a non-parametric tool for detecting any outliers. A k-value of 2.2 will be used to describe the threshold of what is classified as an outlier. Through this method, it was identified that no major outliers exist in the dataset. This means that further statistical analysis is possible without making incorrect inferences from the data.

Non-parametric statistical testing can then be used to assess the level of significance of the dataset. In order to calculate the values for the before and after PSD and SUDS scores, the Wilcoxon Signed-Rank test will be used. This is a non-parametric test for paired data, to test the level of significance of the differences between two groups. Paired data means the data was gathered from the same participants in both conditions. A single tailed test is used as there is a predicted outcome based on the hypothesis.

For the PSD data, a p-value of 0.125 was calculated. This suggests that the results are not significant at the p < 0.05 level of significance. Meanwhile, for the SUDS data, a p-value of 0.0867 was calculated. This is not significant at p < 0.05, however if the 0.10 level of significance is used, then these results could be considered statistically significant. This would, however, increase the risk of falsely rejecting a null hypothesis, and therefore, the 0.05 level of significance will be assumed for all data.

Finally, the correlation coefficient between the change in SUDS ratings and change in alpha PSD was calculated using Spearman’s Rank. This resulted in a correlation coefficient of 0.866, which suggests that a large positive correlation between the two variables exists. This confirms the analysis of the scattergram and suggests that there is some kind of relationship between increasing alpha PSD and decreasing subjective ratings of distress. Due to the nature of correlations, however, it cannot be assumed that increasing alpha PSD is the cause of decreasing ratings of distress. Other factors may be acting as confounding variables, such as the type of VR experience used, or potential bias. The significance of this should then be tested using a t-test. This resulted in a p-value of 0.1667. which suggests that the results are not significant at the 0.05 level of significance, as the p-value is greater than 0.05. The risk of rejecting the null hypothesis is 16.67%, which is too high.

## 5.3 – Evaluation

The experimental design should be analysed against the success criteria that was defined in chapter 3. The table below shows each criteria, and whether it can be considered to be met. This will be followed by an analysis and explanation of each criterion.

Table 2: Table of Success Criteria

|  |  |
| --- | --- |
| Criteria | Achieved |
| A majority of participants state that the experiment was overall, a positive experience | Yes |
| A majority of participants state that the experiment helped them in inducing pleasant emotions and reducing distress | Yes |
| No participants had any major difficulties that disrupted the experiment | No |
| No participants expressed major discomfort during the experiment | Yes |
| The results suggest a positive correlation between the variables | Yes |
| The results suggest a strong positive correlation between the variables | Yes |
| The results suggest statistical significance of the experiment | No |

The results following an analysis of the success criteria suggest 2 out of the 7 listed criteria were not met. During the experiment, one participant accidentally pressed a button on the controller of the Oculus Quest 2 that closed the VR experience. This meant that the experiment had to be restarted for that participant, which is classed as a major disruption. This could have been mitigated through a more in-depth briefing of the controls of the VR headset. Additionally, the investigation has not shown statistical significance of the results, which suggests that further study is necessary to assess whether the results may be a consequence of random chance.

Despite this, all participants stated that the experiment was a positive experience, and that they felt more pleasant and relaxed after undergoing the VR experience. No participants expressed any major discomfort that disrupted their participation in the experiment from any of the apparatus used. There was some mild discomfort from the electrodes of the EEG headset, however. And finally, the results suggest a strong, positive correlation between the dependant variables.

## 5.4 – Discussion

To finalise this chapter of the project, the primary results will be summarised and discussed. The experiment found a non-significant, large positive correlation between the decrease in SUDS ratings, and the increase in alpha frequency PSD. All participants reported lower levels of distress and displayed higher alpha PSD than their baseline readings after the VR intervention.

One of the main weaknesses of these results are that they do not show statistical significance, which could have been mitigated through a larger sample size. Additionally, non-parametric tests were required as the sample size was not great enough to assess for the normality of the distribution of data. It can be inferred, therefore, that the sample size of the experiment was the main limiting factor in further success, as the null hypothesis is not able to be rejected at this stage.

Despite this, a large positive correlation shows that a relationship between the two variables exists, and if the experiment had acquired a larger sample, statistically significant results may have been possible. This suggests that the alternative hypothesis also cannot be rejected, and further, larger scale testing may need to occur.

Finally, the evaluation of the experimental methodology shows that the experiment was mostly a success, with the major drawbacks being the statistically insignificant results, and the participants’ familiarity with the controllers of the VR headset. This could potentially be mitigated in further study through a more in-depth briefing, and the allowing of a period where the participants can familiarise themselves with the controls better before the experiment begins.

# Chapter 6 – Conclusion

## 6.1 – Conclusions

In conclusion, this project consisted of the development of an experimental methodology to assess whether VR could be used to change an individual's affective state and induce positive emotions within them, while decreasing levels of distress. This project implemented this methodology and found that VR was able to increase the alpha frequency band power in a participant’s brain, which is associated with higher levels of valence, meaning increased pleasure. Additionally, it was found that participants reported lower subjective levels of distress in response to VR. These variables were correlated, with a strong positive correlation indicating that increasing levels of alpha power and decreasing subjective feelings of distress may be related.

However, the investigation was not without limitations. The most major issue was that a very small sample size was acquired, of 3. This means the findings of the study cannot be generalised to a wider population and are therefore limited. Additionally, as a consequence of the small sample size, parametric tests were unable to be used as normality cannot be assumed. Finally, the achieved results were not significant at the 0.05 level of significance, which means the null hypothesis cannot be rejected. However, the alternative hypothesis also cannot be rejected due to the strong positive correlation achieved from the statistical analysis. With a wider sample, it is possible that statistically significant results could have been achieved.

Overall, the project can be considered successful, but with issues that need to be addressed in possible future implementations. The success criteria found that 5 out of the 7 criteria had been met, and the only major drawbacks were the sample size, and brief adjustment period allowed for the participants to become acquainted with the controls of the VR headset. Other than this, all participants reported that they felt the experiment as a positive experience and beneficial to them in reducing stress. The strong, positive correlation suggests that VR has the potential to be a capable means of reducing distress in an individual who may be suffering from symptoms of anxiety and depression, such as stress, through the adjustment of affective state.

## 6.2 – Future Work

Future study into this topic would be beneficial. The primary ways in which the drawbacks of the experiment could be addressed is through repeated testing. The experiment was designed to be easily replicable, using commercially available tools that can be acquired easily and at a low cost. A higher level of participant gathering methods could be utilised in future replications, such as wider advertisement of the study as well as combining different sampling methods. This would lead to a higher sample size, which means the study could be conducted on a much wider scale. As a strong positive correlation was achieved, and the p-value was close to being statistically significant, this could lead to an experiment with statistically significant results that can be generalised to the wider population. Such studies could lead to the further research into low-effort, at-home VR techniques for stress reduction, as well as a wider commercial adoption of VR applications and games designed for this specific purpose. This would prove hugely beneficial for a younger population that spends more of its time online, is attracted towards technology, and is found to have poorer overall mental health.

Additionally, the methodology of the experiment can be refined. Some difficulties were experienced as some participants were not fully comfortable with the controls of the VR headset, leading to the experiment needing to be restarted on one occasion. This could have affected the results, as the participant had already begun the VR experience, as therefore had already adjusted to it. Additionally, the participant may feel a higher level of distress, due to suddenly and unexpectedly exiting the VR experience, which could be jarring. To improve the reliability of future results, the methodology could incorporate time before the experiment to allow the participant to fully familiarise themselves with the controls of the VR headset and proceed when they feel ready.

Both techniques would have the advantage of building upon the successes of the experiment. As a strong positive correlation has already been identified, doing wider scale, more robust experimentation can lead to better, more accurate results that reliable inferences can be made from.

## 6.3 – Legal, Social, Ethical, and Professional Issues

### 6.3.1 - Legal Issues

The primary legal issue of this project is the Data Protection Act (Data Protection Act, 2018). The project collected data on participants, and therefore this data needed to be stored in accordance with the DPA. Any data collected on the participants must only be used fairly and lawfully, and only for the purposes that have been explicitly specified. The data that the experiment collected was only used for the intended, explicit purpose that was stated in the project. This was made clear to the participants in their briefing. The data be kept accurate and up to date and should not be kept for longer than it is needed. The data will only be kept for a specified amount of time as stated in the ethical application, and thereafter will be destroyed. The data should also be stored securely in a way that ensures protection against any unauthorised access. The data was stored on university-hosted cloud platforms, protecting against unauthorised access of a locally stored dataset. All data will also be anonymised and encrypted to protect the participants’ privacy, ensuring that their identity cannot be revealed in the event of a security breach. The participants names were not used within the study, and numbers were used instead.

### 6.3.2 - Social Issues

The primary ethical implication of this project is privacy. Some companies such as Valve and Meta have been researching the potential incorporating BCI into VR headsets directly, allowing for increased data collection of the users of VR headsets. This has implications on the issue of targeted advertising, as such companies may be able to identify a user’s neural activation in response to the content displayed on the VR headset. This information could be valuable to advertisers and sold. Another issue could be social isolation. Incorporating low intensity and low effort symptom management techniques into virtual may lead to individuals wanting to spend more time within the virtual world. Due to the idea that VR management techniques require less effort and motivation, users may be less inclined to seek more intensive measures of intervention such as psychotherapies. However positive impacts could also be made, as VR could be utilised as a measure to improve general mental health of a population that already spends much time using technology, through the adoption of the technology as a medium for symptom management for mental issues.

### 6.3.3 - Ethical Issues

As the project will consist of an experiment on a sample of individuals, it was necessary for any health and safety precautions to be taken, to prevent risks to people’s health during the experiment. For example, participants were screened for photosensitive epilepsy, as well as any other conditions that may create risks when using a VR headset, such as motion sickness. Some participants may not have been able to use the controllers due to motor impairments, and considerations should have been made in advance, such as alternative controllers designed for motor disabilities. Due to the project requiring voluntary participants, a bias of the sample in gender or ethnicity may have emerged, which may have led to unrepresentative results.

### 6.3.4 - Professional Issues

Throughout the project, a certain level of professionalism and integrity must be upheld. The BCS Code of Conduct (BCS, 2021) was closely followed, which outlines the key principles that an individual conducting work within the tech industry should uphold. All participants were interacted with in a professional and respectful manner, without any bias or discrimination. Care to uphold the integrity of the work and subject matter was taken.

## 6.4 – Synoptic Reflection

I have gained a vast array of skills from both from my time at Nottingham Trent University as a whole, and from undertaking this project in my final year, that will lead me to success in years to come.

This project has given me experience in undertaking a large responsibility, setting my own goals and deadlines, and trying my best to meet them. I have learnt how to deal with factors outside my control when these personal deadlines cannot be met in a calm and professional manner.

The project has taught me new ways to research and analyse a topic that is adjacent to my own primary area of professional interest, being able to extract useful information from large academic sources and contrasting them with others. I have been capable of highlighting the strengths and weaknesses of topics I am new to, through the quick learning of information and the ability to find sources that provide a balanced review of the topic. This has led me to be able to quickly analyse and discuss underdeveloped areas of research within a broader topic and focus my interest in developing that area.

I have learnt how to conduct a laboratory experiment in line with scientific principles and ethical guidelines, to investigate a specific area of research. This has been done through the designing of an experimental methodology, that needed to be reproducible, falsifiable, and empirical. I have taught myself a variety of tools and resources for the purposes of statistical analysis, such as R and MATLAB, which are programming languages I previously had no experience in. This means I was able to successfully complete a statistical analysis on a set of data, and critically evaluate the usefulness of the outcomes, teaching myself concepts about parametric and non-parametric tests, the requirements for each, and what information the results convey.

All these skills will aide me in future professional endeavours, as I will be able to take these skills with me into further, postgraduate study to complete more effective, robust pieces of research that may have a positive impact on the lives of people. I can also take the skills I have learnt with me into a variety of professions. The primary skills of time management in large projects, knowing how to quickly and reliably learn new technologies and theories, and critically analysing pieces of work and solving problems that arise with the methodology can be transferred into any software development role. The programming languages of R and MATLAB, as well as skills and experience gained in statistical analysis and handling a set of data can be extremely useful in a data analytics or data science role.

In conclusion, this project has helped me achieve my goal of providing research that can help in improving the lives of a population, and the skills and experience I have gained from it can be used in furthering developing my academic and professional career.

# References

Abhang, P. A., Gawali, B. W. & Mehrotra, S. C., 2016. *Introduction to EEG and Speech-Based Emotion Recognition.* s.l.:Academic Press.

Aim, F., Lonjon, G., Hannouche, D. & Nizard, R., 2016. Effectiveness of Virtual Reality Training in Orthopaedic Surgery. *Arthroscopy: The Journal of Arthroscopic & Related Surgery,* 32(1), pp. 224-232.

American Psychiatric Association, 2013. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5.* 5th ed. Washington D.C.: American Psychiatric Publishing.

Auerbach, J. & Miller, B. F., 2020. COVID-19 Exposes the Cracks in Our Already Fragile Mental Health System. *American Journal of Public Health,* 110(7), pp. 969-970.

Baghaei, N. et al., 2021. Virtual Reality for Supporting the Treatment of Depression and Anxiety: Scoping Review. *JMIR Ment Health,* 8(9).

Banos, R. M. et al., 2002. Virtual reality treatment of flying phobia. *IEEE Transactions on Information Technology in Biomedicine,* 6(3), pp. 206-212.

Barrett, L. F. & Russell, J. A., 2015. *The Psychological Construction of Emotion.* 1st ed. New York: The Guilford Press.

BCS, 2021. *BCS Code of Conduct: BCS.* [Online]   
Available at: https://www.bcs.org/membership/become-a-member/bcs-code-of-conduct/  
[Accessed 02 11 2021].

Bestelmeyer, P. E. G., Kotza, S. A. & Belin, P., 2017. Effects of emotional valence and arousal on the voice perception network. *Social Cognitive and Affective Neuroscience,* 12(8), pp. 1351-1358.

Bradley, M. M. & Lang, P. J., 1994. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry,* 25(1), pp. 49-59.

Bull, P., 1978. The interpretation of posture through an alternative methodology to role play. *British Journal of Social Care and Clinical Psychology,* 17(1), pp. 1-6.

Cacioppo, J. T., Bernston, G. G., Larsen, K. M. & Ito, T. A., 2000. The psychophysiology of emotion. In: R. Lewis & J. M. Haviland-Jones, eds. *Handbook of emotions.* s.l.:Guilford Press, pp. 173-191.

Cacioppo, J. T., Gardner, W. L. & Bernston, G. G., 1997. Beyond Bipolar Conceptualizations and Measures: The Case of Attitudes and Evaluative Space. *Personality and Psychology Review,* 1(1), pp. 3-25.

Cahn, B. R. & Polich, J., 2006. Meditation states and traits: EEG, ERP, and neuroimaging studies. *Psychological Bulletin,* 132(2), pp. 180-211.

Cannon, W. B., 1927. The James-Lange Theory of Emotions: A Critical Examination and an Alternative Theory. *The American Journal of Psychology,* 39(1), pp. 106-124.

Colodro-Conde, L. et al., 2018. A direct test of the diathesis-stress model for depression. *Molecular Psychiatry,* 23(7), pp. 1590-1596.

Cullen, W., ulati, G. & Kelly, B. D., 2020. Mental health in the COVID-19 pandemic. *QJM: An International Journal of Medicine,* 113(5), pp. 311-312.

Darwin, C., 1872. *The Expression of the Emotions in Man and Animals.* Chicago: University of Chicago Press.

Data Protection Act, 2018. *Data Protection Act 2018: Legislation.* [Online]   
Available at: https://www.legislation.gov.uk/ukpga/2018/12/contents/enacted  
[Accessed 02 11 2021].

Desmet, P. M., Vastenburg, M. H. & Romero, N., 2016. Mood measurement with Pick-A-Mood: review of current methods and design of a pictorial self-report scale. *Design Research,* 14(3), pp. 241-279.

Dimberg, U., 1982. Facial Reactions to Facial Expressions. *Psychophysiology,* 19(6), pp. 643-647.

Dimberg, U., Thunberg, M. & Elmehed, K., 2000. Unconscious Facial Reactions to Emotional Facial Expressions. *Psychological Science,* 11(1), pp. 86-69.

Ekman, P., 1957. A Methodological Discussion of Nonverbal Behaviour. *The Journal of Psychology,* 43(1), pp. 141-149.

Ekman, P. & Friesen, V. W., 1971. Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology,* 17(2), pp. 124-129.

Evrim, C., 2021. Gamification: 12 Augmented Reality, Virtual Reality Games and Tourism Marketing Applications. In: F. Xu & D. Buhalis, eds. *Gamification for Tourism.* s.l.:s.n., p. 237.

Francis Jr, A. G., Mehta, M. & Ram, A., 2009. Emotional Memory and Adaptive Personalities. In: *Handbook of Research on Synthetic Emotions and Sociable Robotics: New Applications in Affective Computing and Artificial Intelligence.* s.l.:IGI Global, p. 22.

Freeman, D. et al., 2014. The Use of Immersive Virtual Reality (VR) to Predict the Occurrence 6 Months Later of Paranoid Thinking and Posttraumatic Stress Symptoms Assessed by Self-Report and Interviewer Methods: A Study of Individuals Who Have Been Physically Assaulted. *Psychological Assessment,* 26(3), pp. 841-847.

Harmon-Jones, E., Harmon-Jones, C. & Summerell, E., 2017. *On the Importance of Both Dimensional and Discrete Models of Emotion,* Sydney: School of Psychology, The University of New South Wales.

Holzwarth, V. et al., 2021. Towards estimating affective states in Virtual Reality based on behavioral data. *Virtual Reality,* Volume 25, pp. 1139-1152.

Institute of Health Metrics Evaluation, 2021. *Global Health Data Exchange (GHDx).* [Online]   
Available at: http://ghdx.healthdata.org/gbd-results-tool?params=gbd-api-2019-permalink/d780dffbe8a381b25e1416884959e88b  
[Accessed 27 10 2021].

Ioannou, A., Papastravrou, E., Avraamides, M. N. & Charalambous, A., 2020. Virtual Reality and Symptoms Management of Anxiety, Depression, Fatigue, and Pain: A Systematic Review. *SAGE Open Nursing.*

James, W. & Lange, C. G., 1922. *The Emotions.* 1 ed. Baltimore: Williams & Wilkins Co.

Kehel-Forbes, S. M., Meis, L. S., Spoont, M. R. & Polunsy, M. S., 2016. Treatment initiation and dropout from prolonged exposure and cognitive processing therapy in a VA outpatient clinic. *Psychological Trauma: Theory, Research, Practice, and Policy,* 8(1), p. 107.

Kim, H. et al., 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. *Frontiers in Psychiatry.*

Klug, M. & Gramann, K., 2020. Identifying key factors for improving ICA-based decomposition of EEG data in mobile and stationary experiments. *European Journal of Neuroscience,* 54(12).

Kumar, A. & Nayar, K. R., 2020. COVID 19 and its mental health consequences. *Journal of Mental Health,* 30(1), pp. 1-2.

Kumar, M. R., 2021. *Estimating Affective States in Virtual Reality Environments using the Electroencephelogram,* Virginia: Virginia Commonwealth University.

Leslie, G., Picard, R. & Lui, S., 2015. *An EEG and Motion Capture Based Expressive Music Interface for Affective Neurofeedback.* Plymouth, s.n.

Lidner, P., Hamilton, W., Miloff, A. & Carlbring, P., 2019. How to Treat Depression With Low-Intensity Virtual Reality Interventions: Perspectives on Translating Cognitive Behavioral Techniques Into the Virtual Reality Modality and How to Make Anti-Depressive Use of Virtual Reality–Unique Experiences. *Frontiers in Psychiatry,* Volume 10.

Li, L. et al., 2017. Application of virtual reality technology in clinical medicine. *American Journal of Translational Research,* 9(9), p. 3867–3880.

Mahmoud, J. S. R., Staten, R., Hall, L. A. & Lennie, T. A., 2012. The Relationship among Young Adult College Students’ Depression, Anxiety, Stress, Demographics, Life Satisfaction, and Coping Styles. *Issues in Mental Health Nursing,* 33(3), pp. 149-156.

Makeig, S. et al., 1997. Blind separation of auditory event-related brain responses into independent components. *PNAS,* 94(20), pp. 10979-10984.

Malhi, G. S. & Mann, J. J., 2018. Depression. *The Lancet,* 392(10161), pp. 2299-2312.

Maples-Keller, J. L., Bunnel, B. E., Kim, S.-J. & Rothbaum, B. O., 2017. The use of virtual reality technology in the treatment of anxiety and other psychiatric disorders. *Harvard Review of Psychiatry,* 25(3), pp. 103-113.

Massaro, J., 2009. Experimental Design. In: D. Robertson & G. H. Williams, eds. *Clinical and Translational Science.* s.l.:Academic Press, pp. 41-57.

Mazuryk, T. & Gervautz, M., 1996. *Virtual Reality - History, Applications, Technology and Future,* s.l.: s.n.

Miloff, A. et al., 2016. Single-session gamified virtual reality exposure therapy for spider phobia vs. traditional exposure therapy: study protocol for a randomized controlled non-inferiority trial. *Trials,* Volume 17.

Mottelson, A., Peterson, G. B., Lilija, K. & Makransky, G., 2021. Conducting Unsupervised Virtual Reality User Studies Online. *Frontiers in Virtual Reality.*

Orne, M. T., 1996. Demand Characteristics. In: A. Grayson & P. Banyard, eds. *Introducing Psychological Research.* London: Palgrave, p. 395–401.

Ortony, A., Clore, G. L. & Collins, A., 1988. *The Cognitive Structure of Emotions.* Cambridge: Cambridge University Press.

Pandita, S. & Won, A. S., 2020. Clinical applications of virtual reality in patient-centered care. *Technology and Health,* pp. 129-148.

Parcesepe, A. M. & Cabassa, L. J., 2013. Public stigma of mental illness in the United States: A systematic literature review. *Administration and Policy in Mental Health and Mental Health Services Research,* Volume 40, pp. 384-399.

Parsons, T. D. & Rizzo, A. A., 2008. Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis. *Journal of Behavior Therapy and Experimental Psychiatry,* 39(5), pp. 250-261.

Pfefferbaum, B. & North, C. S., 2020. Mental Health and the COVID-19 Pandemic. *New England Journal of Medicine,* 383(6), pp. 510-512.

Pinilla, A. et al., 2021. Affective Visualization in Virtual Reality: An Integrative Review. *Frontiers in Virtual Reality.*

Plutchik, R., 1982. A Psychoevolutionary Theory of Emotions. *Social Science Information,* 21(4), pp. 529-553.

Posner, J., Russel, J. A. & Peterson, B. S., 2005. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology,* 17(3), pp. 715-734.

Price, J. & Budzynski, T., 2009. Anxiety, EEG patterns, and neurofeedback. In: T. H. Budzynski, H. K. Budzinski, J. R. Evans & A. Abarbanel, eds. *Introduction to Quantitative EEG and Neurofeedback.* New York: Academic Press, p. 453–472.

Quesnel, D. & Riecke, B. E., 2018. Are You Awed Yet? How Virtual Reality Gives Us Awe and Goose Bumps. *Frontiers in Psychology,* Volume 9.

Quesnel, D. & Riecke, B. E., 2018. Are You Awed Yet? How Virtual Reality Gives Us Awe and Goose Bumps. *Frontiers in Psychology.*

Reichert, A. & Jacobs, R., 2018. The impact of waiting time on patient outcomes: Evidence from early intervention in psychosis services in England. *Health Economics,* 27(11), pp. 1772-1787.

Riva, G., 2009. Virtual reality: an experiential tool for clinical psychology. *British Journal of Guidance & Counselling,* Volume 37, pp. 337-345.

Rizzo, A. A. et al., 2004. The Virtual Classroom: A Virtual Reality Environment for the Assessment and Rehabilitation of Attention Deficits. *CyberPsychology and Behaviour,* 3(3), pp. 483-499.

Rizzo, A. et al., 2009. Virtual Reality Exposure Therapy for Combat-Related PTSD. *Post-Traumatic Stress Disorder,* pp. 375-399.

Rosenman, R., Tennekoon, V. & Hill, L. G., 2011. Measuring bias in self-reported data. *International journal of behavioural & healthcare research,* 2(4), pp. 320-332.

Russel, J. A., 1980. A circumplex model of affect. *Journal of Personality and Social Psychology,* 39(5), p. 1161–1178.

Saarni, C., 1999. *The Development of Emotional Competence.* New York: The Guilford Press.

Schachter, S. & Singer, J., 1962. Cognitive, social, and physiological determinants of emotional state. *Psychological Review,* 69(5), pp. 379-399.

Sevinc, V. & Ilker, B. M., 2020. Psychometric evaluation of Simulator Sickness Questionnaire and its variants as a measure of cybersickness in consumer virtual environments. *Applied Ergonomics,* Volume 82.

Shamlo-Bigdely, N., Mullen, T., Su, K.-M. & Robbins, K. A., 2015. The PREP pipeline: standardized preprocessing for large-scale EEG analysis. *Frontiers in Neuroinfomatics.*

Sherman, W. & Craig, A. B., 2003. *Understanding Virtual Reality.* s.l.:Elsevier.

Stirling, E. R., Lewis, T. L. & Ferran, N. A., 2014. Surgical skills simulation in trauma and orthopaedic training. *Journal of Orthopaedic Surgery and Research,* Volume 9.

Tang, A. C., Sutherland, M. T. & McKinney, C. J., 2005. Validation of SOBI components from high-density EEG. *NeuroImage,* 25(2), pp. 539-553.

Tang, Y.-Y. & Tang, R., 2020. *The Neuroscience of Meditation: Understanding Individual Differences.* s.l.:Academic Press.

Tanner, B. A., 2012. Validity of Global Physical and Emotional SUDS. *Applied Psychophysiology and Biofeedback,* 37(1), pp. 21-24.

Teymoori, A. et al., 2020. Measurement invariance of assessments of depression (PHQ-9) and anxiety (GAD-7) across sex, strata and linguistic backgrounds in a European-wide sample of patients after Traumatic Brain Injury. *Journal of Affective Disorders,* 262(1), pp. 278-285.

Thayer, J. F., Hansen, A. L., Saus-Rose, E. & Johnsen, B. H., 2009. Heart Rate Variability, Prefrontal Neural Function, and Cognitive Performance: The Neurovisceral Integration Perspective on Self-regulation, Adaptation, and Health. *Annals of Behavioral Medicine,* 37(2), p. 141–153.

Ulrich, R. S., 1981. Natural Versus Urban Scenes: Some Psychophysiological Effects. *Environment and Behaviour,* 13(5), pp. 523-556.

Valmaggia, L. R., Latif, L., Kempton, M. J. & Rus-Calafell, M., 2016. Virtual reality in the psychological treatment for mental health problems: An systematic review of recent evidence. *Psychiatry Research,* Volume 236, pp. 189-195.

Valve, 2022. *Steam Hardware & Software Survey: April 2022.* [Online]   
Available at: https://store.steampowered.com/hwsurvey/Steam-Hardware-Software-Survey-Welcome-to-Steam  
[Accessed 01 05 2022].

Villaroel, M. A. & Terlizzi, E. P., 2020. *Symptoms of Depression Among Adults: United States, 2019,* s.l.: National Centre for Health Statistics.

Vogt, T., Andre, E. & Bee, N., 2008. *EmoVoice — A Framework for Online Recognition of Emotions from Voice.* Berlin, Springer, p. 188–199.

Webb, C. A., Rosso, I. M. & Rauch, S. L., 2017. Internet-based cognitive-behavioral therapy for depression: Current progress and future directions. *Harvard Review of Psychiatry,* Volume 25, pp. 114-122.

Wiergraber, M., Papazoglou, A., Broich, K. & Muller, R., 2016. Sampling rate, signal bandwidth and related pitfalls in EEG analysis. *Journal of Neuroscience Methods,* Volume 268, pp. 53-55.

Willan, J., King, A. J., Jeffery, K. & Bienz, N., 2020. Challenges for NHS hospitals during covid-19 epidemic. *BMJ,* 386(8239), p. 368.

World Health Organisation, 2021. *Depression.* [Online]   
Available at: https://www.who.int/news-room/fact-sheets/detail/depression  
[Accessed 27 10 2021].

Wosik, J. et al., 2020. Telehealth transformation: COVID-19 and the rise of virtual care. *Journal of the American Medical Informatics Association,* 27(6), pp. 957-962.

Wundt, W., 1897. *Outline of psychology.* New York: Thoemmes Continuum press.

Yiannakopoulo, E., Nikiteas, N., Perrea, D. & Tsigris, C., 2015. Virtual reality simulators and training in laparoscopic surgery. *International Journal of Surgery,* Volume 13, pp. 60-65.

Yu, M. et al., 2022. EEG-based emotion recognition in an immersive virtual reality environment: From local activity to brain network features. *Biomedical Signal Processing and Control,* Volume 72.

Zhang, W. et al., 2020. Mental Health and Psychosocial Problems of Medical Health Workers during the COVID-19 Epidemic in China. *Psychotherapy and Psychosomatics,* 89(4), pp. 242-250.