UNIVERSITY OF TECHNOLOGY SYDNEY

The Effects of Physiological Stress on Brain-Computer Interface Systems

by Howe Yuan Zhu

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

under the supervision of Prof. Chin-Teng Lin and Dr. Hsiang-Ting (Tim) Chen

University of Technology Sydney
Faculty of Engineering and Information Technology

March 2022

Certificate of original authorship

I, *Howe Yuan Zhu* declare that this thesis, submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the *School of Computer Science*, *Faculty of Engineering and IT* at the University of Technology Sydney, Australia.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution. This research is supported by the Australian Government Research Training Program.

Production Note:
SIGNATURE: Signature removed prior to publication.

Howe Yuan Zhu

DATE: 17th March, 2022

PLACE: Sydney, Australia

Abstract

Brain-Computer Interface (BCI) devices are an emerging technology that aims to revolutionise the way humans interact with various contemporary technologies. A significant challenge to the practical use of BCI devices is the signal sensitivity to changes in a user's physical and emotional state. Factors such as muscle movement, electrical noise, workload, fatigue, and emotional state, can negatively contribute to the performance of a BCI device when being used in a real-world environment.

This work investigates how a user's physiological stress level may impact the performance of a BCI device on an Electroencephalograph(EEG) signal level. The human stress response is a universal survival mechanism that impacts both the emotional and physiological state of the body. While it is known that acute stress directly affects mental performance, its specific effects on the EEG signal behaviour and P300 response is mixed and sometimes contradictory.

We performed two novel experiments to mimic real-world scenarios that elicit a stress response. Our experiments incorporated complex visual input, auditory stimuli, and proprioceptive feedback into our experimental design to improve the future robustness of BCI system designs. The two experiments explored different types of stressors, with one being a prolonged stimuli stressor (height exposure) and the other being a dynamic stressor (unexpected drone collisions). The first experiment explores a novel elevated walking experiment that utilises a combination of physical and virtual height to induce a stress response. The second experiment investigated the potential use of unexpected drone collisions to elicit a stress response. Both experiments successfully induced a physiological stress response and produced an observable neurological change in the EEG signal.

Our results indicate that prolonged exposure (Height Exposure Experiment) to stressful stimuli creates a significant increase in frontal to parietal beta power and a lower P300 peak amplitude in some participants during a BCI task. On the other hand, a dynamic stressor (Drone Collision Exposure Experiment) tends to produce a short-term increase in frontal and central theta power, along with a negativity response during the stimuli. Collectively these findings provide insights into how different forms of physiological stress affects BCI devices on an EEG signal level and furthers the development towards a practical, real-world BCI device.

Dedication

To the Almighty God that create such amazing indecipherable mysteries for us to explore and grasp in our days. To my wife Louise and family who supported, guided, and shaped who I am today. . .

ACKNOWLEDGMENTS

First and foremost I want to thank my supervisors Prof. Chin-Teng Lin and Dr Hsiang-Ting (Tim) Chen. This research would not have been possible without their patience, guidance, and willingness to explore innovative experiments that other researchers may have considered too perilous or difficult. Thank you for the continuous support and for providing opportunities to collaborate with others to investigate various interesting research projects.

I would like to thank all my lab colleagues in Computational Intelligence and Brain-Computer Interface (CIBCI) group for the help, discussion and general camaraderie in keeping each other sane. Special thanks to Carlos, Thong, Avinash, Dr YK Wang, Xiaofei, Yanqiu, Alex, Jia, Sai, and Daniel.

I am grateful to our talented undergraduate team that I've had the pleasure to have individually worked with over the past years. Thank you to Ehsan, James, Samuel, Mena, Vinayak, Sanjid, Eirene, Omar, Mohammad, Vanessa, and Tim (Trung Le), for their dedication and passion in pursuing their capstones projects.

Thank you to the examiner panel members during my candidature assessments for providing valuable feedback and external perspectives on my research. Thank you to A/Prof. Wei Liu, Dr Benjamin Johnston, A/Prof. Ling Chen, Dr Jaime Garcia Marin, Dr William Raffe, and future assessors of this thesis.

I want to acknowledge the University of Technology of Sydney (UTS) School of Computer Science and UTS Techlab for the facilities and on-site staff support. Our experiments would not have been possible without the excellent lab space and the custom equipment that they installed on site. I am extremely grateful for the financial support provided by the Australian Government Research Training Program funded by the Australian Commonwealth Government, and the UTS FEIT top-up scholarship funded by the Defence Science and Technology Group of the Department of Defence. A special thanks to the SCADS team (Thomas, Mathew, and Jordan) for the research opportunity and for sharing their extensive knowledge and experience of drone systems.

Lastly, I appreciate my parents and brother (Jian Guo Zhu, Hong-Yuan Liu, and Jiang Zhu) for raising, guiding, support, and overall ever patient with me throughout my life. I would like to thank my loving wife, Louise, for her continuous support and love. I would also like to express my gratitude to my friends. This PhD journey would not have been possible without any of the people mentioned above.

List of Publications

Journal:

1. **H.Y. Zhu**, H.-T. Chen, and C.-T. Lin, *The effects of virtual and physical elevation on physiological stress during virtual reality height exposure*, (***under review, 2nd stage***).

2. **H.Y. Zhu**, H.-T. Chen, and C.-T. Lin, *The Effects of Height-Induced Stress on Human Brain Dynamics during a P300 Task*, (***Drafted***).

3. **H.Y. Zhu**, H.-T. Chen, E. M. Magsino, S. M. Hamim, and C.-T. Lin, *EEG evaluation of the Effects and Detection of Drone Collisions*, (***Drafted***).

Conference:

1. **H.Y. Zhu**, E. M. Magsino, S. M. Hamim, C.-T. Lin, and H.-T. Chen, *A drone nearly hit me! a reflection on the human factors of drone collisions*, Extended Abstracts of the 2021 ACM CHI Conference on Human Factors in Computing Systems (**SIGCHI**), Yokohama, Japan, 8-13 May, 2021.

2. **H.Y. Zhu**, H.-T. Chen, and C.-T. Lin, *The effects of a stressful physical environment during virtual reality height exposure*,2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (**VRW**), pp. 468-469, Lisbon, Portugal, 27 March-4 April, 2021.

Others :

1. Y. Tian and A. Minton, **H.Y. Zhu**, G. Notaro, R. Galvan, Y. -K. Wang, H. -T Chen , *A Comparison of Common Video Game versus Real-World Heads-Up-Display Designs for the Purpose of Target Localization and Identification.*, 2021 IEEE International Symposium on Mixed and Augmented Reality Poster (**ISMAR**), Bari, Italy, 4 - 8 Oct., 2021.

2. T. T. L. NGUYEN, **H.Y. ZHU**, H. -T CHEN, *Remote Visual Line-of-Sight: A Remote Platform for theVisualisation and Control of an Indoor Drone using Virtual Reality.*, 27th ACM Symposium on Virtual Reality Software and Technology Poster (**VRST**), Osaka, Japan, 8 - 10 Dec., 2021. **(Accepted)**

Table of Contents

List of Figures

List of Tables

CHAPTER 1

INTRODUCTION

1.1 Brain-Computer Interface Technology

1.1.1 Overview

Brain-Computer Interface (BCI) technology has dramatically increased in popularity over time, with large organisations such as DARPA investing significant funds towards BCI research [132]. Researchers from various multidisciplinary fields have investigated various applications of BCI systems in medicine, rehabilitation, neuroergonomics, security, military, gaming, and education [1]. BCI is a powerful communication tool that could allow humans to interface with various platforms and systems by translating the neurological signals in the brain into usable commands. A common type of BCI is the Electroencephalograph (EEG) based BCI system[204]. These are typically an array of external electrodes that measure the electrical potential on the wearer's scalp. The measurements from the scalp tend to reflect the electrophysiology of the areas of the brain. By applying signal processing techniques, statistical modelling, and machine learning (or deep learning) classifiers, one could translate well known neurological behaviour (such as P300) into a usable user interface, or measurement device [118].

The continued advancements in the BCI classifier have explored various possibilities and produced accurate measures of neurological behaviours [96]. Currently, the most extensive use for EEG devices is within the medical field, primarily for detection of brain activity within coma patients and rehabilitation for patients with spinal injuries [1].

There are advertised BCI devices on the consumer market. However, the performance of these devices tends to rely on spectral activity and make inaccurate (borderline pseudoscience) inferences on the user's mental state. In current times, the primary BCI market is divided into medical devices and research measurement tools [132].

1.1.2 BCI Challenges

From evaluating the current state of BCI research and applications, we have identified a critical barrier that hinders the real-world applications of BCI devices. The main challenge is the reliability of the BCI signal classification in a real-world environment [28, 103, 134]. A common criticism of BCI is that the devices are unreliable during real-world use. This unreliability often outweighs any benefits it may possess. A study by Carabalona et al. [28] investigated the use of BCI as a household light switch. Unsurprisingly, they found that while the BCI application was novel and interesting, the device's unreliability (80% detection accuracy) outweighed the benefits. We have many more advanced options for switching on the light, such as smartphones, sound-based, and automatic light, yet most households still rely on tactile mechanical switches. Kishore et al. [103] made a similar conclusion with traditional eye-tracking based controls outperforming a common BCI control technique (Steady-state visually evoked potential or SSVEP). We are confident that BCI technology will gradually improve in accuracy over time. Nevertheless, there is a need for a more reliable and robust BCI signal classifier.

The underlying challenge of BCI devices is the transference of an individualised system developed in a control laboratory environment to a real-world environment while maintaining a high level of signal classification accuracy. This technical challenge is the reason why consumer BCI devices are often seen as impractical or unreliable [58, 168]. It is far too often that a BCI system with high classification accuracy in a controlled laboratory environment would perform exceptionally poorly in a real-world environment. We have personally observed this issue in our BCI endeavours (outlined in Chapter 7); we can also observe this in previous works [28, 94, 105, 134]. To clarify, we are not dismissing laboratory-based research. There is a definite value in controlled environmental conditions to explore specific independent variables. The current state of BCI is at a position where we have the computational ability, prior knowledge (from past research), and the established tool available to contend with the real-world factors that hinder BCI performance. In order to move forward with BCI development, we must progress toward more realistic experimental designs that incorporate real-world factors in a controlled manner.

EEG based devices are susceptible to noise. This is due to the electrode sensitivity to minute electrical potential changes (μV) [4]. The established research methodology for exploratory analysis minimises noise by controlling the environmental factors to maximise the accuracy of the signal for classification. Traditionally experiments minimise noise through subjects maintaining a stationary seated or standing position with minimal background auditory noise, constant lighting, and consistent emotion behaviour (baseline to calm subjects) [193]. This rationale suggests that body movement, complex auditory noise, dynamic lighting, and changing emotional states are some real-world factors that the BCI system must contend with before being ready for general consumer use.

1.2 Importance of Physiological Stress

1.2.1 The Impact of Physiological Stress on Daily Life

For this thesis, we chose to explore the emotional factors, specifically physiological stress, that affect a BCI system's performance in the real world. This factor was chosen over the previous factors due to the lack of realistic environments for stress elicitation studies and mixed consensus on how stress affects BCI performance [100]. The other previously mentioned factors, such as the impact of body movement on the EEG signal, has been extensively documented with multiple mobile studies that incorporated complex auditory and visual systems [55, 72–74].

We define physiological stress as the cognitive and physical reaction that the body undergoes when the mind perceives a threat or danger. The three types of physiological stresses are chronic, episodic, and acute [127]. Chronic and episodic stress are considered long term (hours, days, weeks) stress responses that often ties with mental health disorders. The acute stress response is the immediate bodily response to a perceived threat. Acute physiological stress is particularly interesting because it is the most common type of stress and is universally experienced in daily life. It is a crucial evolutionary trait that has ensured our survival in dangerous situations [101]. Our physiological stress response is tuned to our perception of danger, and the stress response is a full-body reaction that starts from the autonomic nervous system of our brain [114]. The human's stress response largely dictates cognitive and physiological behaviour, whether it be a cognitive boost from the stresses of writing a thesis or heightened alertness from jump scare in a horror movie [25]. In moderation, stress is essential and often beneficial; however, when in excess, it can cause both mental and physical deterioration resulting in decreased performance and health [194].

1.2.2 Stress, Cognitive and Mental Performance

Past works have extensively investigated the relationship between stress and cognitive performance with the Yerkes-Dodson (YD) law coming into popularity [114, 194]. The YD law dictates that stress/arousal has an open downwards parabolic relation (see Figure 1.1) to cognitive performance [208]. A person will experience low performance at either extremity at lower (bored or drowsy) and higher (panic and intense anxiety) arousal levels while having an optimal performance at a middle arousal level [194]. Over time, researchers found that the difficulty and complexity affect the shape of the YD curve [21]. However, the generalised understanding of the principle still holds true [51]. This law implies that the modulation of a person's stress level can improve their cognitive performance. This would also suggest that at high and low-stress levels, BCI performance would likely change and deteriorate.

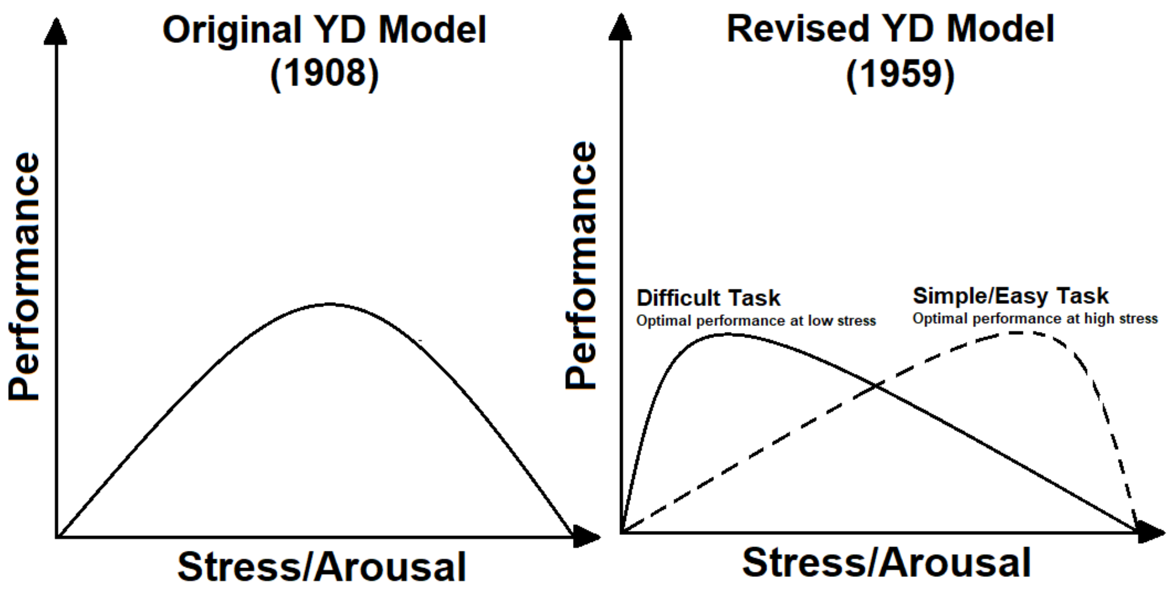

Figure 1.1: A diagram summarising traditional understanding of the YD Law [194]

1.3 The effects of Stress on BCI performance

1.3.1 Divergent Results in Prior Works

A better understanding of how stress affects BCI performance can provide insights into developing methods for adapting to the stress and improving signal classification. One standard method used to measure BCI and cognitive performance is measuring a person's P300 wave peak amplitude. Studies have shown that the peak amplitude can provide an accurate indication for changes in cognitive and behavioural performance [75, 153]. Previous studies have explored the relationship between stress and P300 amplitude; however, the results have varied (often contradicting) between studies. Certain studies have reported that stress can significantly decrease the P300 amplitude, this result was confirmed on multiple different paradigms such as oddball [32, 75, 100], Go/No-Go [53], Dot Probe task [97], and Flanker task [159].

On the other hand, other studies have found an opposing result with stress either increasing [54, 158] or producing no change [69] in P300 amplitude. The impact of this behaviour is quite significant as it would directly affect the accuracy of the P300 signal classifier. Thus, we need to investigate further and understand precisely how stress affects the P300 response and explain the conflicting reports of previous studies. Kamp et al. [100] speculated that the deviation in results was due to the varying effectiveness of different stress paradigms. This idea is plausible as the experience of stress can deviate between individuals, so each study may have elicited varying levels of stress which causes a different group-level P300 behaviour. Our research aims to explore further the effects of stress on the P300 amplitude by using height exposure as the stressor. Height exposure could produce a more consistent or potent level of the stress response, which may provide further insights into this investigation [147].

1.3.2 Stress Paradigms

Evaluating previous stress elicitation paradigms can provide insights into developing a novel yet, reliable paradigm. Most stress-related paradigms rely on the fear and threat response to induce a physiological stress response from the brain [179]. A majority of stress studies rely on glossophobia or social anxiety to induce stress [6, 24, 131, 178]. These paradigms tended to rely on interview blocks where the participant will need to perform an oral task (e.g., a speech or a song). Studies by Allen et al. and Brouwer et al. found that these social anxiety-based paradigms elicited a stress response effectively. While we agree that these paradigms are likely to induce stress, we question the reliability and the level of stress-induced. All the previously mentioned stress-related P300 studies utilised similar social anxiety based paradigms to elicit a stress response. Each study reports success stress elicitation but produces conflicting results. Two possible explanations are that either all the studies on one side of the argument are invalid (unlikely, given both conflicting sides have multiple studies with consistent results), or this paradigm does not consistently produce a stress response. Thus each study explores unique groups of individuals that variate between studies. Intuitively, the second reasoning seems more likely as social anxiety will vary between personalities and public speaking experiences.

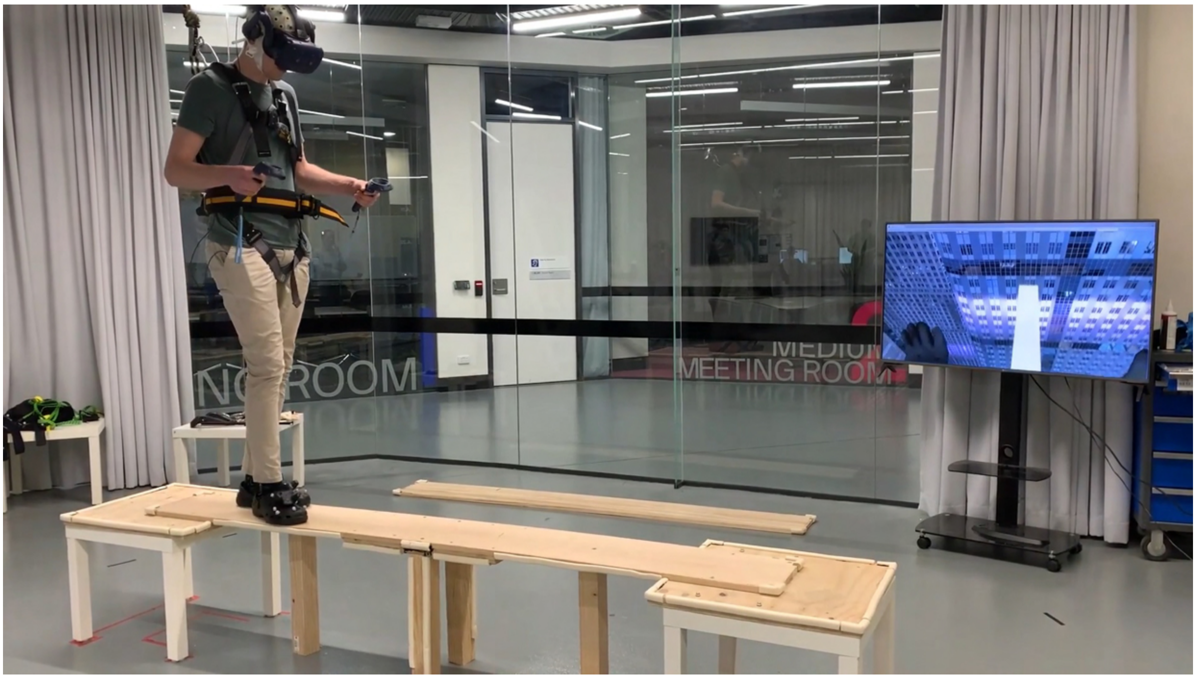

Figure 1.2: An early design of our stress exposure platform.

Height exposure is another paradigm that has, in recent times, gained popularity in stress-related studies [128, 147]. Height exposure relies on a participant's inherent fear of heights (acrophobia). Studies [16, 60] have found that acrophobia is demographically more consistent and potentially more reliable. Researchers often avoid height exposure and opt for social anxiety based paradigms due to the cost and complications of the experimental setup [60]. Social anxiety-based paradigms do not mainly require dedicated space or equipment compared to height exposure paradigms. This dynamic changed with the popularity and improvement of virtual reality (VR) technology. VR created the opportunity for height exposure at a much lower cost. A prominent study by Meehan et al. [128] found that VR height exposure can generate an observable physiological response indicating a stressed state. As time progressed, study [9, 50, 128, 147] found that incorporating physical haptic feedback can significantly increase the stress response (see Figure 1.2). From a traditional BCI research perspective, VR height exposure may incorporate too many noise factors. However, from the advancements of recent mobile VR BCI studies [55, 74, 184], we now have more robust data cleaning and analysis techniques that allow evaluation of the stress behaviours when the participant is walking and observing a complex environment.

Exploring more sudden or rapid threats is an alternate avenue for a stress exposure paradigm. This paradigm differs from the fear/phobia-based paradigm by relying on a reactionary and unexpected threat or danger. One example is the jump scare experiment by Browarska et al. [25], and Pallavincini et al. [145]. Their experiments involved horror movies and games. The participant would play the game, and a jump scare will occur in unexpected periods. Both studies found success in eliciting a stress response from these jump scares. Another example is the Lee et al. [117] which found success by using a driving simulator to simulate unexpected near accident or challenging driving events to produce a stress response.

The current level of understanding and the available EEG analysis tool enables us to explore more realistic and complex paradigms. Previous studies may have considered the EEG data from these experiments as too noisy to analyse [72]. The incorporation of auditory stimuli, complex visual environment, and proprioceptive input can produce a more realistic setting for the experiment and potentially increase the consistency of the stress response [128]. If our novel experiment designs successfully induce a stress response, then it would be beneficial to train and test future BCI systems in a more realistic experimental setup to improve the BCI system's robustness.

1.4 Research Questions, methodology, and Contributions

1.4.1 Research Questions

This thesis presents our works in our investigation into how physiological stress affects a BCI system. We performed two experiments that explored two different types of stressors (prolonged vs dynamic). These experiments were designed to include more real-world factors than previously performed stress paradigms. From the recorded EEG data, we drew further insights into the effects of stress on the EEG signal and potential adaptations to improve the future robustness of BCI systems. From assessing prior research and the significance of the potential findings, we formulated our research questions which dictated our research method and analysis.

The primary two research Questions (RQ) of this thesis are:

- ***RQ1***: Can more realistic stress elicitation experimental designs produce a consistent and reliable stress response?
- ***RQ2***: What are the effects of the stress states on a person's neurophysiological behaviour during common BCI tasks such as P300 or object tracking tasks?

1.4.1.1 Research Question 1

Traditional stress elicitation paradigms in neuroscience research tend to conform to using low-cost social anxiety-based paradigms where participants are stationary. These methods are safe, low cost, and easy to control. However, they lack the various dynamic real-world auditory, visual, and proprioceptive factors in a more realistic environment. Thus there is a gap in the research field between simplistic BCI research in a laboratory environment and practical BCI applications in the real world. In order to push forward the field of BCI research (and reduce the gap), we must innovate and build novel experiment designs that progressively incorporates and investigates these real-world factors. Hence, we build off the notable works of mobile and stress BCI Gramann et al. [74], Do et al. [55], and Peterson et al. [147], to investigate different novel and reliable methods of eliciting a stress response in a more realistic environment. We evaluated the types of paradigms available and coined our classification for two differing types of paradigms that can elicit an acute physiological stress response.

The first type of stress paradigms was named Prolonged Stimuli Stressor (PSS) type paradigms. PSS encompasses most of the traditional fear or phobia based paradigms. These paradigms rely on continuous or prolonged exposure to a stressor that plays on the participant's phobias. In the case of social anxiety paradigms, participants would undergo multiple 5-10 minute interview blocks in which they perform demanding oral tasks [6]. Height exposure is another form of a PSS paradigm. Participants would undergo walking or standing exposure to an actual or VR heightened environment for a period [128].

In contrast, we named the second type of stress paradigms as Dynamic Stimuli Stress (DSS). This type of paradigm elicits an instinctive response when subjects perceive an immediate and unexpected danger or threat. The primary exposure of these paradigms must present a real or sense of real incoming threat to the participant where the brain will need to trigger an immediate fight or flight response to avoid or mitigate the direct threat. Examples of this type of stress can be from the driving experiment by Lee et al. [117] with dangerous unexpected traffic situations, and the jump scares from Browarska et al. [25].

The PSS and DSS classification provides an overview of the two main facets of stress elicitation paradigms. We believe, based on previous works [139, 147], that PSS and DSS will produce differing stress responses due to differences in threat exposure period (continuous vs intermittent) and the type of threat (phobia-based vs sudden danger). We speculate that PSS paradigms can produce a constant stress response that is elevated at the start of exposure and then attenuates over time as the participant is continuously exposed to the stressor. On the other hand, participants performing DSS paradigms will likely experience a sudden and short burst of heightened physiological stress when the brain perceives danger. This burst of stress will likely subside much faster as the brain recognises that the stressor has passed and returns to a calm mental state.

We developed two novel experimental designs that investigated PSS (height exposure) and DSS (drone collision). Our first experiment is a PSS type paradigm that utilised height exposure to induce stress. Conventional height exposure would only use a VR headset to modulate the level of exposure, creating a stressed (extreme height) and non-stressed (ground) state. Our experiment furthers this by incorporating an elevated physical walking platform along with the VR height. The elevated physical platform provides a more reliable and consistent stress response through the increased threat level. The other function of the platform is to investigate if we can modulate differing stress levels by interchanging the physical and virtual height levels. We designed the physical and virtual environment to incorporate walking at heights, a complex virtual urban environment, and background noise stimuli. These factors improve the scenario's realism, and the stress response is likely to be more translatable to real-world stress responses.

The second experiment uses drone collision (and non-collisions) as a form of DSS stimuli. We reason that drones are inherently perceived as a danger due to the loud noise and visual threat of the fast rotating propellers. It is conceivable that an individual would perceive a fast-moving object with rotating blades on a collision trajectory as a threat. For safety reasons, participants did not experience any actual collision with the drone; instead, they stood closely behind a safety net with the simulated collision of the drone colliding into the safety net at the shoulder to face height. We correlate the threat and stress behaviour to be similar to an unexpected jump scare or driving accident. Overall, the drone threat provides a real threat that can be applied to future drone real-world drone interaction or collision avoidance responses.

1.4.1.2 Research Question 2

RQ2 specifically address the gaps or conflicting knowledge in the present literature. As previously outlined, the research community is divided over the exact effects of stress on P300 and BCI on a signal level. Previous studies used similar methodologies yet finding different results among various participant groups. Therefore, using the stress experiments developed for the first research question, we hope to elicit a stress response during a P300-related task study of the effects on the P300 wave. The first experiment (height exposure) aims to directly investigate the conflicting reports by comparing the participants P300 behaviour (through an Oddball task) during the stressed (height) and non-stressed (ground) conditions. We assessed the EEG data collected from the participants and evaluated the metrics commonly associated to the P300 behaviour to determine whether stress increases [54, 158], decreases [32, 75, 100], or does not affect [69] P300 behaviour. The regions of interest are the frontal to parietal areas in the brain, and we specifically investigate the Event-Related Potential (ERP) and Event-Related Spectral Perturbation (ERSP). The insights from this experiment can potentially support or reconcile the conflicts found in previous literature.

The drone experiment does not have direct prior iterations in the literature that we can use as a reference for the experimental method and results. Our analysis will be based on previous studies related to neurophysiological behaviour when tracking ball/object trajectory [138, 139]. We rationale that this is similar to tracking a drone during drone collisions. We plan to follow this exploratory methodology and assess the ERP and ERSP responses. We theorise that the DSS stimuli would produce a faster and shorter stress response that is more instinctive and involuntary than PSS. The detection and classification of this response could impact the design methodology of drone BCI control systems. The findings of the EEG analysis can help mitigate or improve drone behaviour by detecting when it may be on a potential collision trajectory.

1.4.2 Qualitative and Quantitative Measurements

In both experiments, we assess two metrics: the stress level during the experiment and the impact of the stressful condition on the participant's cognitive state. The stress levels are measured through both qualitative responses and quantitative physiological data. We collected self-reported questionnaire responses [20, 121] as well as physiological measures such as heart rate (HR), heart rate variability (HRV) [192], and electrodermal activity (EDA) [35, 52]. Past research has shown the effectiveness of these metrics in measuring changes in physiological stress response.

We evaluated the cognitive state through the recorded EEG data of the participants. The participants wore a gel-based 64 channelled active electrode EEG cap. For the height exposure experiment, we performed a hypothesis-driven analysis as prior works [174] have established the reliability of analysing the midline channels for a P300 response. We would assess the P300 signal on the ERP and ERSP metrics. The drone experiment required an exploratory EEG analysis approach as there were no clear prior findings from the literature. Using the exploratory analysis methodology from Muraskin et al. [139], Singh et al. [184], and Do et al. [55]. We evaluated the EEG data for significant sources during the drone trajectory tracking and observed if the drone collision generated a cognitive conflict response within-participant [139].

1.4.3 Findings and Contributions

In our first experiment (height exposure), we found a significant increase in physiological (HR and EDA) measures, gait behaviour, and questionnaire results when the participant's sense of elevation increases (Chapter 3). This finding is consistent with previous literature [128] suggesting that the elevated physical and virtual platform does successfully induces physiological stress. Incorporating the physical elevation is a novel part of this experiment and provides a more threatening and realistic haptic feedback for the participants. We believe this experimental design can provide a functional testing platform for future BCI research.

The EEG results of the first experiment (Chapter 4) aimed to shed light on the conflicting results of prior results. We found two significantly different groups with our participants. After a median split using the participant's reaction time, we found one group (Group 1) exhibited behaviour consistent with the studies that found a decrease in P300 peak amplitude when stressed. The other group (Group 2) was consistent with the other studies that argue stress increases or does not change P300 peak amplitude when stressed. This divergent result is interesting because it provides a potential explanation that reconciles the contradictory results of prior studies. By assessing the differences between the two groups, we found that Group 1 increased frontal to parietal beta power observed in Group 2. Group 2 virtually had no change in cognitive state from stress, yet, the other metric shows that they still experienced elevated stress levels. Based on the facts presented and our understanding of the YD law, we reason that Group 1 may have found more difficulty completing the oddball task while under stress. This is why their P300 peak amplitude decreased, and they exhibited an increase in beta power. Group 2 did not experience an increase in difficulty when under stress. Therefore the stress increase did not significantly alter their behaviour. This significant finding suggests that the effects of stress should not be universally applied. Instead, we should measure the individual's perceived difficulty of a task, inferring how their current stress level may affect their cognitive performance.

The drone experiment presented a unique and unexplored DSS through drone collisions (Chapter 5). From our questionnaire results (SAM, DASS, and threat ratings), we found that the participants reported higher levels of stress during collision compared to non-collision events. On the other hand, we did not detect any physiological change. These two results seemingly report opposing facts. However, we reason that the lack of change in physiological data may indicate the flaws of the physiological measurement devices (and potentially the measures themselves). The DSS of the drone collision is a considerably fast stimulus that likely produced a sudden and transient stress response. We believe that the physiological measures were not sensitive enough to perceive such a transient response. Therefore, future tests involving DSS-related paradigms should consider using higher fidelity (higher sampling rate and better accuracy) measurement devices or a metric with a more extensive temporal resolution such as EEG.

Our EEG results for the drone experiment are fascinating when comparing the collision and non-collision behaviours Chapter 6. We found two notable IC clusters that indicate frontal to central region activity. We found a significant signal activation that is time lock between the drone at 1m and collision from testing various epoch lengths. The ERP response indicates a cognitive conflict response. This result is corroborated by the ERSP's increase in theta activity. This finding would suggest that drone collisions are an incongruent event in comparison to the non-collision event. Therefore, it is apparent that drone collision can produce a predictable behaviour within the EEG signal, creating the possibility for future applications of BCI interfaces to moderate certain drone behaviours.

1.5 Chapter Organisation

Figure 1.3 illustrate the chapter organisation of this thesis and summarises the key points and findings. This thesis is divided into eight chapters that detail the introduction, methodology, results, and future research applications.

- Chapter 1 entitled *Introduction* outline the details of the background of BCI devices which leads to our research goals on how physiological stress affects BCI performance.
- Chapter 2 entitled *Literature Review* provides definitions and a summary of prior research concerning BCI, physiological stress, and specific research relating to both concepts.
- Chapter 3 entitled *Height Exposure Experiment* presents our VR height stress experiment. In this chapter, we detail the background of similar studies, the methodology of the experiment, and the physiological results. These results provide insight into the efficacy of the experiment in stress elicitation.
- Chapter 4 entitled *Effects of Stress on Cognitive Performance during Heights Exposure* further investigates and presents the EEG results of the height exposure experiment. The results specifically addressed the relationship between stress and the P300 response.
- Chapter 5 entitled *Drone Collision Exposure Experiment* presents the details of our drone collision experiment. Here, we outline the background, motivations, methodology, and behavioural and physiological results. We also compare these results to the first experiment to explore the difference between a PSS and DSS paradigm.
- Chapter 6 entitled *Effects of Drone Related Stress on Brain Dynamics* describes and presents our exploratory EEG analysis results that compare the drone collision and non-collision events in the experiment. Here we present the ERP and ERSP results and discuss the implication of these results for future drone interactions.
- Chapter 7 entitled *Exploratory and Future Works* outlines some of the exploratory prototypes that we develop with undergraduate capstone students. Our future works discuss the planned applications that would integrate our exploratory applications with our findings from the stress experiments.

- Chapter 8 entitled *Conclusions* concludes the thesis by summarising the key findings and assessing if the research questions have been achieved.

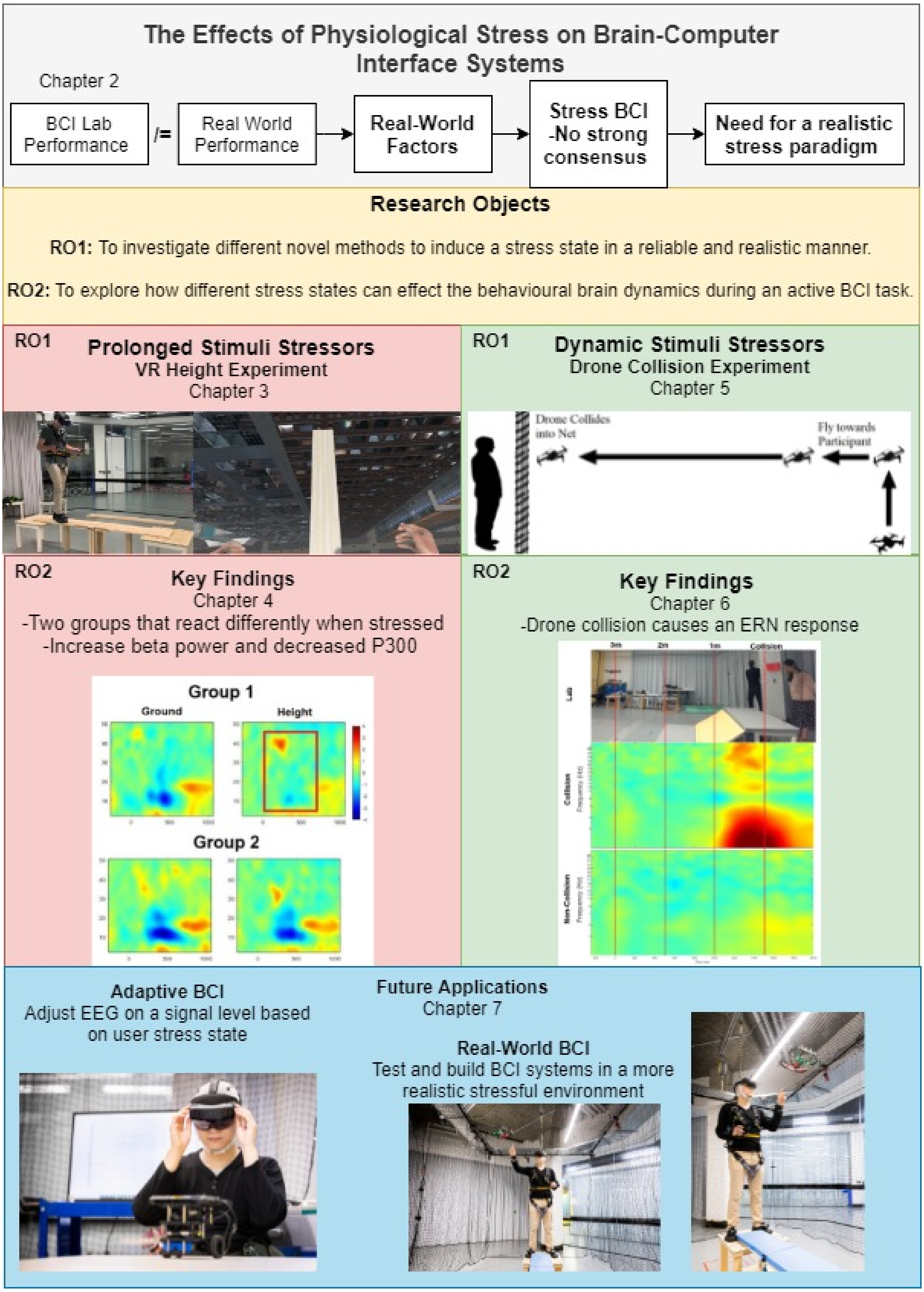

Figure 1.3: A diagram summarising the chapter organisation of this thesis.

CHAPTER 2

Literature Review

2.1 Physiological Stress

2.1.1 Definition

Physiological stress is a complex physiological response that occurs when the brain perceives a threat and triggers a response. Stress is directly related to a human's natural fight, flight, or freeze response [34]. Stress can be both a positive and negative response, and it is a fundamental survival mechanism for humans [115]. At certain stress levels, it can result in alertness, productivity, and focus. However, at higher levels, stress can cause symptoms such as emotional distress, headache, organ system failure (commonly digestive and cardiac systems), mental health issues (hypertension, depression and anxiety), and complete mental breakdowns [115]. Stress can be defined into three categories with varying definitions and symptoms as outlined in Table 2.1 [127]. This thesis will primarily focus on the experience of acute stress as it is a normal response that occurs within a healthy population. The response can be studied in a case study manner through stress elicitation paradigms. Episodic acute stress and chronic stress have more overt pathological factors involved, requiring extensive participant screening and a careful therapy-based and longitudinal approach for studying the stress response. Stressors refer to any form of stimuli that could cause the secretion of stress hormones (part of the stress response) [114]. Stress elicitation paradigms often rely on a reliable stressor (universal across different demographics) that can produce an acute physiological stress response.

Table 2.1: A table outlining the different types of stress [127]

Type	Definition
Acute Stress	Short term (seconds to hours) Reactive to external stimuli Low risk of physical and psychological issues
Episodic Acute Stress	Frequent acute levels of stress at specific periods Connected to lifestyle and mental stimuli Symptomatic for other mental health disorders
Chronic Stress	Experienced very frequently and continuously Connected to other mental health disorders Can be debilitating for a person,Äôs health and lifestyle.

In this thesis, we have coined the terms Prolonged Stimuli Stressor (PSS) and Dynamic Stimuli Stressor (DSS) to better differentiate the difference between our drone collision stressor (DSS) and our height exposure stressor (PSS). We consider both PSS and DSS as forms of stressors that elicit an acute physiological stress response. The difference is the exposure duration and the length of the stress response. PSS are stressors that participants would be continuously exposed to the stressor over a duration of time (seconds to minutes). PSS can be from situations such as being at heights or public speaking. DSS are stressors that occur from sudden and unexpected situations such as accidents, dangerous situations (projectiles), or sudden surprises. Figure 2.1 summarises the definition of physiological stress in this thesis and specifies the relation of PSS and DSS paradigm within the acute stress category.

2.1.2 Demographical Factors

Stress is a universal experienced; however, the perception of stress and coping mechanisms are subjective. Many factors such as life experience, personality, culture, community, lifestyle, and beliefs may affect how one perceives and copes with stress. Past research has identified three key factors: age, gender, and fears that significantly affect the stress response behaviour. Stress is both a neurological (brain) and endocrinological (hormonal) response; likewise, stress can be defined both emotionally and physiologically [114]. A study by Scott et al. [173] compared two sample groups with different age groups (20-80). The study used the PANA (positive affect-negative affect) model. The study found that older adults experience fewer stressors for both PA (positive affect) and NA (negative affect), while young adults experienced more stressors and an increased variability between PA and NA stressors. Suggesting that age is not directly related to the negative or positive effects of stress; instead, age is related to the susceptibility and variability

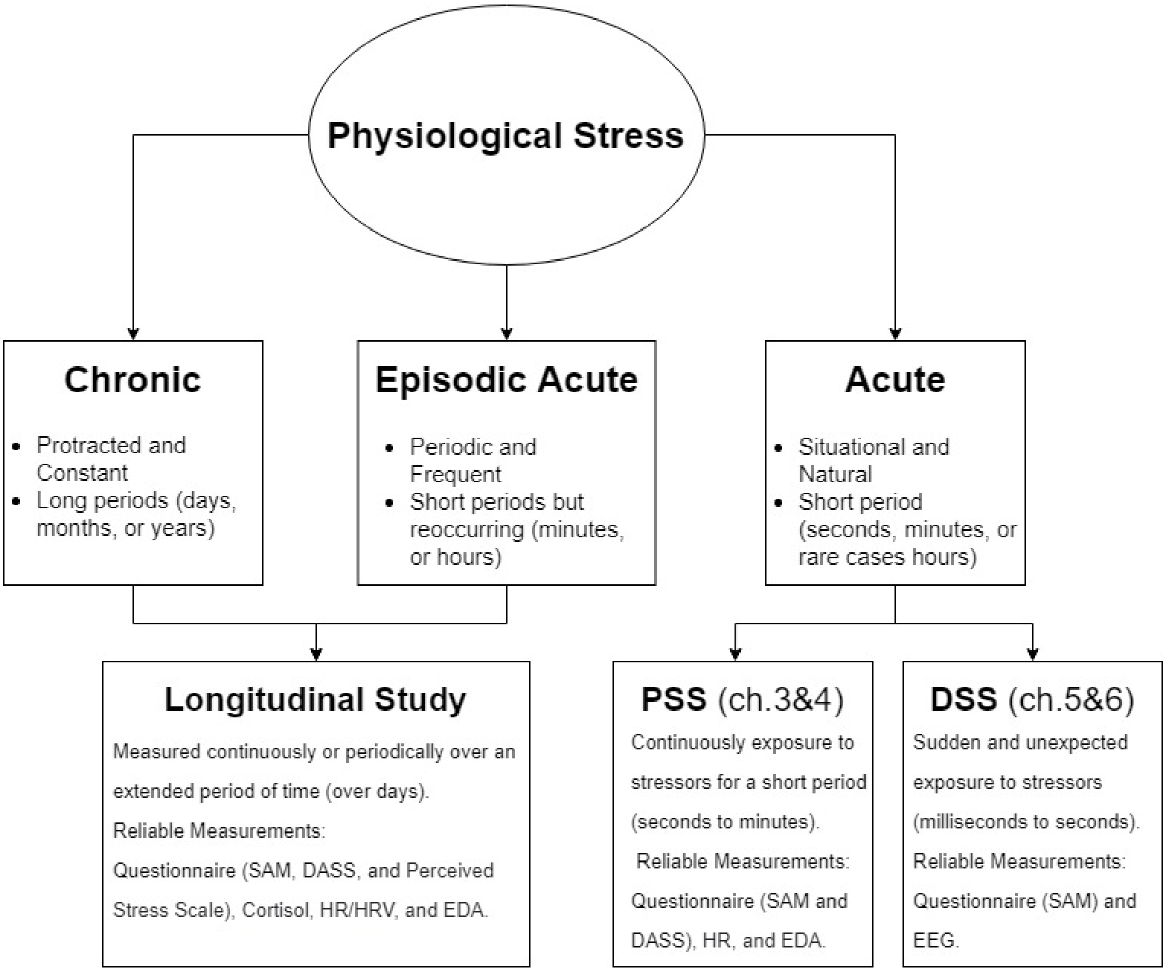

Figure 2.1: An early design of our stress exposure platform.

of the experiencing of stress. A study by Lupien et al. investigated hormonal effects of stress on both animal groups and humans (PTSD and depression) for a broader age group (prenatal, postnatal, adolescence, adult, and aging) [122]. Similar to the previous study, this study found that Glucocorticoids (a hormone secreted during stress) increased during adolescence while decreasing through the adult and aging life stages. Prenatal and postnatal stages were also lesser than the adolescence stage. However, this may be due to the underdeveloped endocrine system as it is before puberty.

The biological sex of a person will affect their susceptibility and reaction to stressors. There is an anatomical and physiological difference between a male and female [198]. Verma et al. [198] concluded that the difference is the existence of the menstrual cycle. The study found that during adolescence, when the menstrual cycle develops, the sex hormone attenuates and the responsiveness of the Hypothalamic-pituitary-adrenal

(HPA) axis increases. The HPA gland determines the rate of hormone secretion (such as Glucocorticoids). This effect results in a heightened response to stressors for females and a more extended downregulation of the HPA axis. Fear is a factor is having been traditionally considered separate from the stress response. However, recent studies are now using neurocircuitry and neurochemistry to correlate stress as an onset response to fear [179]. Shin et al. [179] explore the similarities between the reactions to fear and stress response. Both responses trigger an HPA axis reaction that results in the secretions of stress hormones within the animal population. The author also correlates the similar effect of mild vs severe exposure of stress and fear. These findings agreed with the results of Johnson et al. [99]. However, the author also comments on the similarity between the fear neurocircuitry response and the stress response, specifically the sympathetic nervous and limbic systems (amygdala output circuits).

2.1.3 Modelling Stress

In modern psychology, stress is classified on an emotional spectrum. To correctly classify stress as an emotion, one must first agree on the model to use. There are two prevailing theories in psychology on how emotions should be classified. One approach is the Discrete Emotion Theory suggested by psychologist Silvan Tomkins [196]. This theory suggests eight distinct basic emotions or "affects" that a human possesses, surprise, interest, joy, rage, fear, disgust, shame, and anguish. Tomkins adds that all emotional responses exhibited by a human must comply and exist under these eight categories. An opponent of this theory is James A. Russell, who in turn proposes a dimensional emotion model [166]. This model states that emotions exist on a spectrum with parings such as arousal-deactivation (active vs sleepiness) and positive-negative valance (pleasure vs misery). This theory is widely accepted in the psychology field. However, there are many different models proposed in this field, e.g. PANA, Plutchik's model, and Cowen and Keltner [13]. The widely adopted model is the circumplex model, proposed by James A. Russell. Figure 2.2 depicts the arousal-valence spectrum range of emotions with stress classified explicitly under the heightened arousal, and negative valence level [166].

2.1.4 Physiological and Behaviour Effects and Detection

The stress response is a complex process that involves both the neurological pathways and the endocrine system [115]. The full-body action of the stress response allows for many sites of observation that can detect and measure a person's stress level [37]. The stress

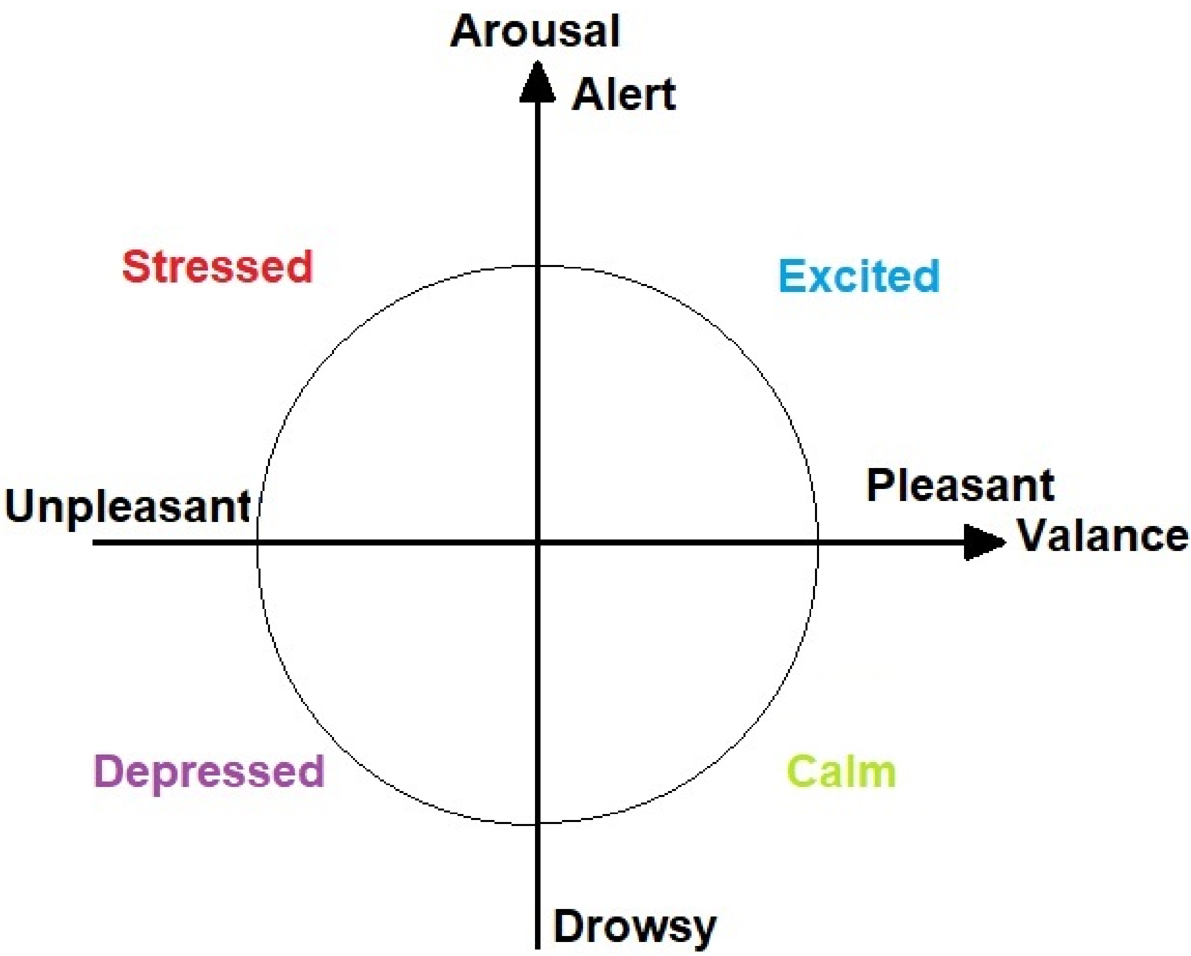

Figure 2.2: The circumplex model proposed by James A. Russell [166]

response begins with the perception of a stimulus that triggers the Amygdala (limbic system) response, which will then excite the hypothalamus to trigger the sympathetic nervous system. The HPA axis will signal a release of hormones to react as part of the stress response. Hormones such as Glucocorticoids are release which causes an excitation of the body. Resulting in heightened breathing, heart rate, blood pressure, and perspiration [37].

The physiological process of the stress response presents multiple markers [8], i.e. that can be measured and used to indicate if a person is stressed or not. Markers such as:

- $\uparrow$ heart rate/heart rate variability (HR/HRV) [192],
- $\uparrow$ blood pressure (BP)[200],

- ↑ skin perspiration [35, 52],
- ↑ respiratory rate (RR) [172],
- ↑ body temperature [37],
- ↓ step size and cadence (gait behaviour) [26, 170],
- ↑ conscious awareness measured by questionnaires such as Self Assessment Manikin (SAM) [20] and Depression Anxiety Stress Scale (DASS) [121], and
- ↑ cortisol levels [82].

Glucocorticoid is a hormone class of cortisol. Cortisol can typically be tested by acquiring a saliva sample from a participant and examining the sample in a medical lab. Herman's study [86] tested the cortisol level of a group of chronic stress participants vs a control group. The study found that glucocorticoid levels are a reliable metric for a participant's stress level. A prominent finding is a parabolic relation between glucocorticoid levels (stress) and a person's cognitive performance, which is known as the Yerkes-Dodson (YD) law [194]. The YD law suggests the existence of an optimum stress/arousal level that will provide peak performance with either low (drowsy or bored) and high (anxiety or panic) arousal levels resulting in reduced performance. Overall, cortisol is an accurate and reliable metric for measuring stress. However, the portability of the test is a significant limitation due to the need for salivary sampling and lab testing.

HR, HRV, and Galvanic Skin Response (GSR) or Endodermal Activity (EDA) are other reliable metrics for stress; it is far more portable (wearable wristband or chest strap) and provides relatively reliable data. Notable studies such as De Santos Sierra et al. [180] investigated the detection of stress using Electrocardiogram (ECG) to measure heart rate and EDA and found that HR/HRV and EDA measures were accurate in the detection of stress responses with heightened readings during the intended stress event. Sun et al. [190] also found similar results with chronic stress patients. Other metrics such as BP and RR (along with HR) were also investigated by Healey et al. [81], finding high accuracy in the detection of stress. However, HR was the most accurate measurement.

2.1.5 Stress Inducing Paradigms

Stress is a topic that researchers in the past deeply explore. Many paradigms exist for the specific purpose of inducing acute stress within healthy participants for clinical research. A reliable stress paradigm should elicit stress consistently across various demographics and in a controlled amount. Based on the literature, there are three common types of paradigms used for eliciting a stress response:

- Glossophobia (fear of public speaking) or Social phobias,

  -Trier Social Stress Test (TSST) [6]

  -Socially Evaluated Cold Pressor Task (SECPT) [131]

  -Sing-a-Song Stress Test (SSST) [24]

  -Maastricht Acute Stress Test (MAST) [178]

- Reactionary tasks (sometimes with monetary loss), and

  -Stroop Task [161]

  -Mental Arthemetic Task [159]

  -Paced Auditory Serial Addition Test (PASAT) [32]

- Acrophobia (fear of heights)

  -In-Vivo height exposure [16, 60]

  -Self Direct (images, videos, or meditative imagination) [12]

  -Virtual Reality (VR) Height [128, 147]

Past neuroscience has tended to favour glossophobia and reactionary tasks as stressors due to the stationary and straightforward nature of the tasks. Stationary experiment carries the benefits of being low cost and reduced noise from muscle movement, visual, and auditory stimuli [106]. However, these more controlled paradigms often lack the real-world factors crucial for the practical implementation of BCI systems. Factors such as complex visual (landscapes with multiple stimuli) and vestibular (background and multi-factored noise), and proprioceptive feedback (walking and muscle sense) are lost. Based on our research questions (***RQ1*** and ***RQ2***), we chose to use acrophobia as our primary stressor. An acrophobia based experiment will allow us to incorporate more real-world factors into a more natural stress response.

2.1.6 Prolonged and Dynamic Stress Exposure Paradigms

Table 2.2 outlines a summary of prevalent stress elicitation paradigms that we have categorised as either PSS or DSS paradigms. We define PSS paradigms as a paradigm that places participants into an environment or situation where they are continually

Table 2.2: A table outlining previous PSS and DSS paradigms

Paradigm	Exposure Period	Task	Stress Source
PSS			
TSST [6]	5 min	Speaking	Glossophobia
SECPT [131]	5 min	Speaking	Social Anxiety
SSST [24]	6.5 min	Singing	Social Anxiety
MAST [178]	10 min	Speaking	Glossophobia
Stroop Task [161]	8 min	Reactionary	Task Failure
Mental Arthemetic Task [159]	10 min	Reactionary	Task Failure
PASAT [32]	1-6 min	Reactionary	Task Failure
Height Exposure [61, 128, 147]	7 s (Chapter 3)	Walking	Acrophobia
DSS			
Horror [25, 145]	5-15 s	Jump scare	Horror
Driving [81]	1 s intervals	Driving Events	Danger
Drone Collision [217]	472 ms±87	Drone Collision	Danger

exposed to a stressor. We classify most traditional glossophobia, social anxiety, and height-related paradigms into the PSS category. For example, our height exposure experiment involves participants being continually exposed to the virtual and physical elevation, eliciting a stress response. Another example would be social anxiety based paradigms [6, 24, 131, 178] which depends on continual social anxiety from the multitude of questions or tasks. Based on the results of prior studies, we suspect that PSS paradigms produce a more sustained stress response due to the continual exposure to the stressor stimuli.

We classify the DSS paradigms as paradigms that utilise unexpected or rapidly appearing threat stimuli short and sudden. In the case of our drone collision experiment, the stressor is the drone abruptly colliding into the participant (into the safety net). We categorise Stress-related driving tasks [81, 117, 125, 160] studies as a sort of DSS. The Matthews et al. [125], and Lee et al. [117] driver simulator paradigm involved continual driving with sudden unexpected negative traffic situations. As expected, these studies found that negative or dangerous traffic situations could elicit a stress response from participants. Although, some of these studies have also suggested that continual driving itself already elevates a person's stress levels [81, 160]. Another example of DSS paradigms is the horror jump scare stress in studies by Browarska et al. [25] and Pallavincini et al. [145]. These studies found an increase in stress levels during jump scares from horror games and videos. These types (driving and horror) of paradigms follow the two critical criteria for a DSS paradigm. Firstly, the stimuli should be unanticipated through the incorporation of non-stressful or neutral portions. In the case of the driving task, the neutral event is the continual driving, and the horror film/game would be the non-horror portions. The second is that the stimuli must provide a sense of threat or play on a real-world related threat (crashing a car or being attacked). It would be unreasonable for the participant not to know the type of stressor since they are likely informed of the research paradigm during recruitment. Thus, the stressor would need to present a sufficient enough threat to elicit a stress response.

2.2 Brain Computer Interface

2.2.1 Definition and Types

A broad definition of BCI is any device that survey, analyse, and interpret the brain signal, then relays information to a machine to carry out a response [38]. The story of BCI begins with a discovery made by the British physicist Richard Canton [148]. In 1875, Canton discovered the existence of electrical signals in the brain of animals. This discovery paved the way for the pursuit of electrically mapping the brain and a better understanding of human neurophysiology. It was only four decades later, a psychiatrist named Han Berger invented the first EEG device, allowing humans to measure the brain's electrical activity for the first time [80]. Berger created the tool and discovered the first neural oscillation frequency, the 8-12 Hz Berger (Alpha) wave. In the present time, BCI has made its way into a wide range of research applications, e.g. military, rehabilitation (Figure 5), gaming, medicine, mental health, robotics and automation, and public services [124]. There are two primary types of BCI devices, invasive and non-invasive [201].

2.2.1.1 Invasive

Invasive or implantable BCIs are BCI systems that are fully or partially implanted into a subject skull [201]. Invasive BCIs is optimal for applications where the subject is indefinitely using the device [1]. The implanted EEG electrodes will bypass the scalp and skull layer and be placed directly in the observed effect's neural area. Collinger et al. [38] found success in improving the performance and usage of a prosthetic through the use of an invasive BCI. The study involved 13 weeks of BCI prosthetic training for a 52-year-old woman with tetraplegia. The participant received two 96 channel microelectrode arrays implants in the brain's motor cortex (M1) region. After the 13 weeks of training, the participant could skillfully control the prosthetic, scoring 91.6% in the training grab and reach tasks. Studies such as Pasquina et al. [146] and Zaaimi et al. [212] found similar success in implanting BCI electrodes.

2.2.1.2 Non-Invasive

There are multiple types of non-invasive BCI devices, and this work will primarily focus on externally worn EEG electrodes [193]. EEG systems support a range of electrode densities (number of electrodes per cap) is worn on the scalp. The placement on the scalp electrode is crucial to the region of the brain that is intended to measure. There are two main designs of non-invasive EEG devices, wet (gel-based) sensors and dry sensors (no gel required) [204]. A study by Guger et al. [79] compared the performance of wet vs dry sensors during a p300 speller task. The results showed that wet sensors perform better than dry sensors; however, dry sensors benefit from convenience and portability.

Implanted electrodes are far more effective compared to wearable electrodes due to the lack of skin and skull impedance between the brain and scalp [201]. Wet sensors can reduce this impedance to a certain degree through the conductive gel. However, the sensitivity to noise is a certainty. The scarcity of invasive BCI research is for two main reasons, the brain's regional coverage of non-invasive devices and the demographic of implantable devices. While invasive devices are accurate, they can only cover a specific region of the brain. It is highly impractical to implant a large electrode array that can match the coverage of non-invasive electrodes [201]. The other key consideration is the state of implantable technology, which has not reached the stage of maturity for healthy individuals to receive implantable BCI devices [111]. This limits the demographics of researched individuals to a rare and particular quality of life circumstances where a person may receive an implanted device.

Non-invasive BCI devices, though less accurate, can cover a larger area of the brain using higher density electrode arrays which is incredibly beneficial to exploratory neuroscience research. Invasive BCI requires extensive participant screening based on medical condition, age, and informed consent [132]; on the other hand, non-invasive BCI devices are far more universal with only factors such as gel allergies and experimental exclusion conditions [111].

2.2.2 P300

The P300 wave is a commonly chosen marker for BCI research. The P300 or P3 wave is an Event-related potential (ERP) component that is identified by positivity in the parietocentral scalp region [150]. This paradigm is highly reliable, with flexible stimuli and a distinct signal peak [39]. The P300 response is a unique signal peak that occurs after(300ms) a target stimuli has occurred; this is observed as a distinct peak in signal amplitude or power around the 300-500ms mark; an example of a P300 ERP response is seen in Figure 2.3. The P300 amplitude can provide an accurate indication for changes in cognitive and behavioural performance [75, 153]. The principle of this signal is based on the brain's ability to recognise stimuli through sensory observation of a specific target and generate a cognitive response to that stimuli [174]. This response can be evident in cases such as recognising a face in a crowd; the minimum response time is roughly 300ms for a person to react to a known face [174]. P300 related paradigms usually involve having multiple non-target stimuli (visual or auditory) with additional unique target stimuli (unique image or distinct sound frequency). These stimuli are observed sequentially, either in series or a matrix. A notable study is the Cattan et al. [29] study, which explores the use of VR to display a P300 paradigm. The study concluded that the combination of P300 and VR is feasible; however, the P300 stimuli must be distinctly noticeable and dominant within the paradigm. Common reliable P300-related paradigms include Oddball [49, 68, 154], Flanker task [107, 155], Go/NoGo [65, 98], and BCI Speller [109, 110].

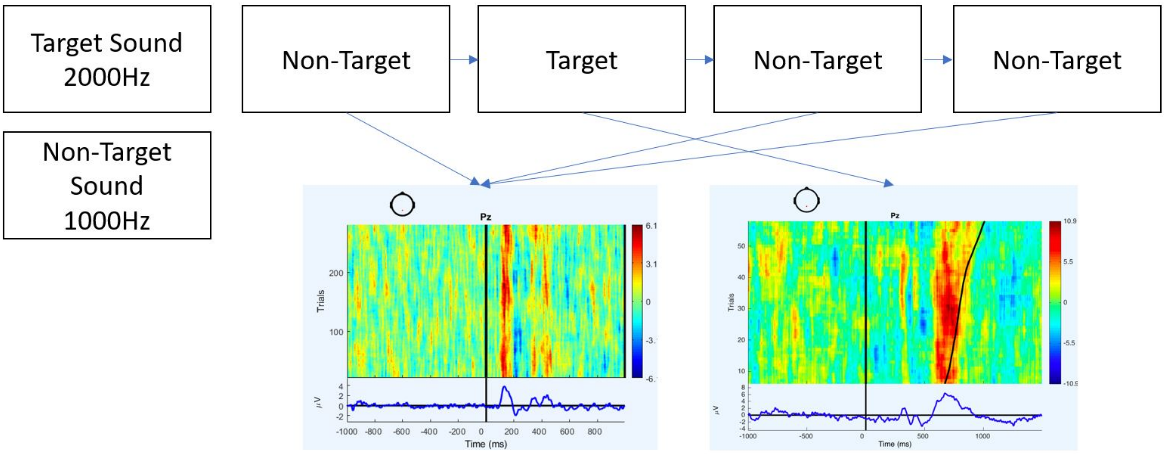

Figure 2.3: An example of an auditory P300 task with the ERP response from the parietal area

2.3 Stress and BCI

2.3.1 Detection

The basis of BCI research is to draw a correlation between an observed effect and the brain's neurophysiological effect. It can be considered a related effect when a neuro-

Table 2.3: A table outlining the key findings of various studies on BCI Stress detection

Ref	Brain Region	Valence	Arousal
Bos (2006) [17]	Frontal lobe	Asymmetry Alpha power	↑ Beta/Alpha power ratio
Reuderink et al. (2013) [162]	Frontal lobe	Asymmetry Alpha power	↑ Beta power ↓ Alpha power
Giannakakis et al. (2015) [70]	Frontal lobe	Asymmetry Activation	↓ Beta power ↓ Alpha power

Table 2.4: A table outlining the varying results of different studies, TSST- Trier Social Stress Test, PASAT- Paced Auditory Serial Addition Test, SECPT- Socially Evaluated Cold Pressor Task

Ref	Stressor task	P300 Task	Sample Size	Stress Effect
Kamp et al. (2021) [100]	TSST	Oddball	63	↓ P300 Amplitude
Ceballos et al. (2012) [32]	PASAT	Oddball	75	↓ P300 Amplitude
Granovsky et al. (1998) [75]	Word Triggers	Oddball	14	↓ P300 Amplitude
Dierolf et al. (2017) [53]	SECPT	Go/NoGo	39	↓ P300 Amplitude
Jiang et al. (2017) [97]	TSST	Dot Probe Task	62	↓ P300 Amplitude
Mingming et al. (2018) [159]	Mental Arithmetic Task	Flanker Task	17	↓ P300 Amplitude
Dierolf et al. (2018) [54]	TSST	Go/NoGo	49	↑ P300 Amplitude
Mingming et al. (2020) [158]	Mental Arithmetic Task	Flanker Task	20	↑ P300 Amplitude
Garcia et al. (2019) [69]	Stroop Task	BCI Speller	7	- P300 Amplitude

physiological behaviour demonstrates sufficient statistical power to a specific effect. To successfully detect stress using BCI, one must first identify the brain's corresponding neurobehaviour traits or biomarkers during the stress response. As previously stated, in the circumplex model of the emotional, stress can be classified as a heightened level of arousal with a negative valence [166]. Table 2.3, lists of literature on emotion classification with a focus on markers for arousal and valence. These studies utilised non-invasive EEG devices and stress elicitation paradigms to detect and classify the stress response. The table indicates that the alpha and beta power bands play a significant role in the frontal lobe for arousal and valence.

2.3.2 Cognitive (P300) Performance

The P300 behaviour can be a valuable metric for a user's cognitive performance, and attention levels [153]. Previous studies have found that the stress response can cause a significant change in cognitive performance and P300 ERP behaviour. However, the results seem to vary (sometimes contradicting) between studies.

As observed in Table 2.4, the effect of stress seems to vary significantly between studies with no apparent common factor as they share similar stressor tasks and P300 tasks. The outlined studies all recruited relatively similar populations with the same researchers (Dierolf et al. [53, 54], and Mingming et al. [158, 159]) seemingly repeating studies to find contradicting results. Kamp et al. [100] made a similar observation in their study and rationalised this contradiction by questioning the effectiveness of the method of eliciting the stress response. They posed that the previous stressor tasks may have been ineffective at adequately eliciting a stress response (implying TSST is more effective than other social anxiety paradigms), resulting in the contradicting result. Overall, it is apparent that unexplored factors cause this subject level difference in results.

2.4 Summary

As observed in the literature review, stress is a vital physiological mechanism in daily life. While many studies have explored the different methods of eliciting the stress response, most methods rely on controlled laboratory environments (social anxiety and reactionary tasks) in contrast to more realistic and translatable paradigms (Height exposure and other DSS). Paradigms such as height exposure and other DSS paradigms has found mixed results [25, 128, 147]. The development of a realistic and reliable stress elicitation paradigm will enable researchers to further explore more translatable research for human-centred computing interfaces and devices ***RQ1***. BCI studies, in particular, have shown varying and often contradicting results (relation between stress and P300) [100]. The development of more reliable stress elicitation paradigms will better enable researchers to further examine and resolve the contradictory findings in previous literature ***RQ2***.

CHAPTER 3

HEIGHT EXPOSURE EXPERIMENT

3.1 Overview

This chapter outlines the rationale, methodology and stress-related results of our height exposure stress experiment. Here we present our novel height (acrophobia) exposure paradigm that incorporates both virtual and physical elevation to produce the sensation of extreme heights. The results assessed the participant's questionnaire response, HR, HRV, EDA, and gait behaviour to determine the effectiveness of this paradigm to answer (***RO1***). The data and work covered in this chapter are also presented in Zhu et al. [215, 216].

3.2 Height Exposure

Before the rise of VR technology, acrophobia was a challenging stressor to exposure to participants in a safe and controlled manner. A psychologist would often require in vivo methods to expose very few individuals to physical height environments [12, 60, 108]. Emmelkamp et al. [61] are one of the few notable studies that explored in vivo height exposure. This particular study used a second-floor ledge as an exposure site and found it effective in eliciting a stress response detected through HR. Alternatives methods to in vivo exposure have been explored, such as self-exposure to pictures and videos of heights ([12]) and imaginal exposure paradigms [60]. These methods often produce mixed results and require a level of subject implicit acrophobia to be effective.

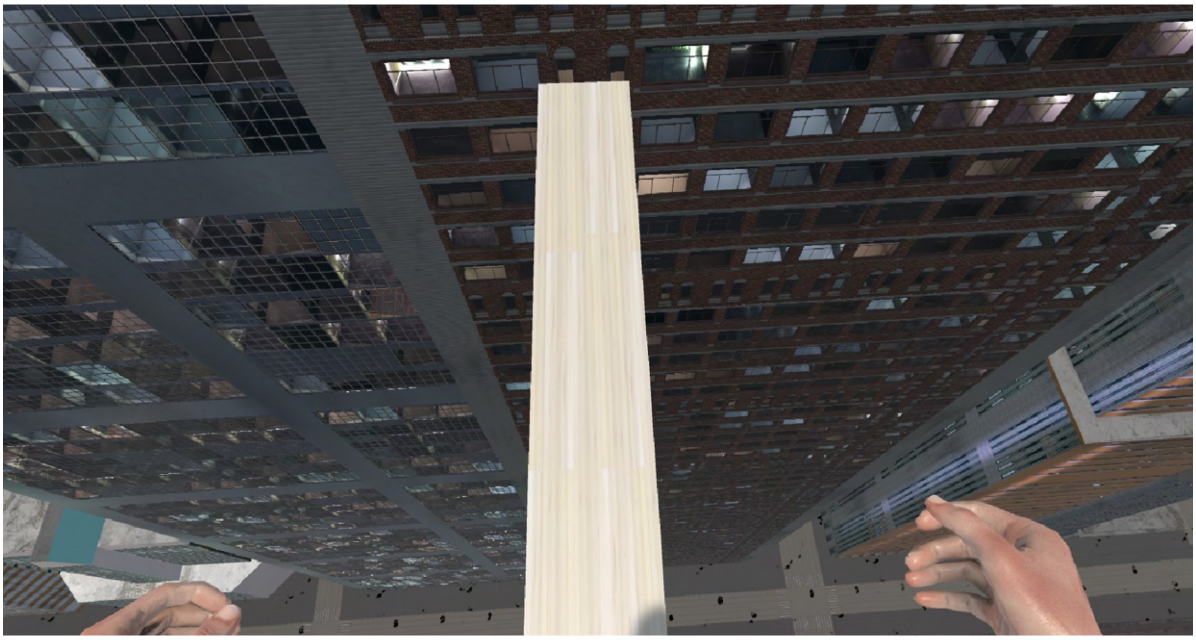

Figure 3.1: An first person view of our prototype virtual height scenario.

The advancement and availability of VR technology opened a new frontier for researchers to create a more realistic, safe, cost-efficient, and controlled environment for extreme height exposure [147, 165]. Multiple studies [35, 60, 181] have compared VR height exposure to the past traditional methods and found comparable results when assessing threat perception and stress levels. VR-based paradigms can provide a virtual environment (VE) that is close to real-world situations and environments. Figure 3.1 provides an example of the first-person view of virtual heights in VR.

3.2.1 Virtual Environment Design

3.2.1.1 Immersion and Presence

VR technology has progressively developed over time and can now render a wide array VEs. Bowman and McMahan [19] identified these two primary factors for the effectiveness of a VE: the level of immersion of the VR system and the sense of presence experienced by the user. Slater [186] defined immersion as the objective degree of sensory fidelity provided by the VR system (and its surroundings) and presence as the perceived subjective response of the human using the VR system. Conventionally, immersion is a function of the hardware capabilities of the VR system (quality of the visual rendering) and the peripheral factors such as auditory noise and haptic feedback [187]. The sense measures presence embodied within a VE, and their situation awareness of the surroundings [19].

3.2.1.2 Visual Realism

Visual realism refers to the level of visual accuracy and quality of the rendered VR display compared to the real-world environment [19]. Slater et al. [187] demonstrated the importance of visual realism by compared two (low vs. high-quality VE) different VE renderings and found a significant increase in a user's sense of presence when experiencing a more realistic VE. Similar studies such as Hvass et al. [92] and Debattista et al. [45] also found similar results. The quality of visual realism is highly dependant on the VR headset and graphics processor unit (GPU) with the appropriate hardware and software capabilities. Selecting the correct VR system can often be complex with the wide range of head-mounted display (HMD) VR systems in the current market. Thus it is essential to have clear guidelines of qualities required to achieve the benchmark visual performance [10]. Bowman and McMahan [19] and Lee et al. [116] outlined these specific qualities as a measure of the capabilities of rendering an accurate display:

- Field of view (FOV),
- Display resolution of inbuilt screen,
- HMD tracking latency and accuracy,
- Realism of lighting within the VE,
- HMD frame rate, and
- Adjustable design for correct fitting (incorrect fitting causes poor display quality)

3.2.1.3 Auditory Stimuli

Auditory stimuli is another critical factor to creating a sensory realistic VE [23]. Auditory stimulation is an indispensable component of the VR system [83]. Sanchez and Slater [167] stipulated that a lack of auditory representation can negatively affect the realism of a VE. Studies by Hendrix and Barfield's [84] and Larsson et al. [112] found that sound can positively affect the user's sense of presence in a VE, and the absence of auditory noise can destroy the sense of presence. This suggests that the simple presence of background noise or appropriate sounds to the VE is enough to increase the realism of the VE.

3.2.1.4 Haptic Feedback

Haptic feedback refers to the user's sense of touch and enables the user to interact with and perceive the VE through tactile senses [27]. Meehan et al. [128] is a prominent study that popularised the use of passive environmental haptic feedback for VR height exposure. The study determined that haptic feedback enhances the user's sense of presence and improves the physiological response to heights. Since then, multiple virtual heights studies have incorporated a floorboard (passive haptic feedback) or other haptics to improve the sense of presence [188]. Later studies such as Asjad et al. [9] and Nagao et al. [140] have further demonstrated the effectiveness of passive haptic feedback in influencing a person's perception of height and presence in a VE.

3.2.2 Virtual Reality Height Exposure

The experimental design presented in this work was inspired by past works in physiological stress and VR height exposure. Early studies, such as Hodges et al. [90], discovered the efficacy of using VR to simulate a heightened environment for acrophobia exposure. Based on the past experiment designs, we have highlighted specific attributes that contribute to a practical height exposure VE [216]:

- VR display should provide a high-quality visual rendering of the VE [19, 77].
- Should have an appropriate auditory stimuli [84, 112].
- The physical space should provide sense of elevation through haptic feedback [128].
- The play area should have a strong correspondence to the physical space [187].

The recent Peterson et al. [147] study explored the effects of stress from height exposure with an experimental setup that meets the established criteria. The setup used a heightened (2.5 cm) beam for walking with VR to visualise height (15 m). A high-quality GPU (NVIDIA Titan X) was used to render the VE, and background noise was introduced for auditory stimulation. The VE also included a virtual avatar for virtual embodiment and identical virtual walking space. This setup was successful in inducing physiological stress from acrophobia. Another notable study is a virtual work-at-height simulator developed by Loreto et al. [50]. The study used a real ladder (virtual height 11 m) as passive haptic feedback. The VR would simulate climbing the ladder to the heights of power lines for training [50]. The unique incorporation of vibrations in the ladder caused

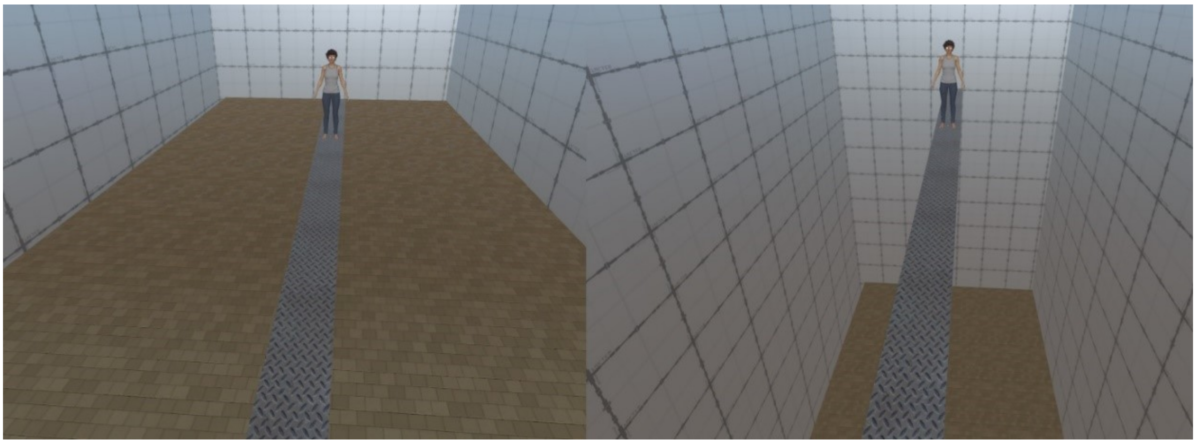

Figure 3.2: An early iteration of VR heights experiment

a significant increase in the perceived realism of the VE. The experimental setup of this work builds upon these two studies [50, 147] by further heightening both the physical (0.65m) and virtual elevation (150m) increase the sense of elevation.

3.3 Methodology

3.3.1 Preliminary Experiment

Before this elevated physical platform, we performed a pilot experiment with purely virtual height modulation (15 m), see Figure 3.2. We collected five healthy participants and recorded their SAM responses. Each participant performed 24 walking trials with randomised (counterbalanced) conditions (15 m vs 0 m). Ultimately, the SAM results showed no significant difference among participants between the virtual height condition (M= 2 SD= 0.54) and the ground condition (M= 1.78 SD= 0.43). This lack of stress response led to the formulation of the VE criteria in subsection 3.2.2 and the redesign shown in Figure 3.3.

3.3.2 Experiment Physical and Virtual Environment

3.3.2.1 Physical Space

One novelty of our experimental design is the elevated physical platform. The platform provided physical height, adding a sense of physical elevation. We set two requirements when designing this platform. The participant's safety was the first requirement, and

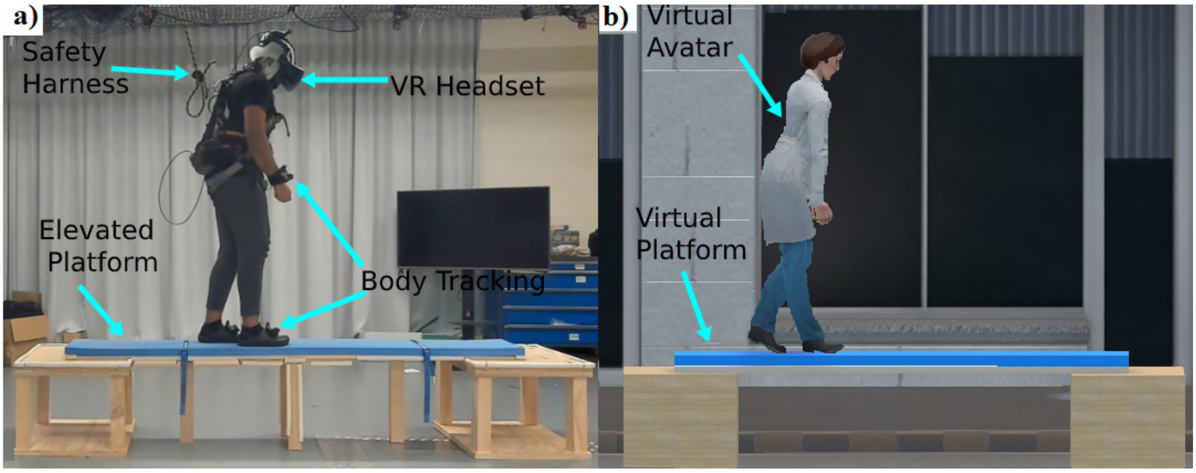

Figure 3.3: a) The physical elevated platform, VR headset, body tracking, and safety equipment. b) The virtual environment used in this experiment with a virtual avatar that is driven by body tracking.

the second was to introduce a sense of fear from the awareness of physical elevation. To ensure participant safety, we lined the corners and edges with protective foam and installed a harnessed fall arrest system (to prevent fall-related injuries). The platform's height (0.65 m) was set to the maximum allowable height based on the height of the rail, safety line, and harness system. We progressively designed the platform and rated the design to support up to 150 kg (experimental exclusion criteria restricted this to 95 kg); this was set through the failure of a previous iteration to support the specific weight.

The second concern was creating a sense of danger/fear on the platform. We introduced a small amount of instability through metal plates on the bottom of the platform. A surface foam layer was introduced to add further postural instability when walking on the elevated platform. We found that instability factors added difficulty to walking, which increased the participant's anxiety levels. The ground platform consisted of a separate board with matched dimensions to the elevated platform. This correspondence ensured consistent gait behaviour during the experiment. Both the elevated (0.65 m height) and ground (0.02 m height) platforms had the exact walking space dimensions of 2.4 m long and 0.3 m wide, see Figure 3.4.

3.3.2.2 Virtual Space

We designed the VE to have a strong correspondence with the experimental setup. The physical plank's dimensions, orientation, and position were measured through the Optitrack motion capture (mocap) system (12 flex13 cameras) and mapped in virtual

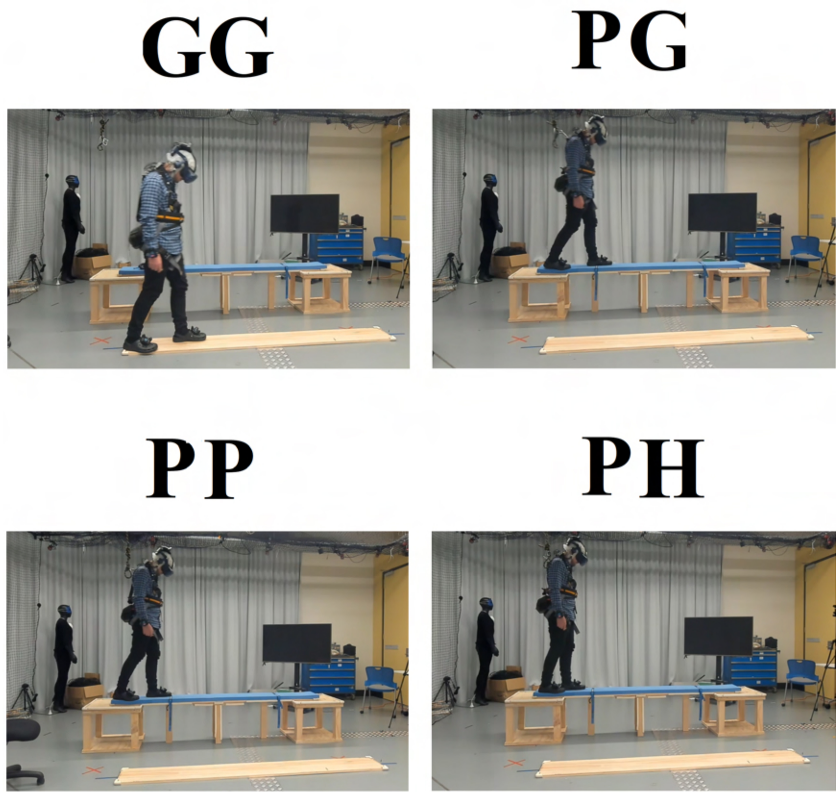

Figure 3.4: A comparison of the participants during walking in the experiment.

space. The usage of mocap ensured the accurate translation between the physical and virtual space. A calibration process was developed using the HTC Vive controllers, which captures any offset movement and vibrations of the physical platform and adjust the VE accordingly.

We chose to set the VE in an urban environment because the buildings and the overall city provided the participants with a believable environment for the virtual height. The buildings emphasised the sense of height through the scaling of the surrounding buildings; see Figure 3.3 for the physical and virtual setup. The high-quality VE was rendered through the HTC Vive Pro VR HMD and used a high power GPU with sufficient processing (NVIDIA GTX1080 GPU, Core i7). A wireless adapter was added to improve the safety and mobility of the participants.

The participants used a virtual avatar as the medium for virtual embodiment. The avatar used inverse kinematics (FinalIK by RootMotion [163]) and a six-point body tracking system (one HMD, one waist tracker, two hand trackers, and two feet trackers) using HTC Vive trackers. The avatar improved the participant's spatial awareness of the VE and their sense of presence. Ambient urban environment noise was played in the

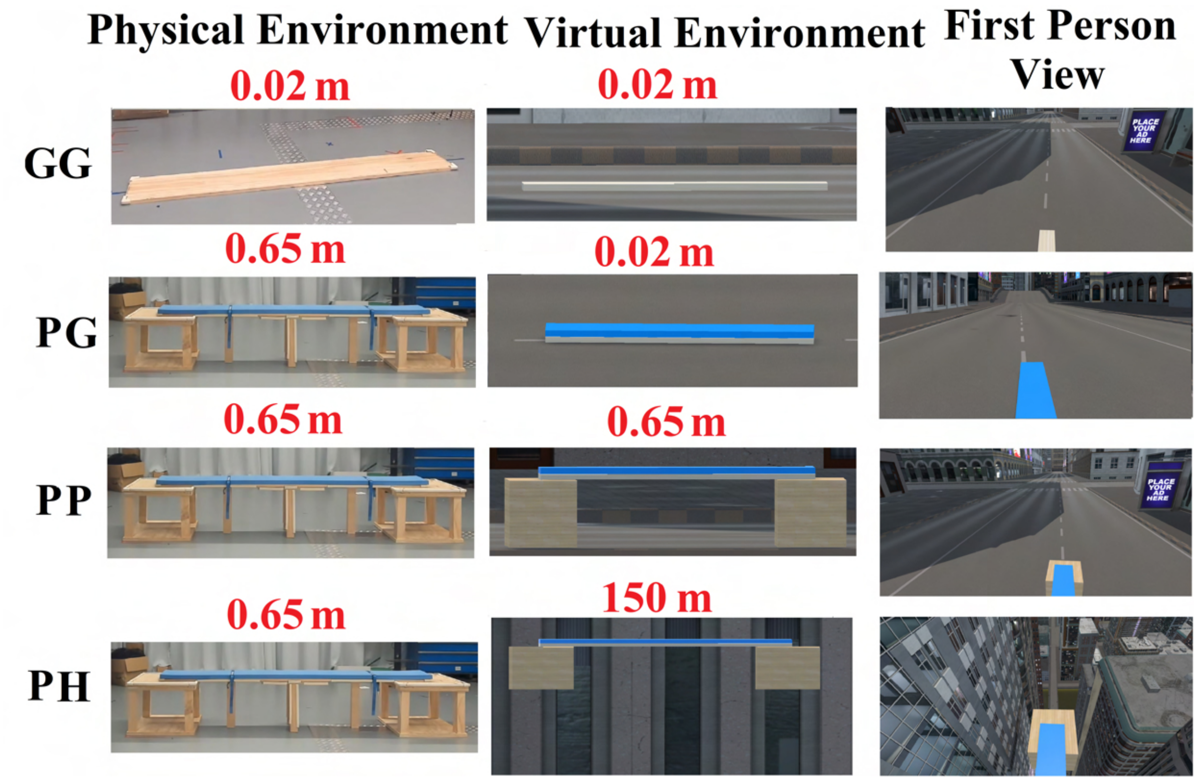

Figure 3.5: The experimental conditions with condition labels, *GG*, *PG*, *PP*, and *PH* (G=ground, P=platform, H=Extreme height). First condition letter: the physical experiment setup; second condition letter: the virtual experiment setup.

background to serve as the auditory stimuli.

3.3.3 Experiment Design and Protocol

This experiment tested four conditions; each condition consisted of a combination of the physical (ground and elevated platform) and virtual (ground, elevated platform, and extreme height) independent variables. Figure 3.5 outlines each condition for this experiment; the conditions are *GG*, *PG*, *PP*, and *PH*. Every participant experienced the four conditions in a randomised sequence (Figure 3.6). Due to timing constraints of the physical setup, the physical ground platform (*GG*) was randomised separately to the elevated platform (*PG*, *PP*, and *PH*). Each experimental condition consisted of 5 trials (1 baseline walk and 4 trials with 25 Oddball trials tasks).

One walking trial constituted the trip from the starting position to the end, and the return walks back to the start (see Figure 3.7A). Participants were instructed to walk in their natural gait. During the return trip for a non-baseline walk, the participant

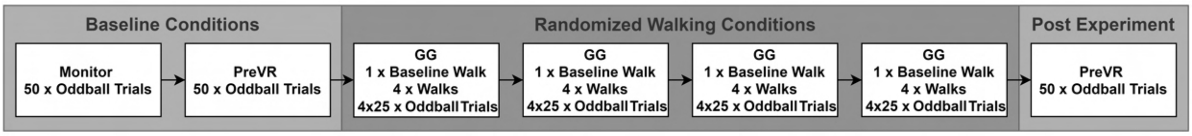

Figure 3.6: The timeline for each condition and the trials per condition.

would perform 25 oddball trials in the middle of the platform. We selected oddball as the P300 task due to the reliability and simplicity of implementation. The oddball task consisted of a serialized sequence of images consisting of green (non-target) and blue (target) circles at a 1:4 target to the non-target ratio (see Figure 3.7B). The participants reacted to the target stimuli through a keypress on the remote control. The interval timings for the Oddball stimuli were decided based on the Sellers et al. [174] study, which assessed P300 performance at various intervals and target:non-target scenarios. The study found that optimal P300 performance occurs when adequately long (interval > 350ms) stimuli reveal interval and easily discernible images. Hence we chose colour variation as the target/non-target difference. 50 baseline trials of the oddball task were collected while the participants were seated (on the computer monitor) and standing (in VR). All events were synchronized through a lab streaming layer (LSL) server. The same standing oddball task was performed after the experiment, along with a post-experiment questionnaire. The participants were given 3 minutes (minimum) resting periods between each condition; the participants may extend the rest time based on need. See Figure 3.6 for the timeline of the experiment.

During the conditions, the participants were encouraged to wear the VR HMD throughout the experiment continuously. This provided a continual sense of presence in the VE. The participants would step onto the physical platform while in a VR calibration area and then be shown the appropriate scene for the condition. The participant would also be clipped onto the safety harness during the elevated walking conditions. During rest breaks, participants were encouraged to close their eyes to prevent eye strain and provided the option to remove the VR HMD if they felt severe discomfort. During the experiment, we maintained social distancing protocols and sanitised all equipment.

3.3.4 Participants

This human research experiment had the approval of the UTS research ethics committee. All participants provided written and informed consent and received 60 AUD compensation for their participation (regardless of outcome or termination). We recruited 20

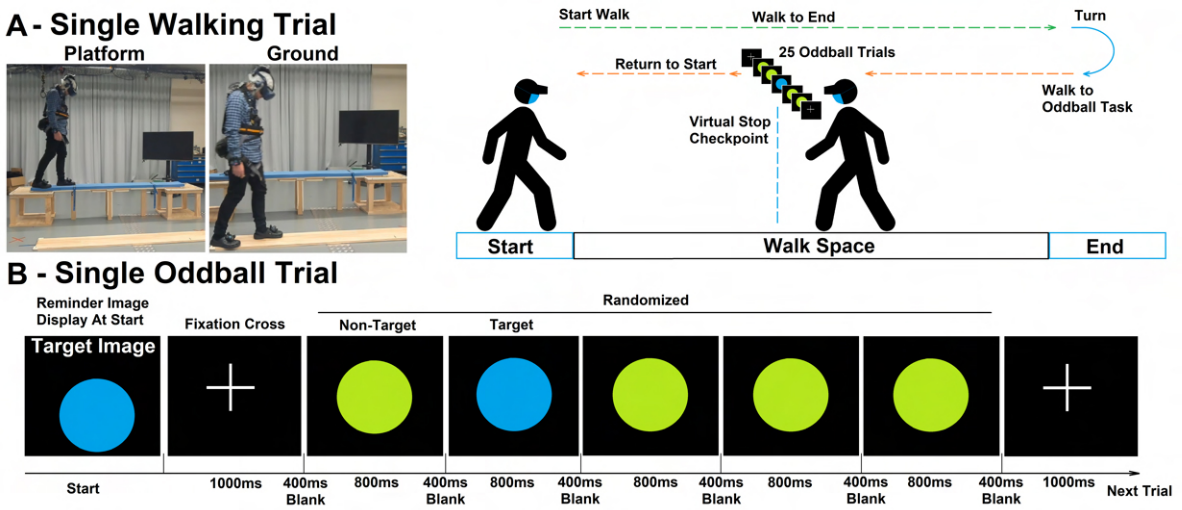

Figure 3.7: (A) An outline of a single walking trial (the real world pictures were taken before pandemic restrictions). (B) An outline of a single Oddball trial with the respective stimuli time periods at the bottom.

adults (5 females and 15 males) between ages 21 to 35. The mean age was 26, and the population variance was 4.70. The key exclusion criteria were:

- an inability to understand the experimental instructions (language and cognitive ability),
- existing medical conditions, such as
  - neurological and cardiovascular disorder,
  - diagnosed mental health issues (depression, anxiety, or chronic stress),
  - sensory (visual, vestibular, or auditory) dysfunction, and
  - gait (unable to perform unassisted walking) disorders.
- weight more than 95 kg (participant safety).

The data analysis only included a dataset of 18 participants. One participant's (male) data were excluded due to incomplete data from hardware failure. The other participant (male) felt overwhelmed by the height exposure and did not complete the *PH* condition after the first walk.

Figure 3.8: The equipment used in this experiment for VR visualisation, tracking, and physiological measuremets.

3.4 Measurements and Analysis

3.4.1 Questionnaires

Before starting the experiment, participants were given a questionnaire that provided insight into their level of acrophobia and their prior experience with heights, video games, and VR. The participant was asked to provide an arousal Self-Assessment Manikin (SAM) rating (1-9) [20, 215]. The participants were shown the SAM questionnaire figures before the experimental conditions. The ratings were verbally (without the figures) asked at the end of each walk. The SAM analysis involved separating the data by condition and averaging across 18 participants. The SAM results were also used to evaluate the condition order effect.

During each resting period for the experiment conditions, the participant filled a modified Depression Anxiety Stress Scale (DASS) questionnaire [121]. The DASS questions involved the participants answering a series of questions (rating 0-3) to gauge their level of anxiety (sense of uncertainty) and stress (heightened arousal and emotional response). The questionnaire was modified to remove the depression component. This modification shortened the questionnaire to eight questions (originally twenty-one, four anxiety questions and four stress questions). The DASS analysis separated the data by condition and components (anxiety and stress), then averaged and compared. The DASS scores were compared to the normative scores based on the previous DASS-42 [121] and DASS-21 [85] studies. The score ranges were scaled to a four-question category range (DASS-42 has fourteen, and DASS-21 has seven per category). The scaled normative scores are shown in Table 3.1. After the experiment, the participants retrospectively ranked (user ranking) each condition from 1st (most stressful) to 4th (least stressful)

Table 3.1: Scaled DASS normative scores based on previous studies. [216]

	Stress			**Anxiety**		
DASS	**Scaled**	**21**	**42**	**Scaled**	**21**	**42**
Normal	0-4	0-7	0-14	0-1	0-3	0-7
Mild	5-6	8-9	15-18	2-3	4-5	8-9
Moderate	7-8	10-12	19-25	4-5	6-7	10-14
Severe	8-9	13-16	26-33	6-7	8-9	15-19
Extremely Severe	10+	17+	34+	8+	10+	20+

3.4.2 Physiological Measurements

Figure 3.8 shows the various apparatus used for this experiment.

3.4.2.1 Heart Rate and Heart Rate Variability

The participants wore a wearable Zephyr bioharness device to measure and calculate their HR and HRV. The device is lightweight (85 g) and accurate. It measures HR through electrocardiography (ECG, sampled at 250 Hz) [141]. The ECG data gave the HR (BPM) and HRV data. The HR calculation used the RR interval time [214]. The HRV calculation involved a 300-sample (5 minutes) rolling average of the standard deviation of the NN interval time [214].

We filtered the HR and HRV data points based on the normal human ranges (30<BPM<120, and 20 ms<HRV<110 ms) and removed high noised portions of the ECG signal (noise>20%). These boundary parameters were determined based on the age group (21-35), body weight (less than 95 kg), and relative physical exertion (standing and walking) criteria [62]. The HR data were normalised per participant through feature scaling (min-max) normalisation to account for the variance in cardiovascular behaviour.

3.4.2.2 Electrodermal Activity

The Empatica E4 wristband was used to measure EDA. Wristbands are portable and non-intrusive and recorded the EDA (sampled at 4 Hz) at an acceptable level of data quality [126]. The EDA (μs) was measured by dry electrodes on the wristband. The data points were separated by condition, filtered by the removal of values outside of the human skin conductance range (0<μs<20) and outlier values (2 standard deviations from the mean). Similar to the HR data, the EDA data were normalised per participant to account for the variance in individual skin conditions through feature scaling (min-max).

3.4.3 Behavioural Measurements

We calculated the participant's behavioural gait parameters through the six-point motion tracking system. Each tracker recorded the global position vector (x, y, z), orientation vector (Quaternion), and time stamping. We excluded the gait parameters of the return trip of each walking trial due to the oddball trial pausing the continuity of the gait movement. The calculation also excluded each trial's first and last step (stepping onto and off the platform).

We calculated three parameters, the step length, step count per trial, and trial completion time to reproduce the results from past studies [26, 170]. The gait parameter calculation involved replaying the tracker data (with inverse kinematics) and the Unity game engine collider system. We filtered outlier values (2 standard deviations from the mean) and averaged them across the 18 participants.

3.4.4 Statistical Analysis

The participant sample size was calculated through Wilcoxon signed rank t test with standard large effect size of 0.8 and an error α of 0.05. After calculating the parameters through G*power, the sample size of 20 was determined to be sufficient for observing the effect of the study. A one-sample Kolmogorov-Smirnov test (compared against cumulative distribution function) was applied to determine the normality of each metric. The test found the HR, HRV, EDA, and DASS scores to be normally (Gaussian) distributed; we applied one-way ANOVA to determine the statistical significance. The gait parameters and SAM ratings were of non-normal distribution; thus, the appropriate test was the Wilcoxon signed-rank test. The statistical tests compared the four conditions pairwise with a significance level ((α) of 0.05, determining statistical significance. The significance stars for the figures are *p<0.05, **p<0.01, and ***p<0.001.

3.5 Results

3.5.1 Questionnaire

3.5.1.1 Depression Anxiety Stress Scale

Figure 3.9 depicts the DASS questionnaire results (Table 3.2) for each condition. The anxiety results found a significant difference between the *PH* condition and the *PP* (F(1,34)=10.56, p=0.0026, and partial η^2=0.24), *PG* (F(1,34)=7.45, p=0.01, and partial

Table 3.2: DASS mean, standard deviation, Normative Rating (NR, from Table 3.1), and Normative Range (from Table 3.1)

	Mean	SD	NR	Range
Stress				
GG	2.500	2.256	Normal	Normal-Mild
PG	2.667	2.029	Normal	Normal-Mild
PP	2.722	2.024	Normal	Normal-Mild
PH	4.611	2.453	Mild	Normal-Moderate
Anxiety				
GG	1.556	1.789	Mild	Normal-Moderate
PG	1.944	2.182	Mild	Normal-Moderate
PP	1.722	1.775	Mild	Normal-Moderate
PH	4.111	2.564	Moderate	Mild-Severe

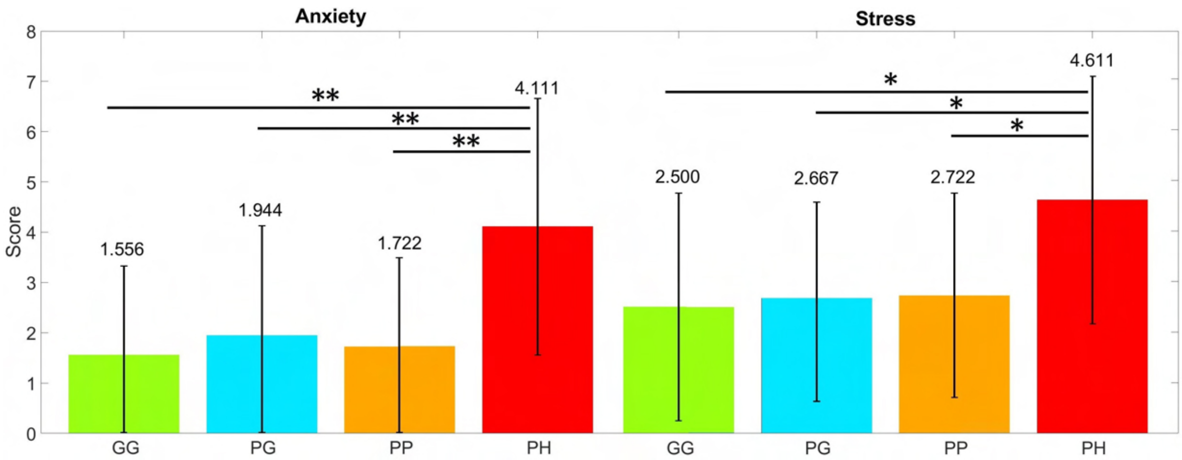

Figure 3.9: DASS anxiety and stress scores with significance.

η^2=0.18), and *GG* (F(1,34)=12.02, p=0.0014, and partial η^2=0.26) conditions. Similarly, the stress results demonstrated a significant difference between the *PH* condition and the *PP* (F(1,34)=6.35, p=0.0166, and partial η^2=0.16), *PG* (F(1,34)=6.72, p=0.0140, and partial η^2=0.16), and *GG* (F(1,34)=7.22, p=0.0111, and partial η^2=0.18) conditions.

3.5.1.2 Self-Assessment Manikin

Figure 3.10a presents the SAM results (*GG* M=1.40 and SD=0.94, *PG* M=2.37 and SD=1.58, *PP* M=2.62 and SD=1.54, and *PH* M=5.03 and SD=2.21) across the 18 participants. The results in the *PH* condition were significantly different (W=0, Z<-3.41, p<0.001, and r>0.80) from all the other conditions (*PP*, *PG*, and *GG*). There is also a significant difference when comparing the *GG* condition to the *PG* condition (W=0, Z=-3.18, p=0.0015, and r=0.75) and the *PP* condition (W=4, Z=-3.05, p=0.023, and r=0.72). Based on the average value and significance, the trend of the SAM ratings was *PH>PP=PG>GG*. Table 3.3 depicts the SAM scores (mean and standard deviation) sorted by the sequence

Table 3.3: The mean and standard deviation of participant's SAM responses (rating 1-9) based on the sequence of conditions

Condition	1st	2nd	3rd	4th
GG	1.76 ± 1.19	N/A	N/A	1.04 ± 0.13
PG	3.85 ± 1.20	2.29 ± 1.59	1.10 ± 0.10	1.80 ± 1.01
PP	3.90 ± 1.10	3.24 ± 1.30	2.09 ± 1.42	1.00 ± 0.00
PH	3.30 ± 1.30	5.87 ± 1.63	5.14 ± 2.06	4.27 ± 2.72

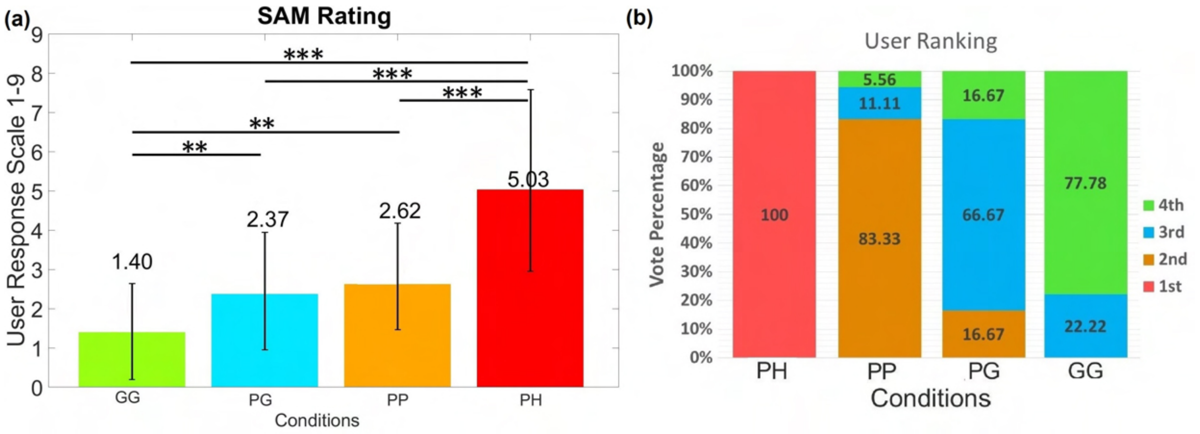

Figure 3.10: Bar plot of the average (a) SAM rating with the standard error bars and (b) Retrospective user ranking (1st-4th) of the conditions based on perceived stress levels

the participant experienced each condition.

3.5.1.3 User Ranking

Figure 3.10b shows the results of the retrospective ranking. The *PH* condition was overall ranked 1st based on the votes of every participant. Based on the majority votes, the subsequently ranked conditions were the *PP* (2nd), *PG* (3rd), and *GG* (4th) conditions.

3.5.2 Heart Rate and Heart Rate Variability

Figure 3.11 illustrates the average (standard deviation bars) HR values (*GG* M=52.56% and SD=6.92, *PG* M=55.52% and SD=14.910, *PP* M=55.86% and SD=8.87, and *PH* M=64.88% and SD=12.53) and HRV values (*GG* M=72.89% and SD=16.74, *PG* M=82.82% and SD=22.98, *PP* M=81.61% and SD=18.52, and *PH* M=87.69% and SD=19.48) of the 18 participants.

The HR response during the *PH* condition was significantly different from that of the *GG* (F(1,30)=11.86, p=0.0017, and partial η^2=0.28), and *PP* (F(1,30)=5.53, p=0.0255, and partial η^2=0.16) conditions. The HR data did not show significant difference when

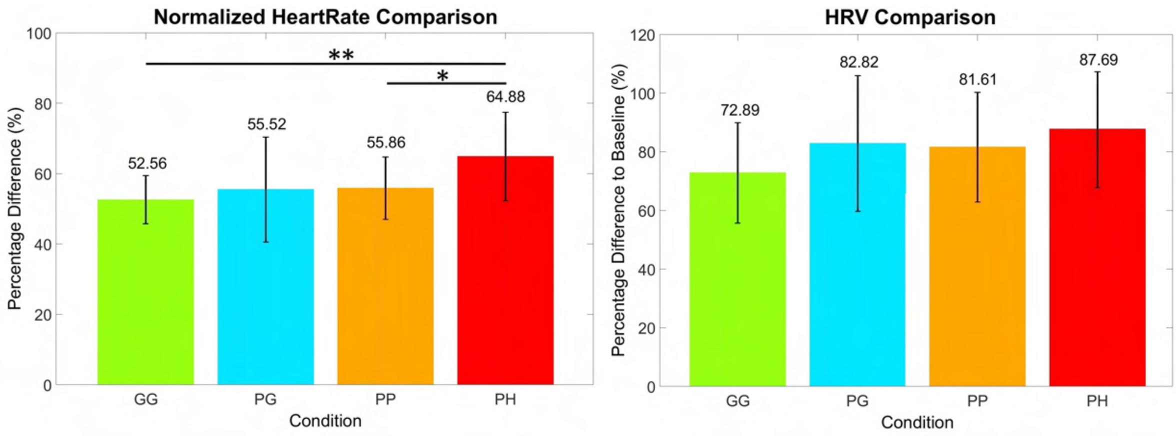

Figure 3.11: A bar plot of the normalised HR and HRV values averaged across all the participants.

comparing *GG* condition to the *PG* and *PP* conditions. There was also no significant difference when comparing *PG* to *PP* and *PH* conditions. We did not find a significant difference in HRV data among the experimental conditions (F(1,30)<2.3, p>0.13, and partial η^2 <0.07).

3.5.3 Electrodermal Activity

Figure 3.12a shows the average (with standard deviation bars) EDA results (*GG* M=33.99% and SD=27.66, *PG* M=39.38% and SD=28.87, *PP* M=41.30% and SD=23.79, and *PH* M=57.53% and SD=14.74) of the participants. There was a significant difference when comparing *PH* to *PP* (F(1,30)=5.04, p=0.0322, and partial η^2=0.14), *PG* (F(1,30)=4.98, p=0.0332, and partial η^2=0.14), and *GG* conditions (F(1,30)=8.89, p=0.0056, and partial η^2=0.13). The *GG* condition showed no significant difference in EDA from *PG* and *PP* conditions. Similar to HR, the condition *PG* had no significant difference from the *PP* condition.

3.5.4 Gait

Figure 3.12b shows the average trial completion time (*GG* M=3.27 and SD=0.88, *PG* M=2.55 and SD=5.60, *PP* M=1.65 and SD=4.77, and *PH* M=3.23 and SD=7.65) of each condition. The notable finding of this measure was the significant difference between the *GG* condition and the *PG* (W=20, Z=-2.67, p=0.0075, and r=0.65), *PP* (W=15, Z=-2.91, p=0.0036, and r=0.71), and *PH* (W=10, Z=-3.15, p=0.0016, and r=0.76) conditions. There

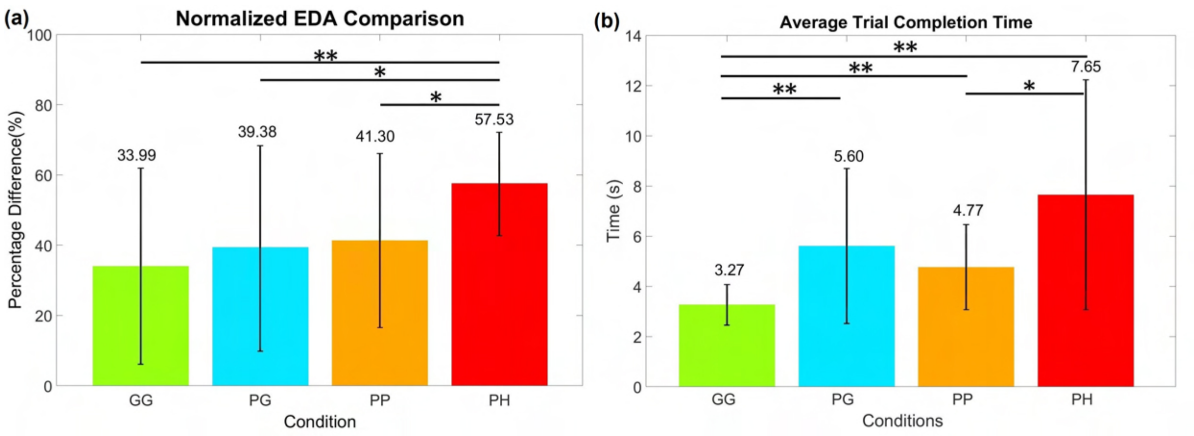

Figure 3.12: A bar plot of the normalised (a) EDA value averaged across all the participants and average (b) trial completion time.

was also a significant difference between *PP* and *PH* (W=29, Z=-2.25, p=0.0245, and r=0.54).

Figure 3.13a depicts the average step distance (*GG* M=0.42m and SD=0.07, *PG* M=0.36m and SD=0.09, *PP* M=0.35m and SD=0.07, and *PH* M=0.29m and SD=0.08). The step length demonstrated a significant difference between the *GG* condition and the other conditions, *PG* (W=152, Z=2.90, p=0.0038, and r=0.68), *PP* and *PH* (W=171, Z=3.72, p<0.001, and r=0.88), on the elevated platform conditions. The step length in the *PH* condition was significantly different from that of the *PG* condition (W=157, Z=3.11, p=0.0018, and r=0.73) and the *PP* condition (W=166, Z=3.51, p<0.001, and r=0.83). No significant difference was found between *PG* and *PP* conditions.

Figure 3.13b illustrates the step count (*GG* M=3.91 steps and SD=0.66, *PG* M=5.68 steps and SD=1.88, *PP* M=5.50 steps and SD=1.27, and *PH* M=7.52 steps and SD=2.99) per trial. There was a significant difference when comparing the *GG* condition to the *PG* (W=3.5, Z=-3.46, p<0.001, and r=0.84, *PP* (W=0, Z=-3.52, p<0.001, and r=0.85), and *PH* (W=1, Z=-3.57, p<0.001, and r=0.87) conditions. There was also significant difference when comparing the *PH* condition to the *PG* condition (W=18.5, Z=-2.75, p=0.0060, and r=0.67) and the *PP* condition (W=10.5, Z=-2.82, p=0.0048, and r=0.68). Similar to the step length, no significance was found when comparing the *PG* condition to the *PP* condition.

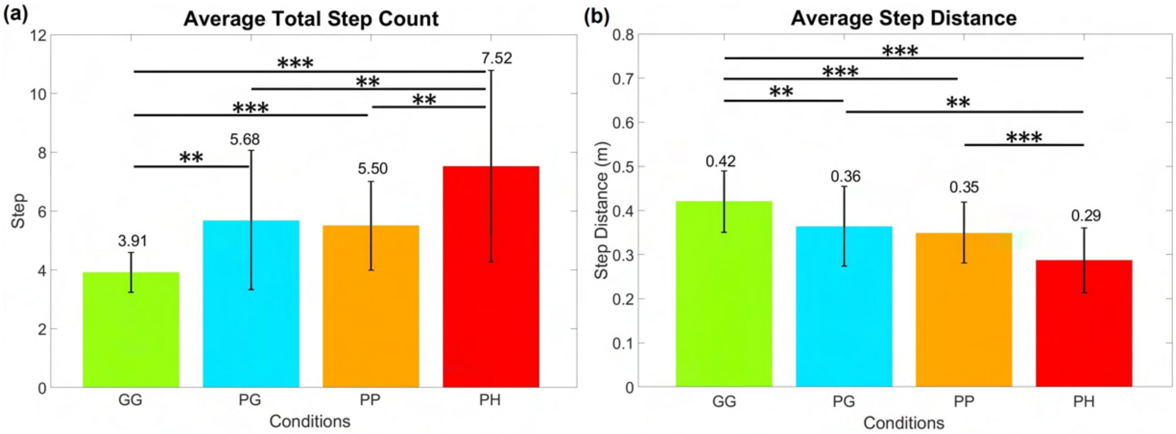

Figure 3.13: Bar plot of the average (a) step count per trial and the average (b) step distance.

3.6 Discussion

3.6.1 Overview of Results

From the questionnaire results, it is evident that the exposure heights induce physiological and behavioural changes in the participants. The *PH* condition notable produces the biggest change in participants with the highest SAM rating, highest DASS score and being ranked the most stressful condition by all participants. In contrast, the *GG* condition received the lowest SAM rating, lowest DASS score and ranked the least stressful condition. Surprisingly, the *PG* and *PP* conditions produces mixed results. The SAM ratings and user rankings suggest *PG* and *PP* to be between *GG* and *PH*. On the other hand, the DASS scores put forward that the conditions are more likely low stress (same as *GG*), which correlates to the HR and EDA findings. Another consideration is the normative ratings of the DASS scores from Table 3.1. The normative ratings suggest that *PH* is a slightly higher level of stress and anxiety rating by one category above the other conditions (Table 3.2).

One difference between the DASS scores and SAM rating is the time point at which it was taken (DASS between conditions and SAM after each trial). The DASS scores are more indicative of sustained symptoms of anxiety and stress after completing the condition. The SAM trial results in Table 3.3 maybe imply that the *PG* and *PP* produced stress levels that subsided at a faster rate compared to *PH*.

A higher level of stress, in general, induces higher HR and EDA and a lower HRV [172, 175, 192]. The *PH* condition has a substantial increase in both HR and EDA compared to the other conditions. Similar to the DASS results, there was no significant difference between the *GG* condition and the *PG* and *PP* conditions in HR and EDA data. One potential explanation could be that in the *GG* condition, the participants engaged in a stepping strategy that required a higher level of physical exertion, i.e. they took larger steps at a higher cadence, which resulted in an increased HR and EDA. The lack of significance in the HRV results may be due to the short duration of each trial (3-7 seconds, Figure 3.12b) compared to the 5 minute rolling average for the calculation. In the future, including a more extended stationary component to each condition may yield a more observable effect for HRV. Peterson et al. [147], and Lee et al. [116] also observed a similar difficulty when analysing HR, HRV, and EDA changes due to the physical exertion of walking at heights.

The gait parameters corroborate the findings of the SAM rating. Like the other metrics, the *GG* and *PH* conditions show a significantly significant difference compared to the other conditions and the *PG* and *PP* conditions showing no significant difference. The results found a correlation between the level of height exposure and the stepping length, total steps, and trial completion time. The increase in threat (height) perception caused the participant to engage a different stepping strategy, becoming more cautious with a reduction in step length, the number of steps taken and trial completion time. This result concurs with the findings of Schniepp et al.[170]. However, the walking surface is a crucial consideration when comparing *GG* condition to the other conditions. The *GG* condition did not have a foam walking surface, and participants did not wear a safety harness. Thus the change in gait in *GG* may be a result of condition variation rather than stress. However, a similar change in gait behaviour occurred when comparing the *PG* and *PP* conditions to the *PH* condition. In this case, the physical environment was consistent across these conditions. Therefore, this behavioural shift may be more likely due to the change in perceived height.

3.6.2 Effects of Virtual Elevation

By comparing the *PG*, *PP*, and *PH* conditions, we can observe the effects of modulating the virtual height as the physical height is consistent. As expected, at the extreme level of virtual height, we observe a notable increase in stress level consistent with previous literature. However, a contention occurred when evaluating the lower levels of virtual height (*PG* and *PP*). Logically, one would expect the *PP* condition at a mid-level elevation to induce a higher level of stress when compared to *PG*. In contrast, we found no significant differences between the *PG* and *PP* conditions for most of the measurements (except user ranking), suggesting an equivalent level of stress.

A potential reason for this unexpected result could be the level of virtual height being insufficient to produce a fear of heights. Wuehr et al. [205] investigated the concept of a height threshold for producing a fear of heights. The study found that exposure to heights is not a linear proportional relationship. Instead, the most significant behaviour change occurs at the virtual height of 20 m and begins to saturate at approximately the 40 m mark [205]. This would infer that there is a minimum height threshold for virtual height to be effective. Heights under this threshold would not produce a strong physiological response compared to the overt response after passing the height threshold. Virtual heights at 0.65 m were likely below this threshold. In contrast, the height in the *PH* condition of 150 m was above the upper maximal threshold. Thus the effect was observable.

In summary, in extreme virtual heights (*PH*), there was a distinct level of high stress; however, at lower levels of virtual height (*PP* and *PG*), there was a less discernible difference in the levels of stress. This finding suggests that low-level virtual height is unreliable for inducing multiple levels of stress when on the elevated platform.

3.6.3 Effects of Physical Elevation

The comparison of the *GG*, *PG*, and *PP* conditions can convey the efficacy of the physical elevation in inducing stress. The condition had different physical elevations with a negligible (below the threat threshold) virtual elevation. The significant difference in the SAM ratings and gait parameters between the *GG* condition to the *PG* and *PP* conditions would infer an increase in stress. However, the absence of behavioural and physiological differences suggest the *PG* and *PP* conditions produces no change in stress level.

From a tactile sensory point of view, the participants should not be able to discern physical elevation while using VR within a VE. A possible explanation could be that the instability of the platform and the tactile surface causing stress from the postural imbalance. Another theory could be that the participants feared falling off the elevated platform as they were knowingly wearing a HMD VR. The evaluation of this is difficult as the participant were also aware of the safety harness, which may have provided a sense of safety [44].

A plausible explanation for this phenomenon could be a presupposition of the height altering the person's perception of height. The implicit prior knowledge of the platform's height can affect the perceived height and fear response experienced [16, 60, 213]. The presupposed knowledge of the platform's height from visually seeing the platform before wearing the VR headset and the memory of physically stepping up onto an elevated surface may produce mixed effects for various participants. Knowing the height could easily comfort some while also creating anxiety for others.

Overall, the results confirm that the elevated physical platform may cause a conscious change perceived threat at low levels of virtual height. However, only the combination of virtual and physical elevations can reliably induce physiological stress levels.

3.6.4 Incongruence between Physical and Virtual Environments

An additional factor is an incongruence between the physical and virtual heights. The incongruent extreme height environment (*PH*) can cause an increase in stress; this is also supported by literature [9, 50, 128, 147]. Would the opposite effect occur if one modulates the virtual height to a ground-level while physically walking on a more threatening elevation? If the virtual height is the determinant of threat, then it would be safe to assume that *GG* and *PG* would share identical behaviour being the same virtual height.

Then, the significant difference in the stress levels between the *GG* and *PG* conditions is particularly surprising. This result contradicts the expected effect of previous VR distraction studies [7, 18] that suggest the VE can induce a sense of safety by detracting the user from the stressful environment. In the *PG* condition's case, it seems the VE failed at distracting the participant from the physical perception of danger.

The rationale conclusion is that no particular (physical or virtual) persistent bias determines the participant's threat perception. Instead, the participants will focus on the most dangerous or higher source of perceived threat, which then dictates their stress levels. In the *PG* condition, the most significant threat was physical height. Conversely, in the *PH* condition, the virtual height was the most significant threat.

3.7 Limitation

3.7.1 Condition Selection

The time length of the experiment was a critical limiting factor for this study. The full experiment was 2 by 3 experimental design that included 2 levels of physical heights (ground and platform) and 3 levels of virtual heights (ground, plank, and extreme height). However, the experiment would have required 3-4 hours to complete. Therefore, two interesting conditions have to be excluded. We chose to remove the two conditions with the ground platform with a virtual platform (GP) and extreme heights (GH). These conditions were chosen because the expected effect was already well established in previous works [9, 50, 128, 147].

3.7.2 Habituation and Condition Sequence Effect

When experimenting with repeated measures using a within-subject design, habituation and sequence effects are major confounding factors. Over time it is expected that the stress response will decrease and affect the participant's behavioural and physiological measures. Thus condition orders may contribute to variation in participant's reactions to stimuli. A between-subjects experimental design could have resolved this issue. However, it would require a much larger sample size, which was challenging to obtain due to the presence of the global pandemic. Instead, the order of the conditions was randomised to mitigate the issue.

The SAM results in Table 3.3 evaluates the effect of the sequence condition. There is a clear descending trend for the conditions *GG*, *PG*, and *PP* based on the sequence of conditions. Interestingly, participants who experienced the *PH* condition first rated their SAM response lower than the other participants. However, the imbalance of condition distribution is the likely culprit for the lower rating. Overall, it is clear that the *PH* condition is least affected by the sequence effect compared to the other conditions. A better distribution of conditions and participants will likely yield more accurate findings to this sequence effect.

3.8 Key Points

Based on all the results, it is clear that height exposure is a reliable and realitic method of eliciting a stress response ***RQ1***. All the metrics consistently indicate that virtual and physical height (*PH*) can elicit an overt stress response in comparison to walking on the ground (*GG*). However, there is uncertainly for the middle-level heights of *PG* and *PP*. In the latter analysis of the EEG data, likely, the difference between *GG* and *PH* is far more overt compared to the *PG* and *PP* conditions.

CHAPTER 4

Effects of Stress on Cognitive Performance during Height Exposure

4.1 Overview

This chapter presents the EEG results and discusses the findings to answer ***RQ2***. The EEG data were recorded from the participants in the experiment outlined in Chapter 3. We evaluated the RT, ERP, and Event-Related Spectral Perturbation (ERSP) to assess the effects of stress on P300. As outlined in Chapter 2 Table 2.4, past studies have produced mixed results, here found two unique subject groups that exhibit contrasting results similar to the varying types of the previous finding.

4.2 Methodology

The data in this chapter was collected from participants in Chapter 3. One participant (male) was excluded (N=17) due to failure to complete the oddball for one condition.

4.2.1 Oddball Reaction Time (RT)

The trials for the oddball task is outline Figure 3.7 in Chapter 3. The participant will complete a total of 550 trials of the oddball task (3x50 trials baseline and 4x4x25 trials during each condition). All oddball trials were completed in a stationary position. We collected the baseline for both seated viewing a traditional computer monitor and standing

Table 4.1: The demographic information for Group 1 (worse performance under stress) and Group 2 (no change in performance under stress).

Group	N	Sex	Mean Age	Variance Age
1	8	5 male 3 female	25.13	4.62
2	9	7 male 2 female	26.33	4.62

in VR. Monitor baseline also serves as the training/practice task. The participants were instructed to press the tactile button on a remote whenever a target image was observed. The RT was recorded through LSL reaction events sent by the remote.

The reactions times were calculated through the latency difference between the target stimuli appearing event and the reaction event. We filtered out trials with incorrect reaction events (multiple or non-target images) and RT outside of the reasonable range (300ms<RT<1200ms) [202]. Additionally, we correlated the RT and SAM results to explore potential trends for how stress effects RT performance.

4.2.2 Dividing Participants into Groups

Figure 4.2 shows the results for RT average across all participants. Based on the results, we did not observe any significant difference between the experimental conditions (the only difference between the first baseline, which is expected). This lack of change is problematic as it suggests that stress did not affect the participant's behaviour. Upon further examination of the participant's behaviour, it became clear that participants exhibited two distinct behaviours. One group (Group 1) performed worse (slower RT) when stressed, while the other (Group 2) performed better or similarly (unaffected) between conditions. When we performed a median split for difference in RT between *GG* and *PH*(*GG*-***PH***), we found two groups with contradictory performances. This methodology of performing median splits based on reactionary performance was observed in other EEG studies such as Miller et al. [130] and Ros et al. [164]. Followings appropriate methodology for median splits outlined by DeCoster et al. [47] and Iacobucci et al. [93], we created an artificial categorisation within our sample forming two groups based on the reactionary performance. Other grouping options were explored, such as SAM, HR, EDA, and gait. However, the split was relatively uneven and did not find any significant difference behaviourally amongst the conditions. This participant grouping was used for the EEG related datasets. Table 4.1 presents the demographic details of the two groups.

4.2.3 Electroencephalograph (EEG)

The participant's scalp voltage was recorded through a 64 channel active electrode cap encoded by the Liveamp EEG system. The Liveamp system was chosen due to the device's portability as it is a wireless system that can support high-density channel arrays. The EEG data were collected at a 500 Hz sampling rate with all 64 channels active and synchronised through LSL.

4.2.3.1 Preprocessing

The preprocessing for the EEG data is aimed to specifically clean line noise, muscle movement, eye blinking, and other walking based artifacts. Our preprocessing pipeline was based on Mokoto's pipeline [133], Singh et al. [184], and Do et al. [55]. The data were processed on Matlab using functions based on the EEGLab toolbox by the Swartz Center for Computational Neuroscience. The walking artefacts were reduced through the use of Artifact Subspace Reconstructions (ASR) [136]. The ASR function aims to reconstruct high noise portions of the EEG data with amplitudes of the 'cleanest' or lower noise portions. We also used the AMICA independent component analysis (ICA) algorithm to compute the components within the EEG data and remove non-brain components such as eye blink and muscle artifacts. The data with the removed component is back-projected and analysed at a channel based level. The EEG data was filtered more aggressively due to the high noise levels from head movement, walking, and the complex stimuli involved.

The pipeline contained the following steps:

1. Bandpass filter 1-50Hz for Anti-Aliasing and remove high and low frequency noise,
2. Down-Sample data to 250Hz to reduce computational time,
3. Clean line noise,
4. Apply ASR,
5. Replace removed channels by interpolating from neighbouring channels,
6. Re-reference data to signal average,
7. Compute ICA weights and Spheres through AMICA,
8. Dipole fit components,
9. Classify components (IC labeling), and

10. Remove any component that is less than 85% classification for brain component,

4.2.3.2 EEG Analysis

The EEG data were first epoched around the target stimuli event. The epoch timeframe of -400 ms to 1200 ms was determined by the no stimuli time before the target appeared (-400 ms), the duration the stimuli appeared (800 ms), the next no stimuli period (400 ms) as shown in Figure 3.7. based on prior literature[75, 150, 153], we selected the central midline channels Fz, Cz, and Pz for analysis.

We assess three key metrics of EEG data to determine the effects of the stress response. Firstly, the ERP response is plotted and compared. The is a plot of the electrode amplitude over the time domain. The trials for each condition is averaged and compared across conditions. The P300 signal peaks were detected by finding the local maxima of the ERP that is between 300ms to the participant's trial reaction time. The peaks were compared by conditions. As highlighted in Table 2.4, the previous research has mixed results on stress's effect on P300 peak amplitude.

The second metric is the EEG topography of activation. The topography provides insights into whether the P300 wave is correctly stimulated. We assessed the topography at 300 - 500 ms and took the average across each group for comparison. The expected result is an activation in the parietal electrodes during the P300 period.

The last metric is the Event-Related Spectral Perturbation (ERSP). The ERSP response measures change in the frequency spectral behaviour over the time domain. The spectral behaviour can provide insights on potential reasoning for behavioural change between conditions. The brain has five main frequency oscillation types that we can observe (data filter to only 1-50 Hz), Delta (1-4 Hz), Theta (4-8 Hz), Alpha (8-12Hz), Beta (13-30Hz), and low Gamma (30-50 Hz). As mentioned in Chapter 2 Table 2.3, these frequency bands (specifically alpha and beta) can offer insights into the emotional state of an individual. The expected P300 ERSP response is shown in Figure 4.1, the onset (before) P300 peak there is an expected alpha power increase and a gradual beta power decrease during the P300 peak. To compare the frequency power bands between conditions, we extracted the power around the P300 peak (300 ms to 500 ms) and separated the power to each respective frequency band.

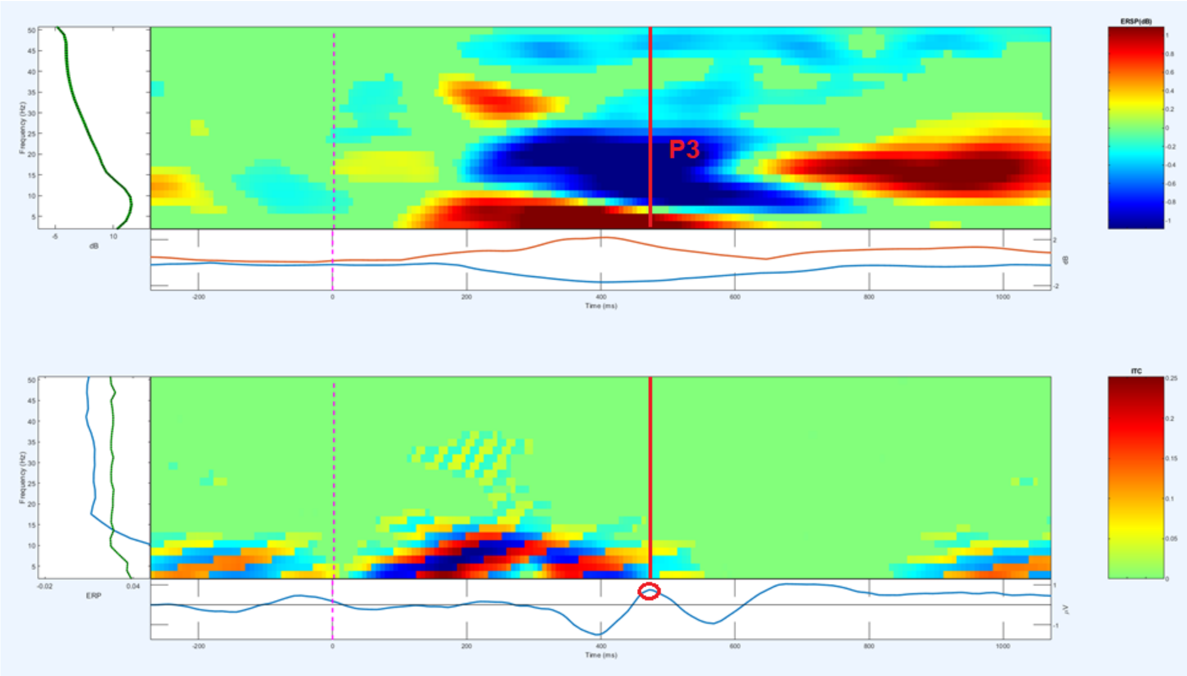

Figure 4.1: An Example of the P300 ERSP response taken from the Pz Channel ERSP response of all conditions and all participants. Significance mask was applied at α <0.01.

4.2.4 Statistical Analysis

Similar to Chapter 3, the normality of the data was determined by a one-sample Kolmogorov-Smirnov test (compared against cumulative distribution function). The RT, SAM, P3 peak, and frequency power band data were not normally distributed, leading to an applied Wilcoxon signed-rank test. The RT and SAM correlation R and P-value were calculated using the Spearman test since the data was non-parametric and monotonic. The EEG topography data was calculated through a repeated measure (each trial) Anova of each channel between conditions. The ERSP response significance mask was calculated through bootstrap permutation ($\alpha < 0.01$) when compared to the baseline, the p values were false discovery rate (FDR) corrected. The significance stars for the figures are *p<0.05, **p<0.01, and ***p<0.001.

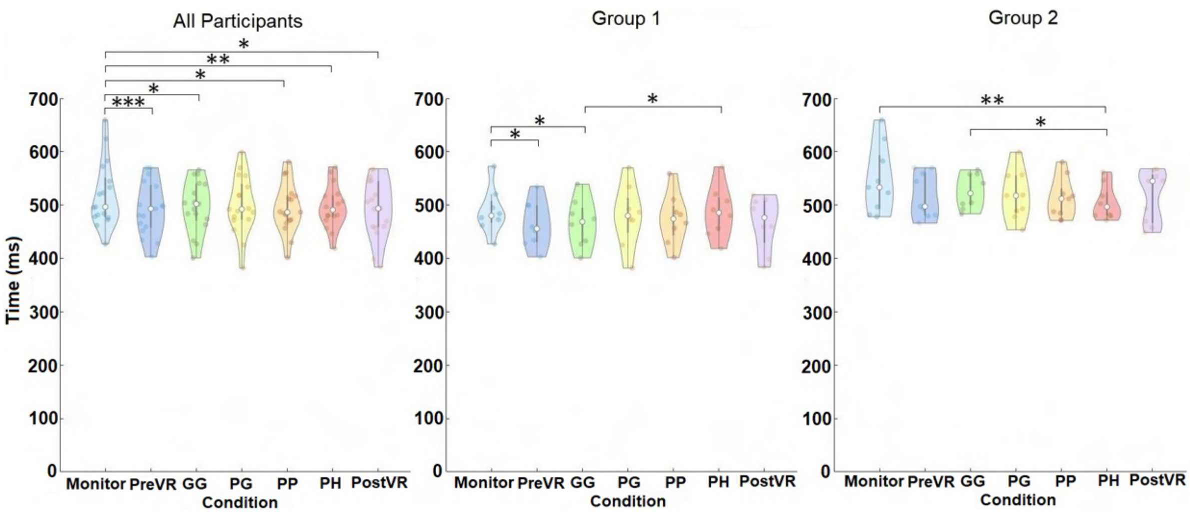

Figure 4.2: Participant's target image RT (0ms=start of visual stimuli) for each condition.

4.3 Results

4.3.1 Reactionary Performance

Figure 4.2 presents a violin plot of the RT results of all participants, Group 1, and Group 2. The RT for all participants (Monitor M=519.65 ms SD=61.04, PreVR M=492.50 ms SD=50.16, *GG* M=497.40 ms SD=48.09, *PG* M=499.82 ms SD=54.12, *PP* M=494.16 ms SD=46.19, *PH* M=495.93 ms SD=39.94, and PostVR M= 492.91 ms SD=54.87) found no significance between the walking conditions with only decrease in RT between the baseline on the computer monitor to the other conditions.

Group 1 (Monitor M=488.77 ms SD=61.04, PreVR M=463.99 ms SD=44.16, *GG* M=465.91 ms SD=44.95, *PG* M=479.42 ms SD=58.45, *PP* M=474.09 ms SD=48.18, *PH* M=486.54 ms SD=47.92, and PostVR M= 466.36 ms SD=51.42) found a significant increase in RT between *GG*and *PH*(W=0, Z=-2.52, p=0.012, and r=0.89) suggesting a slower RT when in virtual height.

Group 2 (Monitor M=547.11 ms SD=63.42, PreVR M=517.84 ms SD=42.33, *GG* M=525.40 ms SD=31.35, *PG* M=517.96 ms SD=45.62, *PP* M=512.01 ms SD=38.40, *PH* M=504.27 ms SD=31.84, and PostVR M= 516.50 ms SD=48.77) showed a significant decrease in RT between *GG*and *PH*(W=43, Z=2.43, p=0.015, and r=0.81) suggesting a faster RT when in virtual height.

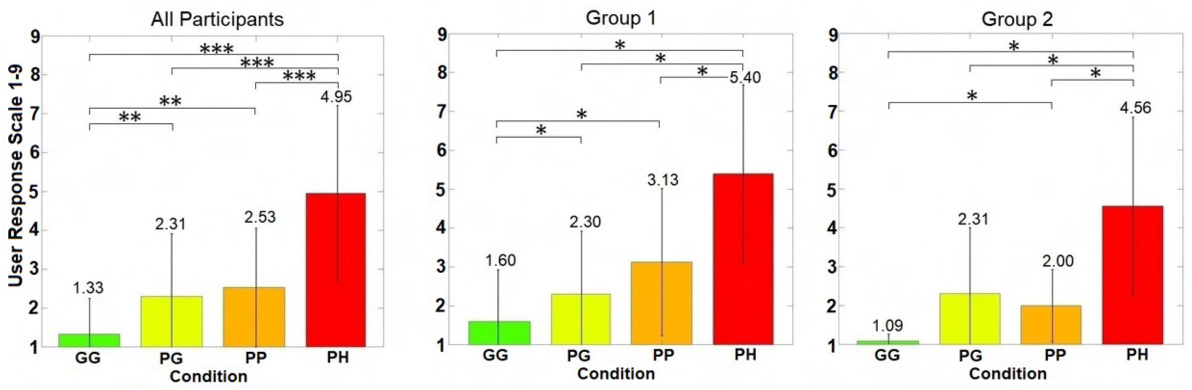

Figure 4.3: Participant's SAM responses for each condition.

4.3.2 Self Assessment Manikin

Figure 4.3 shows a plot of the SAM results for all participants, Group 1, and Group 2. The SAM response of all participants (*GG* M=1.33 and SD=0.92, *PG* M=2.31 and SD=1.60, *PP* M=2.53 and SD=1.53, and *PH* M=4.95 and SD=2.25) found that the *PH*condition were significantly different (W=0, Z<-3.30, p<0.001, and r>0.80) from all the other conditions (*PP*, *PG*, and *GG*). There is also a significant difference when comparing the *GG*condition to the *PG*condition (W=4, Z=-2.91, p=0.004, and r=0.70) and the *PP*condition (W=0, Z=-3.06, p=0.022, and r=0.74).

The SAM response for Group 1 (*GG* M=1.60 and SD=1.32, *PG* M=2.30 and SD=1.62, *PP* M=3.13 and SD=1.89, and *PH* M=5.40 and SD=2.28) found that the *PH*condition were significantly different (W=0, Z<-2.37, p<0.018, and r>0.84) from all the other conditions (*PP*, *PG*, and *GG*). There is also a significant difference when comparing the *GG*condition to the *PG*condition (W=0, Z=-0.78, p=0.027, and r=0.78) and the *PP*condition (W=0, Z=-0.78, p=0.027, and r=0.78.

The SAM response for group 2 (*GG* M=1.09 and SD=0.18, *PG* M=2.31 and SD=1.69, *PP* M=2.00 and SD=0.93, and *PH* M=4.56 and SD=2.29) found that the *PH*condition were significantly different (W=0, Z<-2.37, p<0.018, and r>0.79) from all the other conditions (*PP*, *PG*, and *GG*). There is also a significant difference when comparing the *GG*condition to the *PG*condition (W=3, Z=-1.87, p=0.062, and r=0.62) and the *PP*condition (W=0, Z=-2.21, p=0.027, and r=0.74).

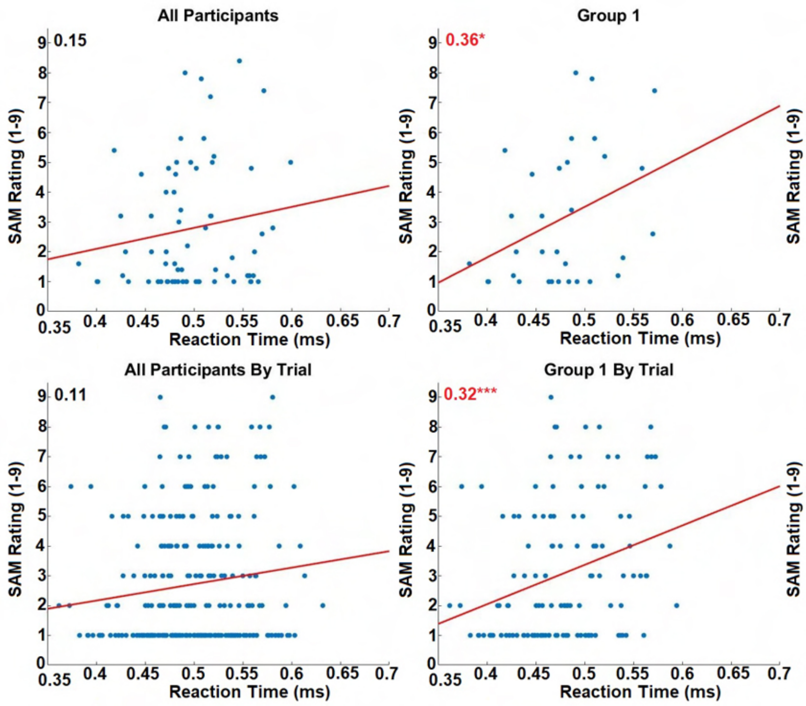

Figure 4.4: Participant's RT and SAM correlation plot. R Value is on the top left of each figure (Spearman test).

4.3.3 Correlation SAM-RT

Figure 4.4 illustrates the colleration between the SAM rating (arousal/stress) and RT for all participants, Group 1, and Group 2. No significant correlation was found in all participants (by subject R=0.15 P=0.2378 and trial R=0.11 P=0.0641) and Group (by subject R=0.05 P=7505 and trial R=0.04 P=6450) in both subject and trial levels.

A significant correlation was found in Group 1 for both subject (R=0.36 and P=0.0428) trial level (R=0.32 and P<0.001) data. The trend suggests that as SAM increases, the RT will also increase reflecting decreased performance.

4.3.4 Topography and Event-Related Potential

Figure 4.5 depicts the topography of each condition between 300 ms to 500 ms for Group 1 and 2. The topography confirms that the P300 task does activate the parietal region of the brain. We did not observe any significant difference between conditions suggesting that the region of activation is not significantly changed between stress levels.

Figure 4.7 and Figure 4.9 presents the ERP results from the Pz channel along with a comparison of the P300 peak amplitudes. When comparing the P300 amplitudes for all

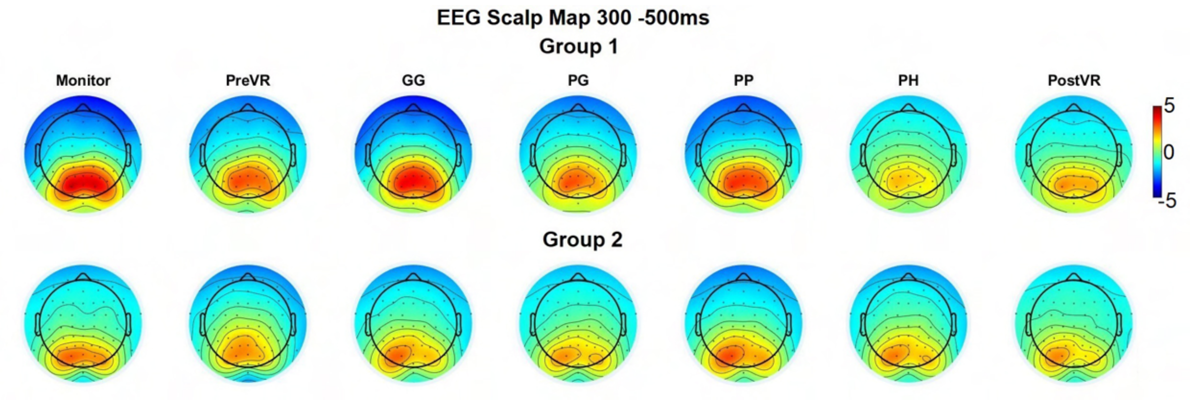

Figure 4.5: The EEG Scalp Map topography for each condition around 300 ms-500 ms (0 ms at the start of target stimuli)

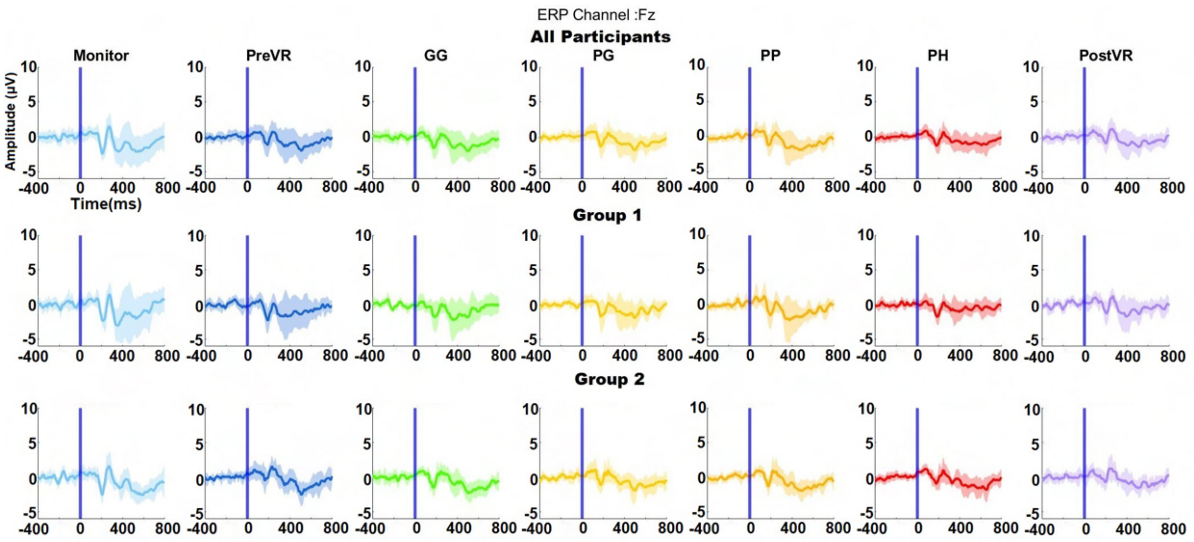

Figure 4.6: A side by side comparison of ERP response from the Fz channel. Solid line is the average ERP and the shaded area is the standard deviation.

participants (Monitor M= 5.12 μV SD=2.91, PreVR M=4.38 μV SD=1.90, *GG* M= 4.55 μV SD=3.23, *PG* M=3.93 μV SD=2.79, *PP* M=4.46 μV SD=2.41, *PH* M=3.17 μV SD=1.98, and PostVR M=3.65 μV SD=3.32) we did not observe any significant difference between the experimental conditions (only difference from baseline). The P300 amplitudes for Group 1 (Monitor M= 6.01 μV SD=3.64, PreVR M=4.77 μV SD=1.60, *GG* M= 5.67 μV SD=4.25, *PG* M=4.67 μV SD=3.30, *PP* M=5.29 μV SD=2.81, *PH* M=3.09 μV SD=1.91, and PostVR M=4.11 μV SD=4.58) found a significant difference between *GG*and *PH*(W=32, Z=1.96, p=0.049, and r=0.69). This difference indicates that the P300 amplitude decreases when at virtual heights. No significance was observed when comparing the P300 amplitudes

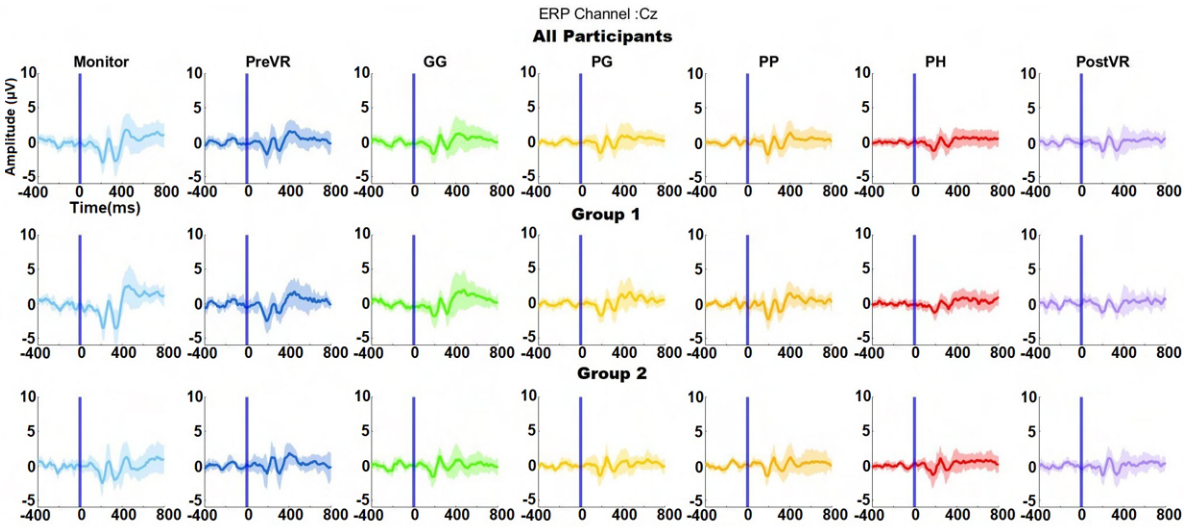

Figure 4.7: A side by side comparison of ERP response from the Cz channel. Solid line is the average ERP and the shaded area is the standard deviation.

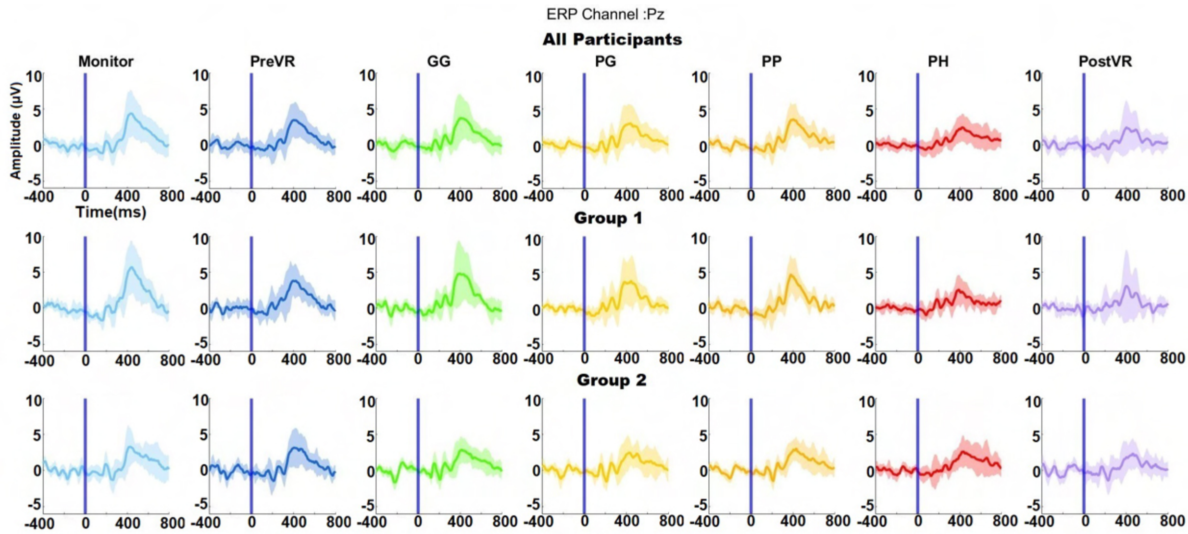

Figure 4.8: A side by side comparison of ERP response from the Pz channel. Solid line is the average ERP and the shaded area is the standard deviation.

for Group 2 (Monitor M= 5.12 μV SD=2.91, PreVR M=4.38 μV SD=1.90, *GG* M= 4.55 μV SD=3.23, *PG* M=3.93 μV SD=2.79, *PP* M=4.46 μV SD=2.41, *PH* M=3.17 μV SD=1.98, and PostVR M=3.65 μV SD=3.32). This finding suggests that there is little change in ERP response between conditions.

We also evaluated the Fz (Figure 4.6) and Cz (Figure 4.8) channel's ERP response but we did not find any significant difference within the P300 peaks.

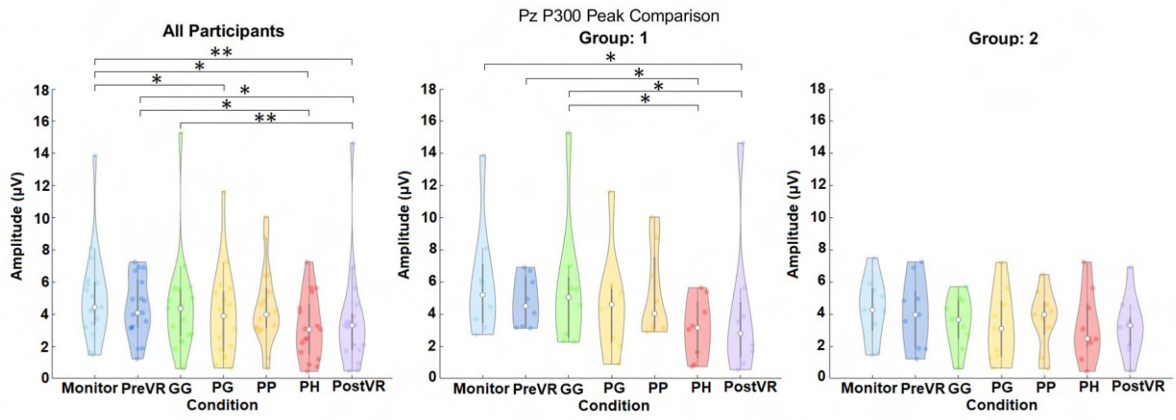

Figure 4.9: A comparison of the detected P300 peak from the Pz ERP response.

4.3.5 Event-Related Spectral Perturbation

Figure 4.10 shows the Fz channel ERSP response between Group 1 and Group 2. When comparing the power bands (Table 4.2) for Group 1, we found a significant increase in alpha between *GG* and *PH* (W=2, Z=-2.24, p=0.025, and r=0.79) and beta (W=0, Z=-2.52, p=0.012, and r=0.89). We did not find any significant spectral change in Group 2.

Figure 4.11 shows the Cz channel ERSP response between Group 1 and Group 2. For Group 1, we observed a significant increase in beta power between *GG* and *PH* (W=0, Z=-2.52, p=0.012, and r=0.89) and between *GG* and *PG* (W=3, Z=-2.10, p=0.036, and r=0.74). There was an increase in alpha power between *GG* and *PH* (W=1, Z=-2.38, p=0.017, and r=0.84) and between *GG* and *PG* (W=0, Z=-2.52, p=0.012, and r=0.89). We also observed an increase in Theta power between *GG* and *PP* (W=3, Z=-2.10, p=0.036, and r=0.74). Group 2 found an increase in gamma power between *GG* and *PH* (W=0, Z=-2.52, p=0.012, and r=0.89) and between *GG* and *PP* (W=3, Z=-2.10, p=0.036, and r=0.74). In Group 2, we found increase in alpha power between *GG* and *PH* (W=3, Z=-2.31, p=0.021, and r=0.77).

Figure 4.12 shows the Pz channel ERSP response between Group 1 and Group 2. For Group 1, we observed a significant increase in beta power between *GG* and *PH* (W=2, Z=-2.24, p=0.025, and r=0.79) and between *PG* and *PH* (W=3, Z=-2.10, p=0.035, and r=0.74). We also observed an increase in theta power between *GG* and *PP* (W=3, Z=-2.10, p=0.036, and r=0.74). Group 2 found an increase in gamma power between *GG* and *PH* (W=5, Z=-2.07, p=0.038, and r=0.69).

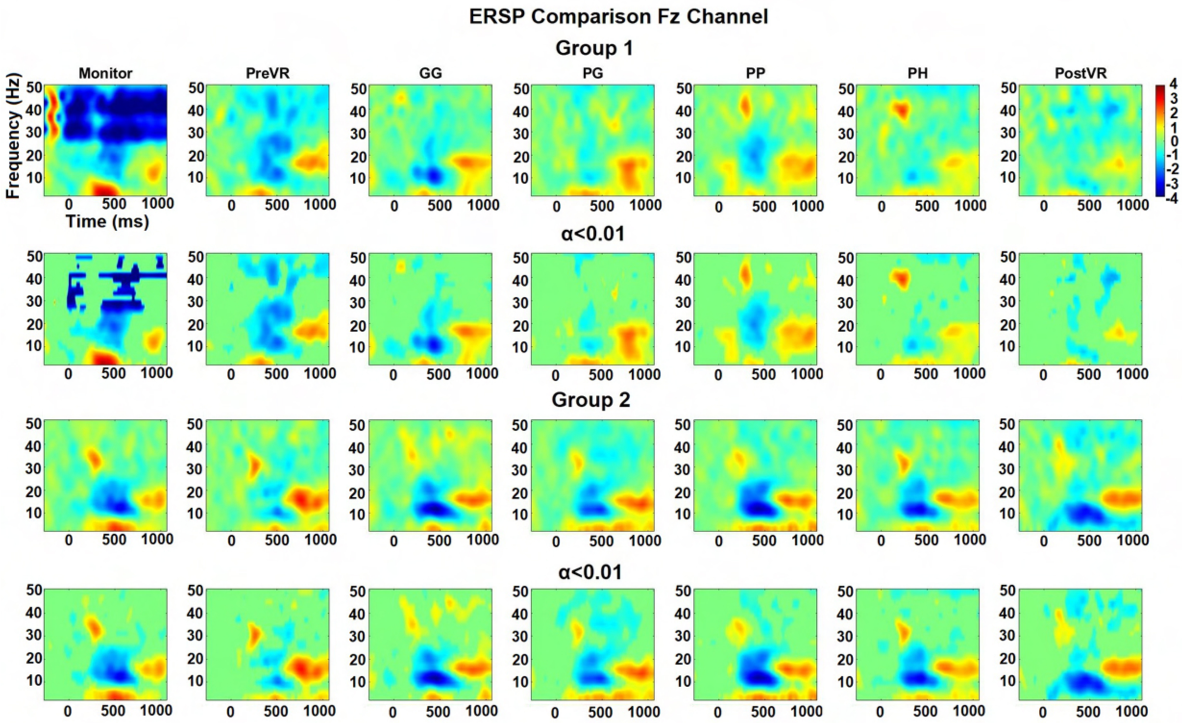

Figure 4.10: The ERSP response for the Fz channel for each group and condition. The plots for the second and fourth row are the significant area on the ERSP above with the criteria of $\alpha < 0.01$ with FDR correction.

4.4 Discussion

4.4.1 Validity of Stress and P300 Response

There are two key factors for the validity of the investigation into the effects of stress on cognitive performance (***RQ2***). Firstly, we must ensure that the experimental conditions have correctly elicited a stress state. This challenge was highlighted in Peterson et al. [147], where the study found that it was the wearing of a VR HMD rather than the VE that caused the stress response and behavioural change. Based on the previous results in Chapter 3 and the SAM results for Group 1 and Group 2 (Figure 4.3), it is apparent that the new sample size (removed one participant) and the group median split did not alter the stress response behaviour between conditions. This is evident with the SAM result exhibiting a consistent trend from Chapter 3, with the stress levels for each condition ranked as *PH>PP=PG>GG*.

The second factor is whether or not the oddball paradigm has correctly produced a P300 wave response. The P300 wave is distinct in its parietal region activation and

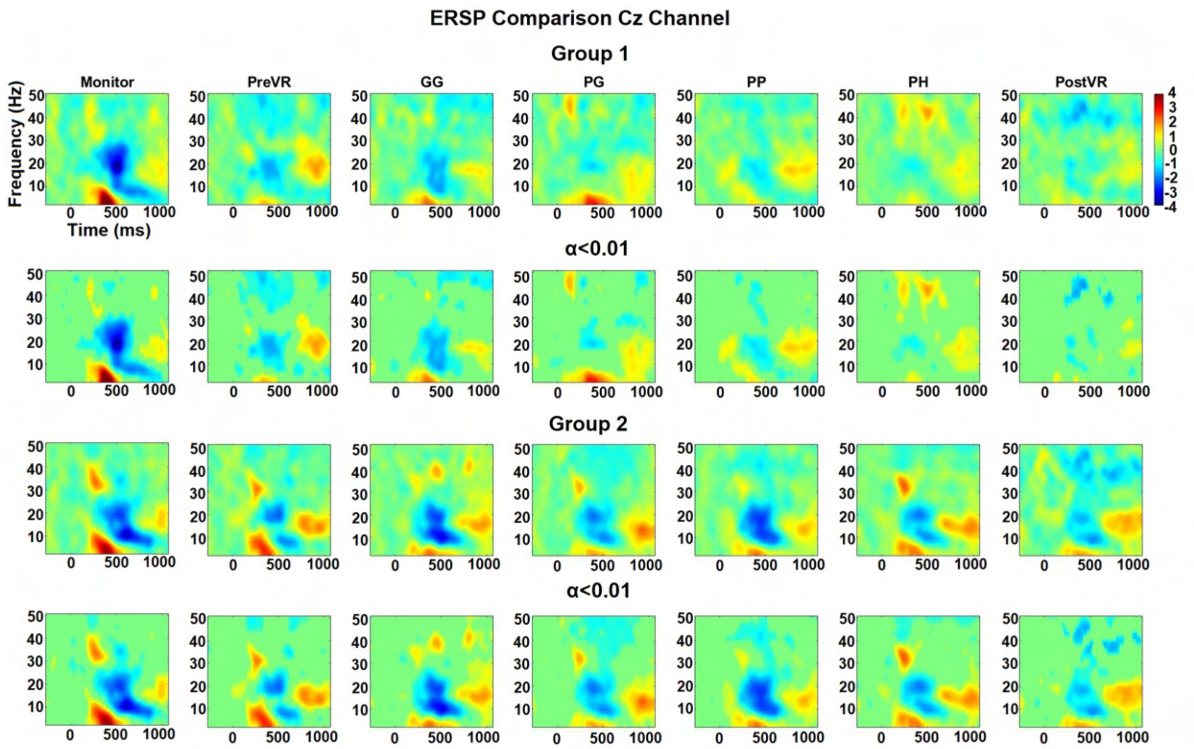

Figure 4.11: The ERSP response for the Cz channel for each group and condition. The plots for the second and fourth row are the significant area on the ERSP above with the criteria of $\alpha < 0.01$ with FDR correction.

distinct ERP wave peak [152]. The EEG scalp topography (Figure 4.5) corroborates this result as a parietal activation can be observed in all the experiment conditions and baselines. The P300 waveform can also be observed in the ERP responses in Figure 4.7. The waveform conforms to the traditional shape with clear identifiable P100, P200, and P300 peaks [154] within all the conditions.

4.4.2 Relation between Stress and P300

No significance could be found when observing the RT and P300 ERP responses on the sample group level (All Participants). This result would have suggested that stress does not change the behaviour of P300 and level cognitive performance. However, if we observe the data from the group level perspective, we find two starkly different behaviours. As outlined in the SAM results, both Group 1 and Group 2 experience a stress response as the height increases. Both groups also demonstrate a P300 wave from the ERP and topography. The RT data (Figure 4.2) provides insights into the output performance of

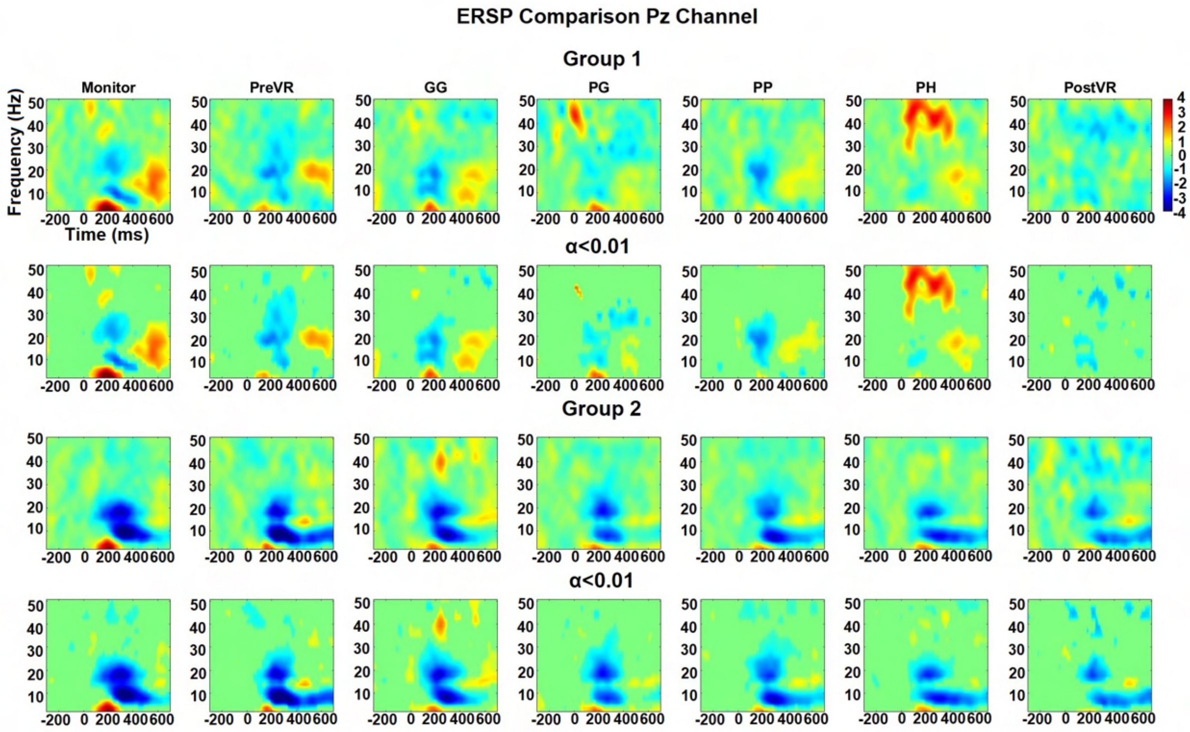

Figure 4.12: The ERSP response for the Pz channel for each group and condition. The plots for the second and fourth row are the significant area on the ERSP above with the criteria of $\alpha < 0.01$ with FDR correction.

both groups, while the ERP response reveals the cognitive behaviour before the output response.

Interestingly, during the stressful condition (*PH*, Group 1's performance significantly deteriorates (slower RT), conversely Group 2's performance improves (faster RT) when stressed. The correlation of SAM to the RT data provides a clue into the relationship between stress level and the reactionary performance for each group. Group 1 found a strong correlation (see Figure 4.4) between the SAM and RT data. Based on the linear fitting, it can be observed that as the stress level increased, the participants would gradually perform worse. Group 2 did not demonstrate any correlation between SAM and RT, which suggests that their reactionary performance was not affected by stress.

The ERP response and P300 peak data corroborate findings of the RT data for both Group 1 and Group 2. It is evident that Group 1's P300 peak amplitude significantly decreases when stressed. This outcome signifies that stress does indeed decrease the amplitude of the P300 peak. However, Group 2 displays a distinct consistency level in ERP and P300 amplitude for the oddball task. Incidentally, the results of either

Table 4.2: The ERSP Frequency Power band data extracted around the P300 period for Group 1 and Group 2.

	Delta (dB)	**Theta (dB)**	**Alpha (dB)**	**Beta (dB)**	**Gamma** (dB)
			Fz Group 1		
GG	1.28±0.82	-0.78±1.53	-3.14±3.03	-1.31±0.96	-0.41±0.57
PG	1.29±0.81	-0.28±1.45	-1.72±2.43	-0.67±0.76	-0.12±0.53
PP	1.76±2.11	0±1.36	-2.15±2.26	-1.33±1.18	-0.01±0.69
PH	0.91±0.97	-0.56±1.45	-1.92±2.38	-0.57±0.77	-0.05±0.26
			Fz Group 2		
	Delta (dB)	**Theta (dB)**	**Alpha (dB)**	**Beta (dB)**	**Gamma** (dB)
GG	1.53±1.56	-0.27±2.19	-3.05±2.82	-1.36±0.86	-0.12±0.88
PG	1.33±0.86	-0.22±0.99	-2.19±2.18	-1.15±0.55	-0.62±0.53
PP	1.53±1.06	-0.38±1.38	-2.96±2.07	-1.48±0.46	-0.40±0.70
PH	1.33±0.56	-0.21±1.08	-2.53±2.37	-1.18±0.80	-0.14±0.38
			Cz Group 1		
	Delta (dB)	**Theta (dB)**	**Alpha (dB)**	**Beta (dB)**	**Gamma** (dB)
GG	1.62±094	-0.33±1.03	-2.22±2.90	-1.18±0.84	0±0.74
PG	1.81±1.57	0.47±1.66	-0.69±2.66	-0.72±0.62	-0.02±0.63
PP	1.68±1.52	0.27±1.10	-1.40±2.52	-0.91±1.23	-0.45±0.84
PH	0.66±0.84	0.12±1.33	-0.68±2.12	-0.26±0.29	0.58±0.96
			Cz Group 2		
	Delta (dB)	**Theta (dB)**	**Alpha (dB)**	**Beta (dB)**	**Gamma** (dB)
GG	1.51±1.56	-0.62±2.41	-2.92±3.07	-1.54±0.94	-0.15±1.28
PG	1.22±1.13	-0.13±1.30	-1.75±2.08	-1.21±0.73	-0.54±0.43
PP	1.47±1.40	-0.29±1.88	-2.42±2.57	-1.65±0.94	-0.46±0.65
PH	1.40±0.92	0.10±1.55	-1.90±2.65	-1.14±0.97	-0.08±0.46
			Pz Group 1		
	Delta (dB)	**Theta (dB)**	**Alpha (dB)**	**Beta (dB)**	**Gamma** (dB)
GG	1.59±0.79	-0.53±1.48	-2.28±3.20	-1.15±0.70	-0.02±0.63
PG	1.62±1.38	-0.29±1.82	-1.44±2.71	-0.93±0.51	-0.05±0.69
PP	2.32±1.69	0.42±1.48	-1.26±1.94	-1.13±0.75	-0.10±0.81
PH	1.23±1.38	-0.22±1.38	-1.26±2.12	-0.43±0.71	0.51±1.50
			Pz Group 2		
	Delta (dB)	**Theta (dB)**	**Alpha (dB)**	**Beta (dB)**	**Gamma** (dB)
GG	1.68±1.43	-0.52±2.16	-2.16±2.25	-1.46±0.69	0.07±1.28
PG	1.71±1.28	-0.20±1.93	-1.40±1.88	-1.26±0.88	-0.42±0.68
PP	1.63±1.43	-0.54±1.76	-1.95±1.81	-1.41±0.74	-0.38±0.74
PH	1.52±1.01	-0.61±2.08	-1.94±2.75	-1.03±1.30	-0.05±0.48

Group correlates with past literature's as previously emphasised in Table 2.4. Previous studies such as Jiang et al. [97] and Mingming et al. [159] reasoned that Group 1's ERP behaviour is due to the stress response acting as a distracter which reduces the participant's attention resources engaged. This explanation is consistent with established worked that affirm the relation between P300 amplitude and attention [76]. On the other hand, this rationale does not explain the seemingly contradicting results of Group 2, which correlated other studies [54, 69, 158]. The studies that support Group 2 rationalise that stress increases an individual's alertness, improving performance and increasing or maintaining consistent P300 peak amplitude.

Both of these explanations are further upheld by the ERSP and powerband results (Figure 4.10, Figure 4.12, and Table 4.2). Group 2 exhibited a remarkably consistent spectral behaviour that is typical of normal P300 behaviour. This consistency would suggest an undisrupted mental state. The same consistency is not observed in Group 1. The ERSP response illustrates a clear inconsistent spectral activity with the Beta power significant increasing (Fz and Pz) during the *PH* condition. Based on prior research, an increase in Beta power may be an indication of a change in mental activity and resource allocation [88]. Schmidt et al. [169] propose that frontal-central beta is closely linked to memory resource allocation and thought processing. This finding may indicate a change in the neurological pathway for target recognition when stressed in which Group 1 would rely more heavily on memory mental resources than Group 2.

In summary, the two groups we found exhibited seemingly contradictory yet rational and supported by literature, P300 behaviours. An elementary approach would be to dismiss the results of a group and support the other. However, the validity of past works and results from this experiment does not provide sufficient evidence to dismiss either behaviour. We argue that both behaviours can be true when within the proper circumstance or subject types within our sample.

4.4.3 Validity and Demographic Driven Explanations

In order to rationalise both behaviours, we must analyse and identify the potential circumstances either behaviour might occur. Kamp et al. [100] identified two factors that may influence the experience of physiological stress and P300 behaviour. The first factor was proposed by Kamp et al. [100], which was the effectiveness of the stressor to induce adequate levels of stress. The paper suggested that TSST may have been more effective at eliciting stress, and thus the decrease in P300 amplitude is the more likely behaviour. However, in the case of the Dierolf et al. [54] study, TSST yields a contradictory result (increase in P300 amplitude). Hypothetically, if there is a difference between Group 1 and 2 due to the effectiveness of the stressor, then there should be an observable and significant difference in the SAM response both between conditions and between groups. This was not found in Figure 4.3, since both Group 1 and Group 2 exhibited signs of stress and unanimously ranked the *PH* as most stressful (Figure 3.10). Nevertheless, there is a subjectivity to the experience of stress. Group 1 and Group 2 participants can vary in the degree of stress/arousal while still being stressed.

The second factor is the participant's demographics, specifically age and sex [100, 122, 198]. Both age and sex can greatly affect the perception of stress and P300 amplitude. Table 4.1 presents the age and sex distribution between the two groups. We found no notable difference in demographics. Therefore, the participant's demographics are unlikely to explain the difference between Group 1 and Group 2.

4.4.4 Yerkes-Dodson Law

We propose to explain the difference between Group 1 and Group 2 through concepts derived from the YD law. The YD law asserts that at optimum levels of stress, a performance improvement is observed and at higher/extremities, a performance deterioration occurs [194]. Researchers have also found that the difficulty or perceived effort of a task will affect the YD curve shape [21]. As illustrated in the traditional YD Model plot in Figure 4.13, difficult tasks that require more cognitive resources often require less/low stress to reach optimal performance and then gradually decrease as stress increases. Simpler or easier tasks that can be performed more autonomically tend to have a more gradual increase in performance as stress increases and requires a much higher/extreme level of stress to cause a drop in performance. Based on the YD law, we propose two hypothetical explanations (see Figure 4.13).

YD Hypothetical 1 (Figure 4.13) takes the assumption that Group 1 and Group 2 has experienced a different subjective level of stress. Then, one can argue that Group 2 experienced a more optimal level of stress, and thus they are at the near-optimal performance level and is yet to suffer the performance decrease. This trend can be observed by the improved performance (faster RT) in the *PH* condition and the consistent ERP and ERSP response. Conversely, Group 1 may have experienced a higher level of stress, thus passing the optimum levels with the increase in stress decreasing performance and P300 amplitude (similar to the correlation in Figure 4.4). This hypothetical is problematic to find evidence within this study to support this claim. The stress metric used in this study does not indicate any difference in the level of stress. Therefore, the stress level for this hypothetical is mainly inferred based on the RT performance and the stress level implications made by Kamp et al. [100].

Our second hypothetical (YD Hypothetical 2, see Figure 4.13) takes the position that stress has affected the perception of the difficulty and mental resource allocation of the task, thus causing the division between Group 1 and Group 2. A similar concept was explored by Sellers et al. [174] who found that the degree of difficulty of the P300 task would affect the P300 amplitude (higher difficulty caused decreased amplitude). Based on the ERP and ERSP result, it is apparent that Group 1 has allocated more mental memory resources (Beta increase) and is more distracted by the stress (less attention resource shown by the decreased P300 amplitude [97, 159]) compared to Group 2, who remained unaffected. Therefore, as the difficulty of the task changes, the subject's stress and performance curve would also need to be adjusted. In the case of Group 1, the change in brain dynamics signifies an increase in difficulty. Thus, indicating that

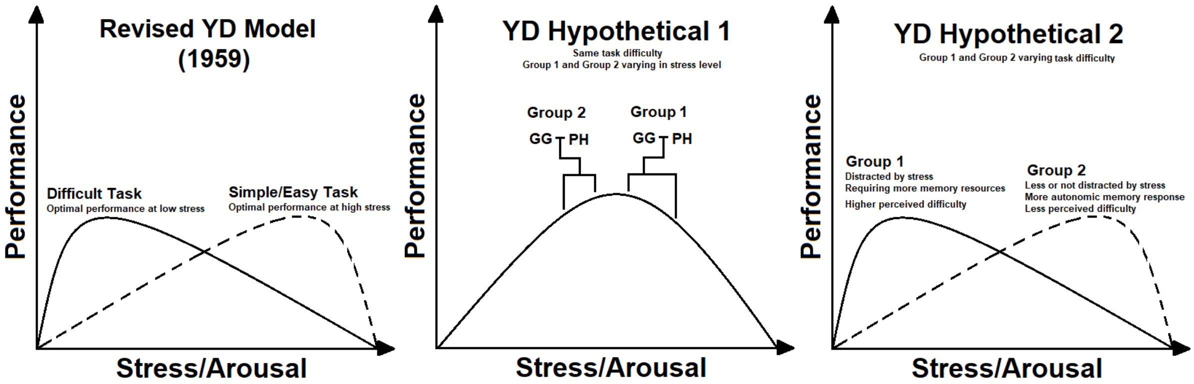

Figure 4.13: A generalisation of the tradition YD law curve from literature [21, 194] and two hypothetical YD law based explanations

they are more susceptible to decreased performance at a lower threshold of stress. When consistent brain dynamics is observed (in the case of Group 2), it can be assumed that the participants who are not experiencing significant difficulties with the task is likely to be less susceptible or even perform better when stressed.

Overall, both hypotheticals can reconcile the conflict in previous works as we have found both results to be true in different subjects. We believe the YD Hypothetical 2 is better supported by the results. This explanation suggests that we can use the RT performance and brain behaviour to interpret the effort/difficulty of a task. By recognising the difficulty of a task, we can gauge and estimate how stress will affect their performance in certain circumstances. Our strategy differs from previous research by metricising and classifying individual behaviour types rather than suggesting a population-based trend. By metricising the beta powers, we can observe how stress affects people on an individual level. Future BCI systems can adapt the behaviour accordingly, e.g. a BCI P300 classify may adjust the threshold amplitude (expecting lower amplitudes) when detecting a stress response, RT performance decrease, and beta power increase. On the other hand, no classifier adjustments is necessary if no power change is detected.

4.5 Limitation

4.5.1 Group Sample Size

Overall, this experiment contains a low sample size and thus may not accurately reflect the population's behaviour. The target sample for this study has initially been forty participants. However, due to the length of the experiment (2-3 hours) and the restrictions during the COVID pandemic, we were only able to recruit twenty participants and further excluding three due to technical (hardware failure), behaviour (not correctly completing the oddball task), and early termination (the *PH* conditions being too stressful). The seventeen participants were split into two groups which further reduces the sample size. Hence, we do not want to speculate into the population claims but rather primarily assess the difference between Group 1 and Group 2. In the future, we hope to collect a larger sample group with broader demographics to assess these results' reproducibility better.

4.5.2 Unbalanced Factors

Another key limitation is the unbalanced size, condition order, and demographics of Group 1 and 2. Ultimately, this split grouping was a serendipitous discovery during analysis. The meant that the size, condition order, and demographics were not matched. Unfortunately, this group split could not have been anticipated as prior studies yielded a singular behaviour, and pilot testing did not show this divergent behaviour. The unbalanced nature may hide factors such as fatigue, habituation, and demographic effects from our data analysis. Futures studies should consider classifying participant behaviour and ensure counterbalancing of conditions.

4.6 Key Points

The findings of our EEG results does not singularly affirm one side of the divide in the literature on the P300 behaviour when stressed. Instead, we have found both behaviours that have been previously reported. At first glance, these behaviours seem contradictory. However, we believe the two groups reconcile the conflicting reports of previous works (Table 2.4). By assessing the difference between the groups, we found that the effect of stress on P300 behaviour (***RQ2***) is a function of the participant's perceived difficulty, memory, and attentional resource allocation rather than a definitive causal relationship. Subsequent works on stress and BCI should classify participants by the cognitive brain dynamics to account for the effects of stress.

CHAPTER 5

DRONE COLLISION EXPOSURE EXPERIMENT

5.1 Overview

This chapter continues the exploration of novel methods of eliciting a stress response (***RQ1***). We propose a new terminology (Prolonged Stimuli Stressor (PSS) and Dynamic Stimuli Stressor (DSS)) to better differentiate between the stressor in Chapter 2 and the stressor in this chapter's experiment. The critical difference is the nature of the stimuli with previous works using prolonged exposure to the stimuli (continuous social anxiety or virtual height). Our drone experiment falls under the DSS label because the stimuli are instantaneous and do not rely on continuous exposure.

Here, we present our unique drone experiment's background, design rationale, methodology, and stress-related results. The experiment is centred around the human response to drone collisions. We created an environment that exposed individuals to simulated drone collision events and measured their stress through questionnaires (SAM, DASS, and Threat scales), HR, HRV, and EDA. Twenty participants experienced first-hand exposure to a drone collision (or near-collision) and non-collision events. This specific form of stressor has not been previously investigated and yields unique findings to stress research. We hope to formulate an informed approach to incorporating BCI into drones controls in society from these reflections. The methodology and some results are covered in the Zhu et al. [217] paper. This work was done with the assistance of two undergraduate capstone students, Eirene Margaret Magsino, and Sanjid Mahmood Hamim. The capstone students primarily assisted in piloting the drone and setting up the experiment during data collection.

5.2 Motivation

Amusingly, the primary inspiration for this drone collision experiment came from a drone-related incident in 2019. During the capstone showcase that year, a student was presenting a drone BCI system (described in Chapter 7) that we developed. While presenting, the student accidentally leaned on the controller, which activated the drone. The drone proceeded to collide with a fellow student. Thankfully no one was seriously injured. Even though this incident occurred within seconds, the noise and attraction of the drone caused an alert reaction within the room. Other presenters, judges, and attendees would disengage their conversations/activities to watch (and potentially avoid) the drone. So as the old saying about lemons and lemonade took our findings of the incident and created a controlled experiment to observe this effect better.

Drones, in particular, are an emerging technology that has developed exponentially over time. We believe that one day, drones have become a more ubiquitous part of society. However, in the present form, drones are considered more of an unusual sighting, with many people still cautious about them [41, 56, 66]. The relationship between humans and drones is a critical factor in the acceptance and integration of drones into society [95]. We believe drone collisions is a key point of caution when a drone is sighted. This could be due to the amount of notable drone-related accidents, a high potential for severe bodily harm, and significant property damage [5, 203]. This reputation can lead to distrust, and a cautious attitude toward drone integration [95]. Other common factors that affect a person's acceptance of drones are privacy concerns [56, 66], and noise pollution [89]. This paper presents a study of 20 participants who experienced a controlled level of exposure to a drone collision (or near-collision) event. From these reflections, we hope to understand better whether drone collisions can elicit a stress response and how this would affect drones' real-world integration.

5.3 Background

5.3.1 Drone Incidents and Perception

Drone-related incidents and other adverse events can shape the perception of drones being dangerous. A report by the U.S. Federal Aviation Administration (FAA) found that there are around 300-400 drone-related incidents per year [3]. These incidents include drones colliding with people, property, and other aircraft (commercial airliners, hot air balloons, and helicopters). A notable incident is the 2016 alleged drone to commercial airliner (Airbus 320) collision at the Heathrow Airport [43]. This incident was one of the first major drone-related incidents in the U.K [43, 203]. The extensive media coverage emphasised the dangers and law restrictions that apply to drones. The following incident two years later (Gatwick Airport) further strained the public's view on consumer drones. The alleged drone sighting caused a two-day closure of the airport resulting in over a thousand affected flights and two arrests (later considered wrongful arrests) [78]. The incident drew international attention and government regulatory bodies proactively instituting harsher restrictions on drone flights [113]. Currently, flying a drone near restricted airspace is not only a criminal offence with severe legal penalties but also a potential offence classed under terrorism which carries a maximum penalty of life imprisonment [40]. Other incidents such as the injury of the Australian triathlete by a drone in 2014 [42], and the numerous occasions of a drone breaching privacy laws [41, 56, 66] has caused a negative or cautious reputation to the ordinary citizen. Not all people may have heard or recalled these specific incidences; the ramification and public perception label drones as a dangerous hobby for consumer use. With the increased market advertising and more detailed regulations, drones will likely improve public perception and become a ubiquitous technology.

5.3.2 Related Drone Human Interaction

A variety of studies have investigated the human interaction with drones flying in proximity to the person. Abtahi et al. [2, 30] investigated the willingness of participants to touch and hold a drone. Participants were asked to hold two drone designs with safety propeller guards and the other without guards. Participants performed various proximity interaction tasks with the drones, which required touching and holding the drone. The study found drones without safety equipment to be more intimidating or threatening. A similar study performed by Yeh et al. [207] finding that the appearance of safeness for a drone directly contributes to the acceptable distance between a person and a drone [207]. Our study explores the human response to being exposed to a drone flying in a more aggressive stance (flying towards collision). These studies suggest that participants are willing to directly interact with a drone when the drone exhibits safe behaviour and appearance.

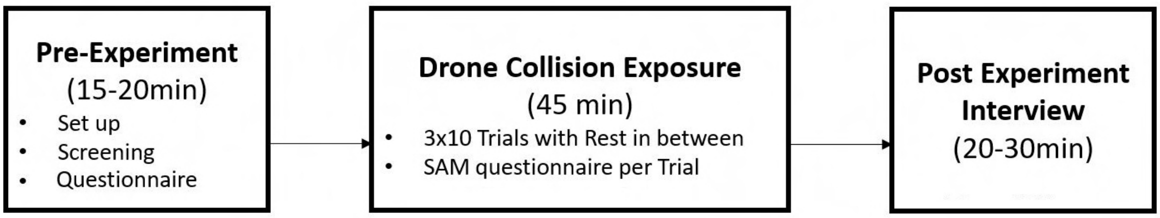

Figure 5.1: The timeline of the drone experiment with the timing and experiment tasks.

5.4 Methodology

5.4.1 Experiment Overview

This experiment's primary goals are to investigate if drone collisions can elicit a stress response and the source of the drone threat (***RQ1***). We flew the drone directly at participants (standing behind a safety net) using an available consumer drone. The drone would either collide into the net or stop mid-flight. The experiment consisted of three stages (see Figure 5.1), the pre-experiment (setup and questionnaire, 15-20 min), drone collision exposure (3x10 trials, 45 min), and post-experiment interview (20-30 min). We collected questionnaire response, reaction response, HR, HRV, EDA, and EEG data from the participant. Our post-experiment interview provides valuable qualitative insight into the participant's opinions and perceptions of the drone.

5.4.2 Drone Collision Exposure Protocol

The drone collision exposure provides valuable first-hand drone exposure to the participant. Figure 5.2A illustrates the two experimental conditions for drone exposure. Based on our established criteria for DSS, the collision condition is the intended stressor, while the non-collision condition is the neutral event. Each participant experienced 3 sets of 10 trials of either a drone collision or a non-collision event. The conditions were pseudo-randomised to reduce the drone's predictability and prevent acclimation to the collisions. The two conditions were distributed at a 1:1 ratio (15 collisions and 15 non-collision) and randomised with a min-max of 3 - 7 collision/non-collision trials per block.

During the experiment, the participant was positioned 0.3m behind a safety net to prevent the drone from physically harming the participant. The distance was determined through repeated testing with motion capture (Optitrack) to measure the net's

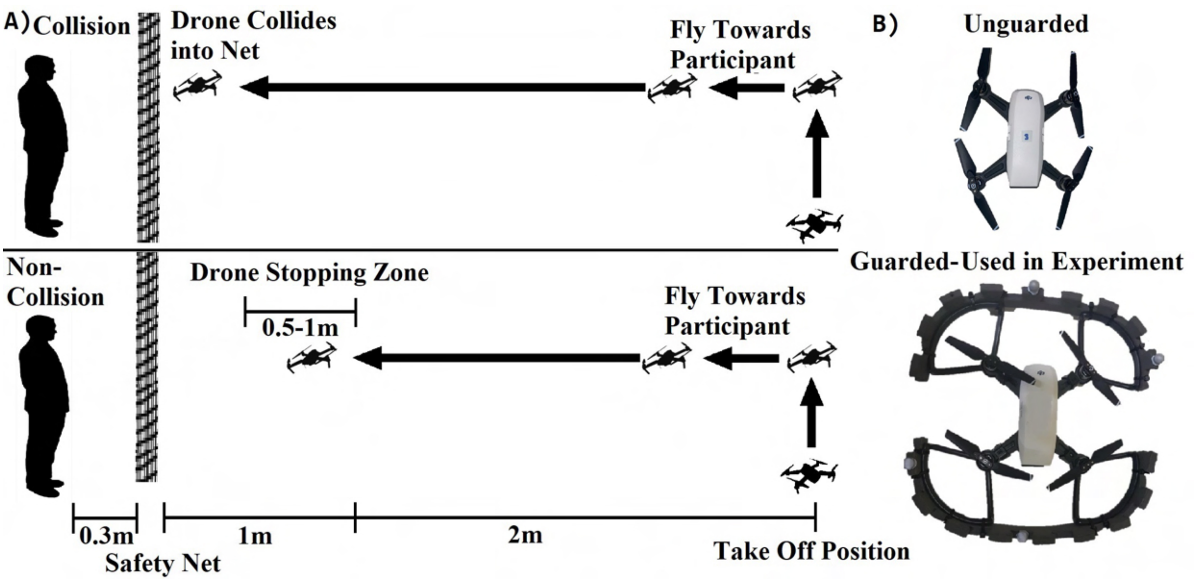

Figure 5.2: A) An outline of the two conditions (Collision vs Non-Collision) and how the experiment was performed and Conditions B) A comparison between regular DJI Spark (used in interview) and the modified version (used during exposure and interview)

displacement (measured 0.25m) during drone collision. The participant was asked to stand (seated during rests) during the experimental trials. The drone will take off 3m away from the net for each trial and fly toward the participant's upper body (shoulders to the head area). The non-collision condition involves the drone stopping/hovering at the 0.5-1m (stopping variance) areas in front of the net (0.8-1.3m to the participant). During the collision condition, the drone will not stop and will collide into the net. Participants were given a minimum of 3-minute rest breaks between every ten trials.

5.4.3 Apparatus

We chose the DJI Spark quadcopter as the drone for the experiment [210]. DJI Spark drones are popularity consumer drone (DJI reported to have around 70% U.S. market shares in 2019 [171]) that is simple and easy to use. The DJI Spark can be considered a typical drone with recognisable features for the participants. The drone was modified with propeller guards (see Figure 5.2B) to protect both the participant and the drone (prevent propellers from getting caught in the net) during collision events. Motion capture reflective markers were placed on mounts integrated into the propeller guards. The drone was manually piloted through an android flight app (DJI Go) by a researcher. The speed of the drone was set to 2.5m/s [210]. Figure 5.3 shows the experimental set-up in the

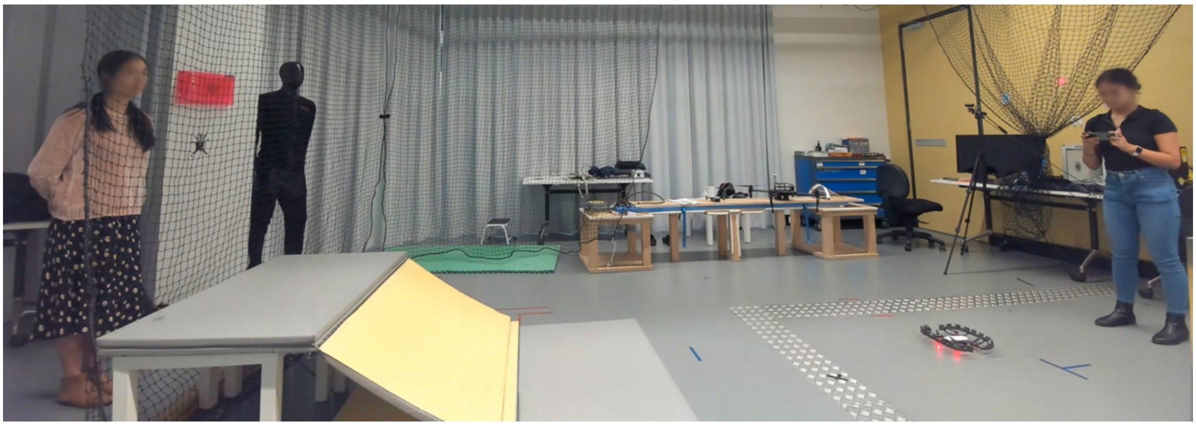

Figure 5.3: The laboratory setting for the drone experiment. (Photo was taken a photo-shoot, the researchers and participants wore protective equipment during the pandemic period)

laboratory space.

The drone and net were actively tracked through the Optitrack motion capture system. We synchronised the mocap data through LSL to provide event markers for the drone's displacement from the net at take-off, 3m (hovering), 2m, 1m, and during the collision. The measuring devices used in this experiment is identical to the setup outlined in Chapter 3. Using the same set up we collected HR/HRV (bioharness), EDA (wristband), and EEG data (64 channel scalp electrodes). The experiment and interviews were video and voice recordings. The experiment was conducted during the pandemic. Therefore extra sanitary and protective measures were taken when interacting with the participants.

5.4.4 Measurements and Analysis

5.4.4.1 Self Assessment Manikin (SAM))

At the end of each trial, the participant provided a verbal SAM rating (1-9) on their current arousal (mental activeness), valance (positive and negative emotions), and dominance (control of environment) level [20]. The full SAM questionnaire was used because this experiment was much shorter than the one presented in Chapter 3. The SAM analysis involved separating the data into the two conditions and averaging across the participants.

5.4.4.2 Drone Threat Perception

During each rest break, the participants were asked to provide a rating (0 to 3) for their perceived danger level for the drone's visual, sound, and collision. The sound and visual questions evaluated the drone at three stages: drone take-off (seeing and hearing the propellers start), flight (seeing the drone fly towards them and increasing loudness of the drone), and collision factor (loudness and visually dangerous). The responses were averaged compared between visual and sound threats. The participant also directly rates the perceived threats for collision vs non-collision.

5.4.4.3 Post Experiment Interview

The post-exposure interview further investigated the perceived hazards and threats from the drone. We aimed to identify the key concerns from the participants and whether the participants felt acclimatised to the drone after the collision exposure. The participant was given two drones to hold (with and without propeller guards, see Figure 5.2B), They were asked to observe and highly any notable features while holding the two drones.

5.4.4.4 Heart Rate (HR), Heart Rate Variability (HRV), and Electrodermal Activity (EDA)

The data analysis process for HR, HRV, and EDA was similar to the process outlined in Chapter 3. We firstly filtered out abnormal values outside the human range (30<BPM<120, 20 ms<HRV<110 ms, and $0<\mu s<20$), and removed low SNR confidence data points (for bioharness). The data were normalised per participant by feature scaling, which included the resting period.

5.4.4.5 Statistical Analysis

Similar to the previous study, we calculated our sample size using Wilcoxon signed rank t test with standard large effect size of 0.8 and an error α of 0.05. Determining that the sample size of 20 was sufficient for observing the effect of the study. The one-sample Kolmogorov-Smirnov test was used to determine the normality of our data. We found the HR, HRV, EDA, and threat response questions to be of normal distribution. Therefore, we applied pairwise a One-Way ANOVA test. The SAM data was tested through a pairwise Wilcoxon signed-rank test. The significance criteria were set to (α) of 0.05 for determining statistical significance. The significance stars for the figures are *p<0.05, **p<0.01, and ***p<0.001.

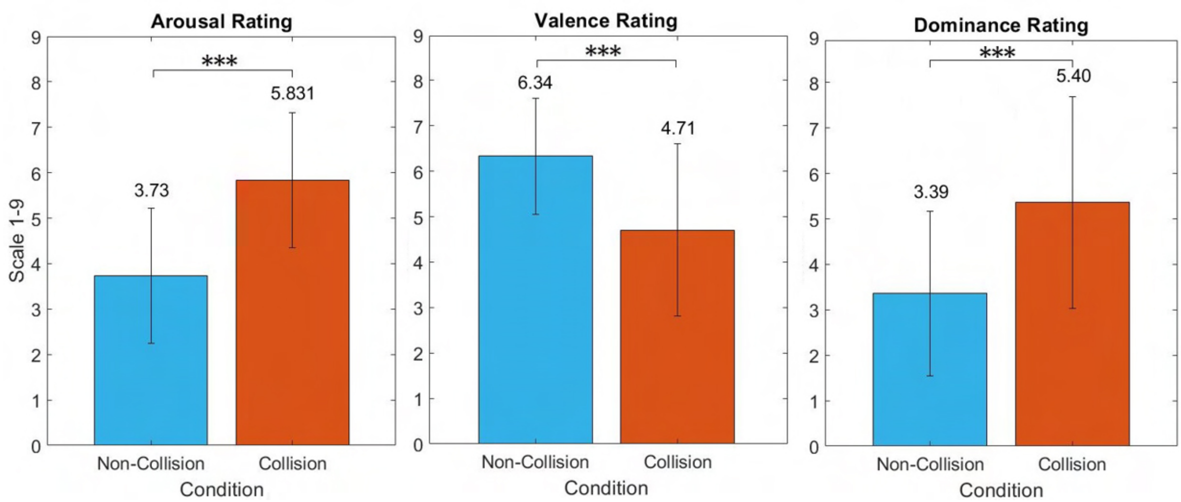

Figure 5.4: The arousal, valence, and dominance SAM rating from the participants.

5.4.5 Participants

The experiment was conducted in person under strict hygiene guidelines (COVID-19 restrictions) and with the approval of the UTS research ethics committee. All participants provided written informed consent for the experiment and video recorded interviews. Each participant received monetary compensation for 1 hour and 30 minutes (experiment duration). The participants were screened within the target age range (18-35 years) and must possess the language capacity to complete the interview. We recruited 21 participants for this experiment. One participant was excluded due to low signal quality from the EEG data. Hence, the data analysis involved twenty participants (N=20, 7 females and 13 males), the mean age was 26, and the population variance was 4.13.

Based on pre-experiment questions, we found most participants (18, 2 selected "monthly") reported to either have "never" seen a drone before or rarely stating a "yearly" frequency. This suggests that our participants were not familiar with encountering drones, especially in dangerous situations.

5.5 Results

5.5.1 SAM

Figure 5.4 presents the average values of the participant's SAM responses. We found a significant difference between collision and non-collision in the arousal level (Non-Collision M=3.73 SD=1.49, Collision M=5.83 SD=1.48, and W=0, Z=-3.92, $p<0.001$, and

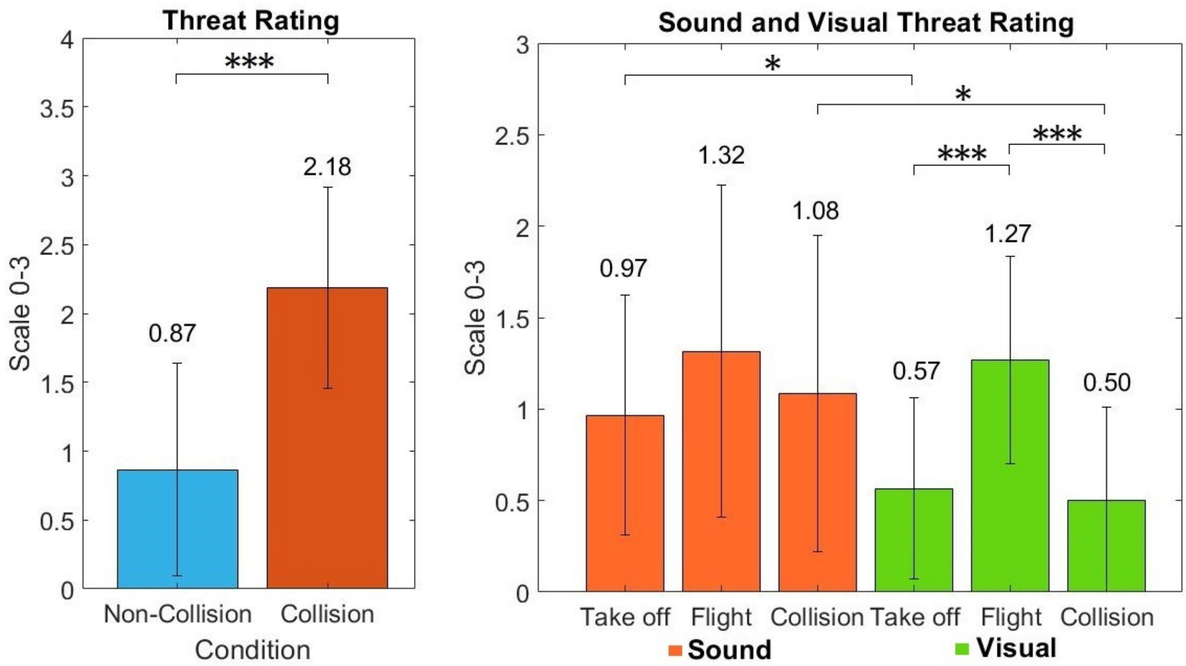

Figure 5.5: The threat scores for collision vs non-collision and Sound vs Visual.

r=0.88), valence (Non-Collision M=6.34 SD=1.28, Collision M=4.71 SD=1.90, W=170, Z=3.68, $p<0.001$, and r=0.82), and dominance (Non-Collision M=3.39 SD=1.83, Collision M=3.39 SD=1.83, W=0, Z=-3.82, $p<0.001$, and r=0.85). The increase in arousal (heightened activity), decrease in valence (negative emotions), and increase in dominance (less control), suggests that the participants are experience an increase in stress.

5.5.2 Threat Scores

The threat ratings for collision vs non-collision and the sound vs visual threats are illustrated in Figure 5.5. We found a significant increase (F(1,38)=30.60, $p<0.001$, and partial η^2=0.45) in perceived threat for the collision (M=2.18 SD=0.73) condition compared to the non-collision condition (M=0.87 SD= 0.78).

When comparing the sound and visuals, we found a significant difference between the sound (Take off M=0.97 SD=0.66, and Collision M=1.08 SD=0.86) and visual (Take off M=0.57 SD=0.50, and Collision M=0.50 SD=0.51) during thee Take off (F(1,38)=4.72, p=0.036, and partial η^2=0.11) and Collision (F(1,38)=6.74, p=0.0134, and partial η^2=0.15). We did not find a significant difference between the flight stage of sound (M=1.32 SD=0.91) and visual (M=1.27 SD=0.57) threats. We also found a significant increase in visual threat during flight when compared to the take off (F(1,38)=17.21, $p<0.001$, and

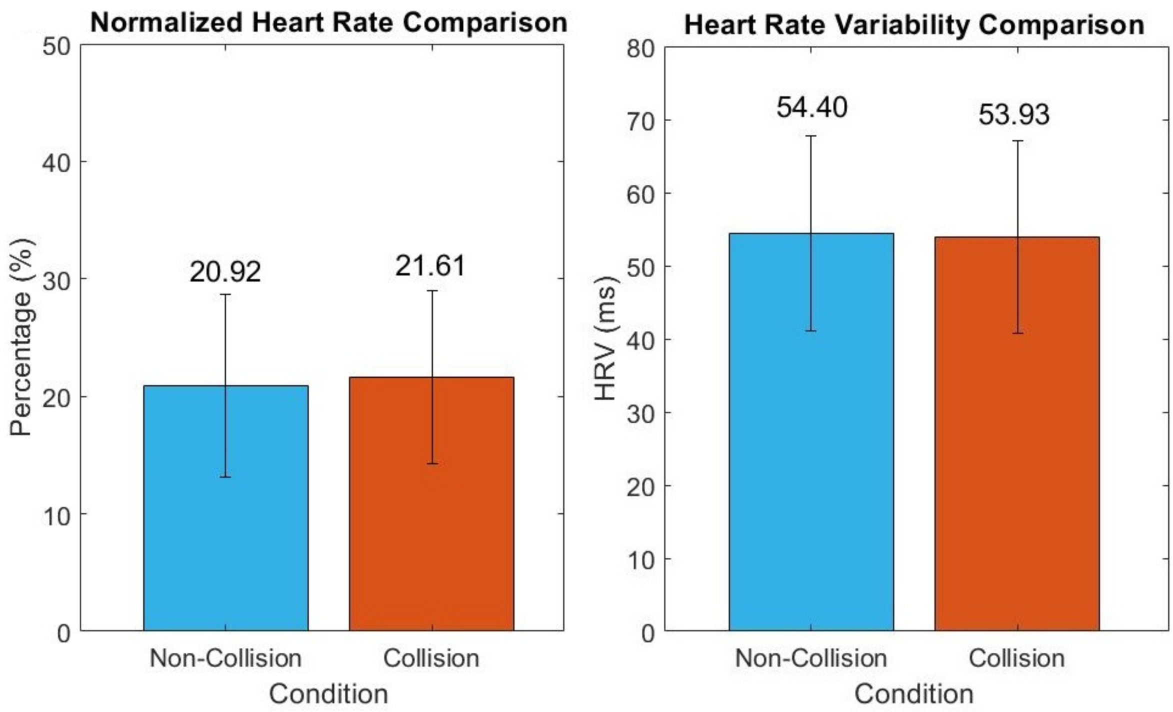

Figure 5.6: The average participant HR and HRV results for collision vs non-collision.

partial η^2=0.31) and collision (F(1,38)=20.06, p<0.001, and partial η^2=0.35) stages.

5.5.3 HR/HRV

We did not find any significance difference between the HR (non-collision M=20.92 % SD=7.78, and collision M=21.61 % SD=7.37) and HRV (non-collision M=54.40 ms SD =13.31 and collision M=53.93 ms SD=13.22). Figure 5.6 shows the average plots for both HR and HRV. Based on the results, the HR and HRV does not change between the collision and non-collision condition.

5.5.4 EDA

Similar to the HR/HRV findings, we did not find any significance within the EDA data (non-collision M=33.13 % SD=25.59, and collision M=33.25 % SD=25.11). This would suggest no increase in EDA between collision and non-collision (see Figure 5.7. This indicates that the collision condition does not elicit an increase in EDA when compared to the non-collision condition.

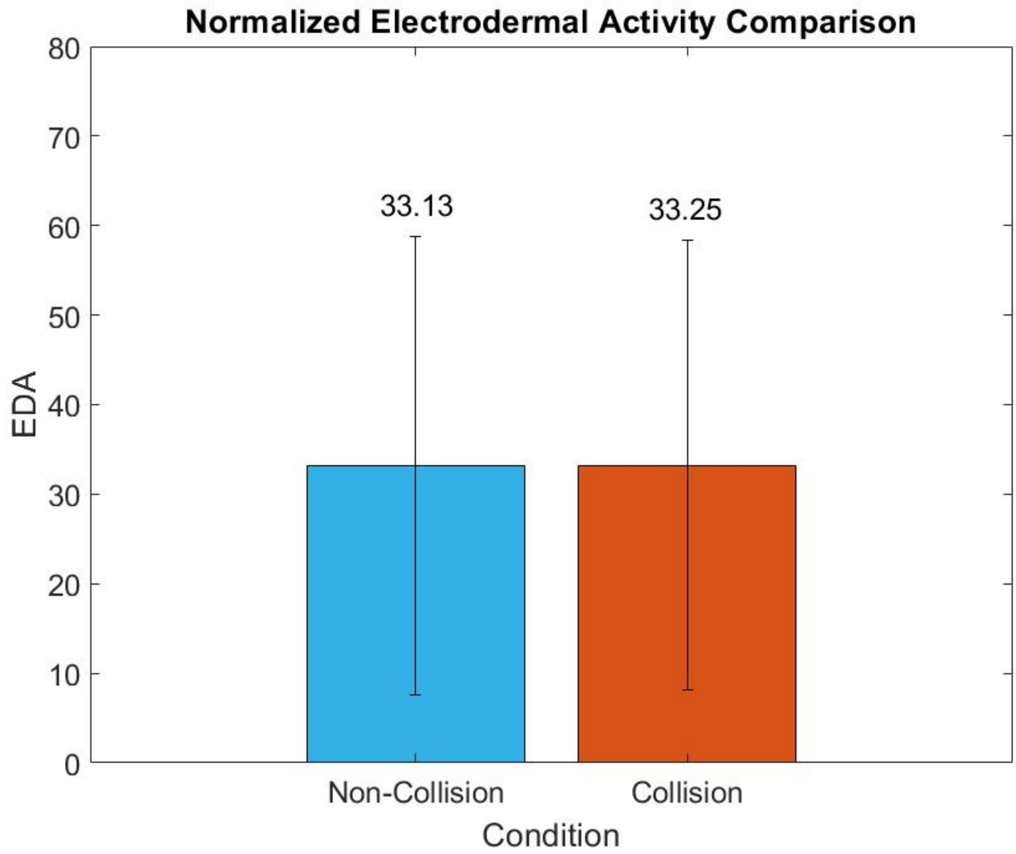

Figure 5.7: The average participant EDA results for collision vs non-collision.

5.6 Discussion

5.6.1 Elicitation of the Stress Response

Overall, the SAM Figure 5.4 and the user threat rating Figure 5.5 indicates that the participants feel threatened by the drone during the collision. The SAM behaviour of increased arousal, negative valence, and increased dominance ratings are consistent with the expected stress response. An increase in perceived threat is evidence supporting a stress response. On the other hand, the HR, HRV, and EDA data found no significant difference between non-collision and collision. From the questionnaire results and post-experiment interview, we believe the participant did experience a stress response (heightened alertness/arousal). However, several factors may contribute to the lack of observable effect in the physiological measures.

One obvious factor could be the strength of the stressor stimuli (drone collision). Similar to our results in the VR height stress experiment (Chapter 3, the lack of change in the *PG* and *PP* conditions), the lack of physiological significance could be an indicator that the drone collision was not threatening enough. The participants may consciously recognise their increased alertness and self-report their stress state, but the response's strength was not overtly visible in physiological data. Without another follow up experiment, we will not be able to fully gauge if a higher magnitude stressor could produce a more overt stress response. Possible ways to increase the magnitude of stressors could be to increase the drone's size, reduce the drone's protective equipment (if safety can be maintained), or introduce some light safe haptic element to the experiment (if safe and the ethics committee approves it).

Another factor could be the habituation and participant is predicting the collision event. The binary nature (two conditions) and repetition of the stress stimuli may have produced a masking effect for the physiological measures. The counter-balancing of trials blocks and rest breaks in between 10 trials are one form of mitigation for the habituation effect. We further investigated this in the post-experiment interview by discussing how their perception of drone collisions changed over time. Fifteen participants (N=20) stated that they felt more comfortable with the drone collision and less threatened after the 30 trials. Conversely, five participants cited that the drone's unpredictability stopped them from feeling comfortable with it throughout the experiment. This would suggest that the predictability of drones contributes to whether or not the participant becomes comfortable with drone collision.

5.6.2 Dynamic vs Prolonged Stimuli Stressors

With the previously stated factors in mind, we believe another crucial reason for the lack of physiological change is due to the nature of DSS compared to PSS. The benefit of PSS is the attenuated stress response from the continuous nature of stimuli. To give perspective, a trial of stress experiment involved 7 seconds of walking and 150 seconds of oddball trials. In a single trial was participants would remain exposed to virtual height through the whole period. Another example is the TSST task [102] which involves 5-minute blocks of interviews to induce physiological stress. In contrast, the drone collision difference only occurs within the 1m range (when the participant recognise the drone is not stopping); this leaves a small window (1 meter to collision time was 472ms±87) for the collision reaction. Participants often react to the collision after the fact. This poses a technical limitation in the physiological measures. The HR metric generally operates at a seconds rate with 80 bpm (1.33 beat per second) being an average healthy adult range (at 60 bpm - 100 bpm range) [11]. A DSS task with a small-time window may cause a minor variation in rate rhythm that HR readings can easily miss. HRV suffers a similar issue with the requirement of a 5-minute rolling average; any HRV variation would be averaged out from the shortness of the response period. The wristband may also overlook this physiological response as the EDA is sampled at 4Hz (250 ms reading) compared to the shorter response of the drone collision Table 2.2. Future reproductions of this type of experiment should consider using a measuring device or higher fidelity to capture the short-term physiological response accurately.

Another consideration is that the non-collision event may have produced a certain amount of stress from flying toward the participant. This is similar to the study of Healey et al. [81] which suggests that the stress level in the driving task events may be less pronounced due to driving itself being already stressful. Similarly, the participant has felt threatened by the drone itself and the uncertainty of the collision. This factor is a potential drawback for DSS based paradigms as it may be harder to create apparent, distinct stimuli neutral/passive stimuli compared to the ease of PSS.

In summary, the drone-based DSS experiment can produce a threat response and cause a conscious stress response (from the self-reported questionnaires). However, the stimuli' sudden and short duration nature is problematic for less sensitive physiological measures such as HR, HRV, and EDA. EEG is a more responsive measure that can yield exciting findings, outlined later in Chapter 6.

5.6.3 Threats of a Drone

Identifying the main reasons for the threatening nature of drone collisions can provide valuable information to improve future iterations of this experimental design. Based on prior literature [2, 89, 203, 207], we found that the drone's noise, physical appearance, and privacy as the main concerns a person may hold when encountering a drone. While privacy is an interesting factor, we chose not to deeply explore it in this experiment as it is unlikely to contribute to drone collisions. Understanding the visual and auditory drones can potentially modulate these factors to alter the participant's perception of a drone.

Firstly, the threat scores (see Figure 5.5) provides a quantitative measure for visual and sound. Based on the results, the sound is rated higher than the visual appearance during the experiment's take off and collision stages. Interestingly, sound and visual threats are matched during the flight phase of a trial. It is not surprising that sound dominates during take-off since the drone is farthest away from the participant. Surprisingly, participants would rate the sounds higher than the visual threat during the collision. This may be because participants tended to flinch or move their heads during a collision or because the drone at its closest was not as threatening. In any case, the drone is most threatening during its flight toward the user, with the sound of the propellers as the over more dominant threat.

In the post-experiment interview, we asked the participant to hold a guard and unguarded drone (see Figure 5.2) and provide any notable observations. The participants mentioned the propellers (mostly when unguarded) as the most dangerous component of the drone. The common concern was the potential for personal injury. All the participants noted the propeller's visually threatening aspect. However, most participants noted that the small propeller was unexpectedly loud in volume.

Another observation was the size and weight distribution of the drone. Interestingly, four participants noted that the guarded drone seemed larger in-flight (drone exposure) than a stationary held position. When observing the drone's actual size, most participants noted the appearance felt safer than during the experiment.

Similar to previous studies, we found that the shape and level of protection of the drone guards can relate to the level of threat perceived by participants [2, 207]. These results stipulate that the propeller is the primary (agreed by all participants) threat perceived by participants due to the potential to inflict bodily harm.

In future drone designs, the propeller design should incorporate both shapes, level of protection, and noise level to create a drone that can be perceived as less threatening. Options such as low noise propellers may improve the reputation of drones being loud and dangerous.

5.7 Limitations and Future Works

5.7.1 Variance of Drone Behaviour

One major limitation of this experiment is the imprecision of the drone movement. The drone flight can be affected by airflow, altitude (air from the ground), flight path, and pilot's human error. These factors lead to variations in individual trials, such as specific non-collision trials finish closer to the net while others further out. Other variances included flight path deviation, different trial times (more prolonged than others), and times when the drones caught on the net. These variances may have contributed to the lack of difference between the collision and non-collision conditions and potential issues for analysing time series data metrics.

To address this, we marked out specific positions and targets in the experimental setup to maintain consistency of take-off and collision positions. We used foam padding on the ground to minimise updrafts from the ground. The drone operator also extensively practised the experiment beforehand to minimise the variance further. Nevertheless, these trial variances still occurred during data collection. Future iterations of this study should an automated flight approach using the motion capture to actively track and adjust the drone's flight path to reduce the variance between trials.

5.7.2 Demographics and Cultural Perception on Drones

The demographics and cultural factors were not considered during the experiment's conception, but some participants brought them up. An example of this concern was a study by L.E. et al [57], which explored the relationship between Chinese culture and drone interaction. A few participants mentioned encountering drones more commonly in the U.S. and Asian counties, which affected their sense of comfort and safety around drones. Another participant noted the implication of drones being used as weapons which would affect the people from cultures and countries within conflict zones. Future studies can further explore a broader demographics and cultural range to understand better how others perceive drones. We suspect that the drone will become more gradually accepted over time, and these perceptions will inevitably shift.

5.8 Key Points

From the presented details, we can summarise that participants perceive drone collisions as threatening and dangerous. From our qualitative and self-reported questionnaires (SAM and threat scale), it is clear that participants did experience some level of stress. This stress response may not have been detectable due to the transient nature of DSS based paradigms and the technical issues (sensitive and sampling requirements) physiological measurements. The issue with physiological measures is likely to persist with other DSS paradigm. Thus PSS paradigms are more beneficial for physiological measurement systems. DSS paradigm may require physiological measurement devices with much higher fidelity (high sample rate and electrode pads) to measure the stress response accurately.

From these findings, we have partially answered ***RQ1***. We believe the questionnaire responses provide a strong indication that the participants have experienced a change in physiological state. However, the limitations of the physiological measures hinder our ability to observe them accurately. In Chapter 6, we will highlight our EEG findings; EEG metrics are more responsive and thus may yield more valuable information for DSS based paradigms.

CHAPTER 6

EFFECTS OF DRONE RELATED STRESS ON BRAIN DYNAMICS

6.1 Overview

This chapter presents the EEG results of the drone study to achieve ***RQ2***. As previously established in Chapter 5, we found great difficulty in observing the stress response of our physiological metrics. This was likely due to the insufficient sensitivity of the measuring devices. EEG, on the other hand, is more sensitive to rapid/minute changes. Here, we present the methodology, results, and discussion for our EEG analysis. From the exploratory analysis of the EEG data, we found distinct dipole component clusters that correlate to the drone collisions. Within these clusters, we found that the drone collisions significantly causes central and frontal activation. The EEG results indicate detectable neural activity during collisions event when compared to the non-collision event.

6.2 Methodology

6.2.1 Exploratory EEG

The novelty of this experiment crucially changes the methodology of the analysis. In the instance of the height exposure experiment (Chapter 3 and Chapter 4), the methodology was hypothesis-driven with established preliminary analysis and results that we can aim to reproduce. In particular, the channel-based analysis in Chapter 4 would be complicated as we are not relying on a well-known paradigm with specific regions of interest (ROI). This is primarily due to the P300 response being extensively studied with established ROI, and ERP behaviours [153]. In contrast, the drone experiment does not possess established prior neurophysiological behaviours. Instead, we would need to rely on exploratory EEG analysis techniques to identify the various brain activation sources during the experimental conditions and cluster them on a group level. This can be done through ICA, which will identify any independent component (IC) and its dipole locations within the EEG signal [189]. Then, using the Matlab EEGLAB clustering tool, we can obverse any notable clusters on a group level.

Although there is direct prior work relating to EEG drone collisions, we can derive ideas from previous studies that use the exact mechanism. In this case, we interpret the works by Muraskin et al. [138, 139] which investigated the neural behaviour of baseball players when facing a ball pitch (ball thrown at them). Their participants were exposed to a computer monitor that played videos of optimal pitches (fastballs) or non-optimal pitches (curveballs). The participant would need to decide to swing or not swing. They rationalised that this situation was similar to a Go/NoGo experiment design. The study found that participants tended to exhibit alpha power changes during optimal pitches trials in the motor area and P300 like behaviour in the frontal area. Following this concept, we can treat the collision task as the Go (or swing) and the non-collision as NoGo (or not swing).

6.2.2 Preprocessing

The preprocessing for the EEG data for this experiment was less aggressive than the method outlined in Chapter 4. The participants mostly remained stationary throughout the experiment, meaning we will not need to contend with walking, heavy head movement, and other motion artefacts. Thus, we did not apply ASR as it may reconstruct low noise portions of the data due to the assumption of a high amount of noise. We also replaced AMICA with RUNICA, which is a faster and less computationally intensive ICA algorithm. The preprocessing pipeline steps were:

1. Bandpass filter 1-50Hz for Anti-Aliasing and remove high and low frequency noise,
2. Down-Sample data to 250Hz to reduce computational time,
3. Clean line noise,

4. Re-reference data to signal average,
5. Compute ICA weights and Spheres through RUNICA,
6. Dipole fit components,
7. Classify components (IC labeling), and
8. Remove any component that is less than 85% classification for brain component,

6.2.3 Data Analysis

After running the ICA, we ran the Matlab dipole fitting (DIPFIT) [48] and IC labelling toolbox, which provides an estimated source position and classification. We then sorted the data into epochs around the collision and non-collision conditions. We specifically epoched the data at the 1m (1 meter from the net) event with the time range of -800 ms to 2000 ms. We chose to epoch around this event because the drone would start to brake or continue forward at that 1m point. Before that point, the participant should not predict whether or not the drone will collide into the net (besides 50-50 guessing). Therefore, the critical difference between the collision and non-collision occurs at the 1m distance point. We also tested epoching at the 3m range, but we found that the statistically significant change still occurs around the 1m event (assuming the drone travels at around 2m/s or 1 meter every 472ms±87, see Figure 6.1).

We then generate an EEGLab study to assess the metrics on a group level. We clustered the EEG data using K-means clustering, producing 6 clusters and removing outlier ICs at 3 standard deviations from cluster centroid. The K-means clustering method can produce variations between runs, leading to varying results within the IC clusters and final results. To better optimise the results, we follow the methodology of Gramann et al. [74] and Do et al. [55] by implementing a repeated K-mean method. We ran 2000 repetitions of the K-mean clustering and assessed each iteration. Based on the repetition, we evaluated the most prominent clusters (most subjects and ICs) and clustered ICs based on the best results.

After discerning the relevant IC clustered, we examined the location of the clusters producing a clustered dipole map for both positions (Talairach space) and density (depth). These clusters indicate source activity in the various regions of the brain. Then, we assess the ERP response to observe if a P300 wave is visible (as suggested by previous studies [139]). The ERP significance was calculated through EEGLab's Study statistic permutation function, which uses a one way ANOVA test with the criteria of $\alpha < 0.05$. We also assessed the ERSP response for spectral changes between collision and non-collision. The ERSP response statistical significance was also calculated through the one way ANOVA test with FDR correction.

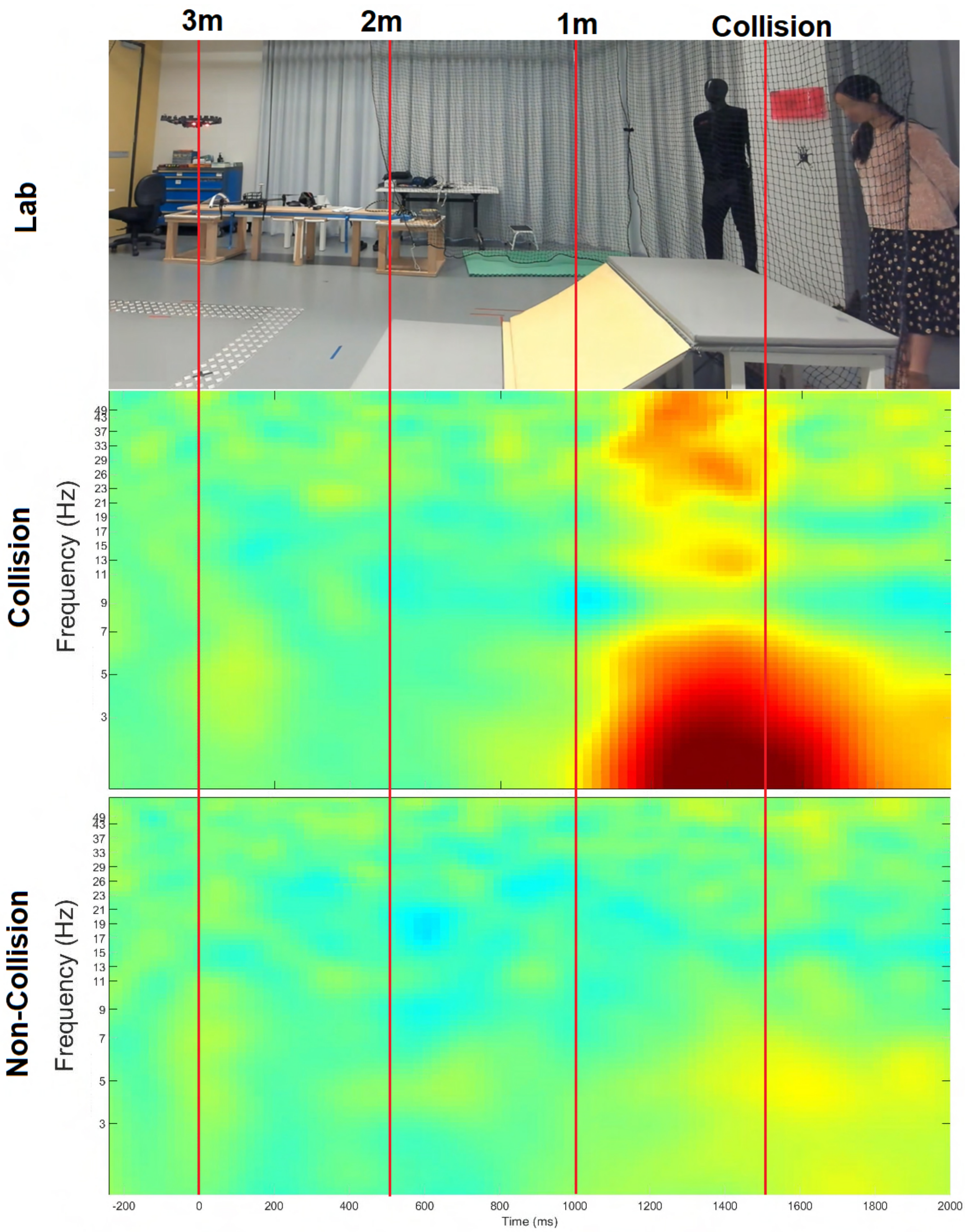

Figure 6.1: The ERSP response of a ClS cluster that is epoched and the 3 meter event. We overlaid it with the experimental setup with the estimated distance and time marked on the image and ERSP. At seen in the ERSP response, the main difference occurs at the 1m range.

6.2.4 Dividing Participants into Groups

The methodology in Chapter 4 hints at a potential method of classifying an individual's susceptibility to stress. The significant difference between the two groups suggests an individualised difference in the effects of stress on cognitive performance. This possibility is evident in Group 1 exhibiting a reduced performance (higher RT) when under stress, while Group 2 performed slightly better (lower RT). By evaluating the participants' RT performance in both stressed and non-stressed states, we theorise that it is possible to measure and predict their cognitive performance when under stress potentially. From our previous analysis in Chapter 4, we identified that the RT and the beta oscillation power are two potential indicators that can be used to gauge the participant's level of distraction or perceived difficulty of a task. Future stress elicitation studies may benefit in performing within-subject grouping by either screening the participants through performing an RT task under stressful conditions or dividing the participants in the post hoc analysis.

We applied the group splitting methodology of using RT as the metric for the median split. We used the participant's RT response from the button press during drone collisions. The participants were instructed to press the button when they felt the drone would collide with them. The median RT time for the collision was 535 ms (after the 1m mark) which would have been around the period the drone collided with the net. Two groups were formed from the RT median split. Group 1 (N=10) consisted of participants who responded to drone collisions after the median period. These participants likely did not anticipate the drone collision (or accurately predict the collision) and reacted during the collision event. In contrast, Group 2 (N=10) responded before the median time, suggesting that they expected the drone collision before it occurred. We performed the same EEG data analysis as Chapter 6 for each group and assessed the region activity, ERP, and ERSP.

6.3 Results

6.3.1 Dipole Clustering

Figure 6.2, depicts the two main dipole clusters (out of six) that we highlighted. These clusters were chosen based on the subjects that it included (ClS3, and ClS8 included 18/20 subjects) and the at least 2 IC per participants (ClS3- 45 ICs, and ClS8- 38 ICs). From the cluster's centroid location we can observed the regions and depth of brain

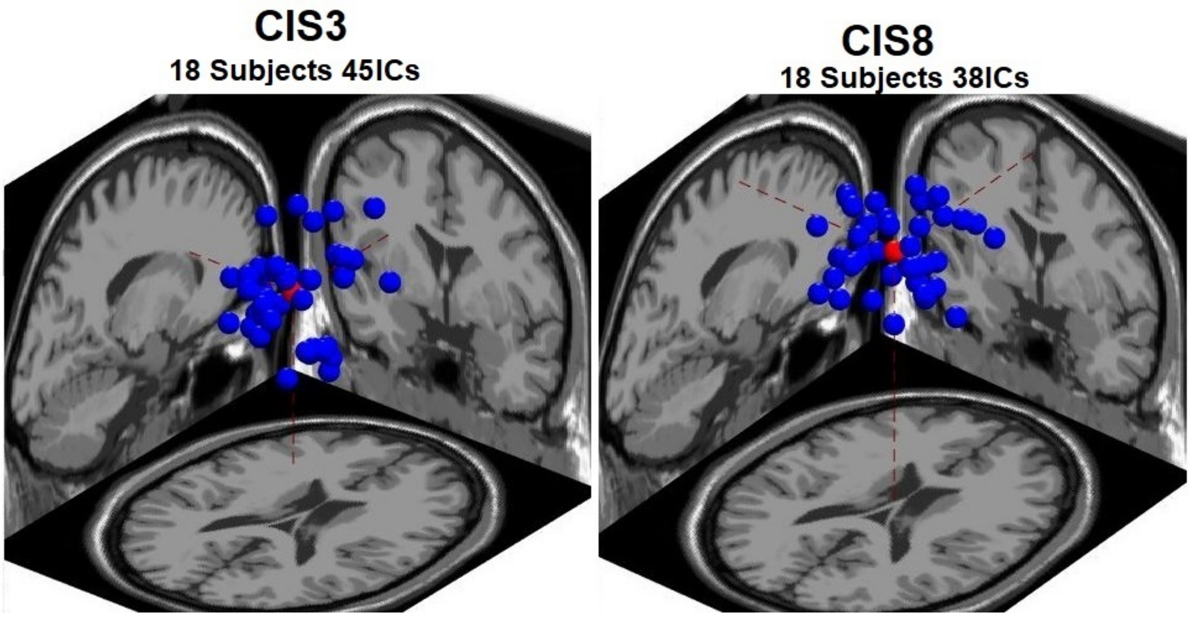

Figure 6.2: The two chosen K-means independent clusters found after 2000 iterations of repetition. The subjects and ICs are included in the subheading for each ClS.

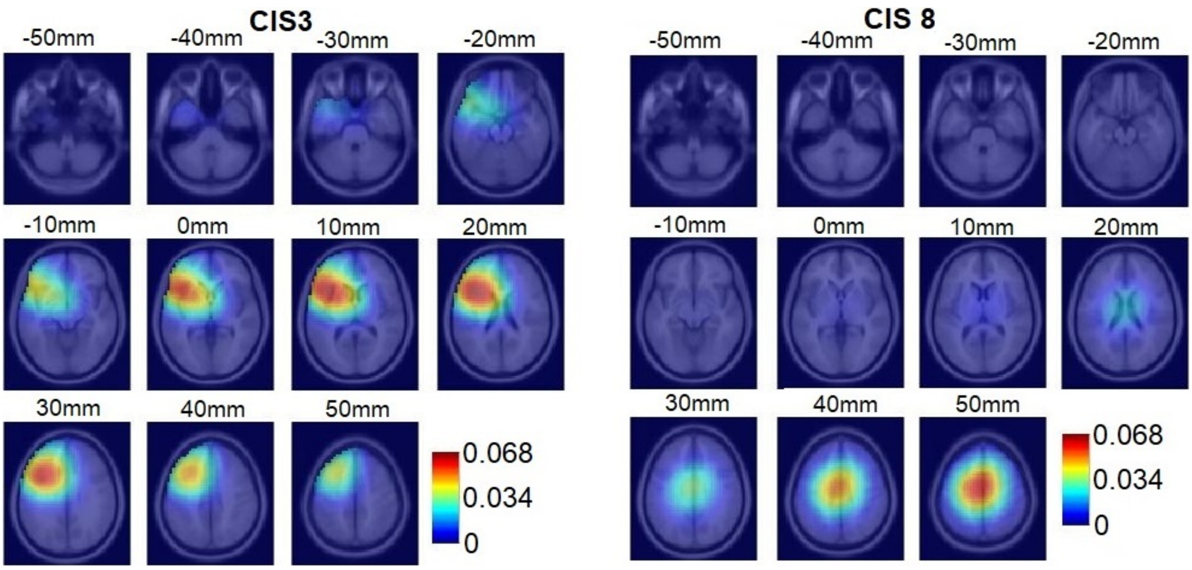

Figure 6.3: The dipole density of the ClS3 and ClS8 clusters.

activity (see Figure 6.3. The ClS3 cluster shows a left side frontal activity. In the ClS8 cluster, we see a central (motor and sensory cortex) activity.

Figure 6.4 and Figure 6.5 presents the IC dipole locations and regional density for the IC cluster found in Group 1 and Group 2. Both clusters contained a sufficient amount of unique subjects and ICs per subject (at least 2 IC per subject). Both the Group 1 and

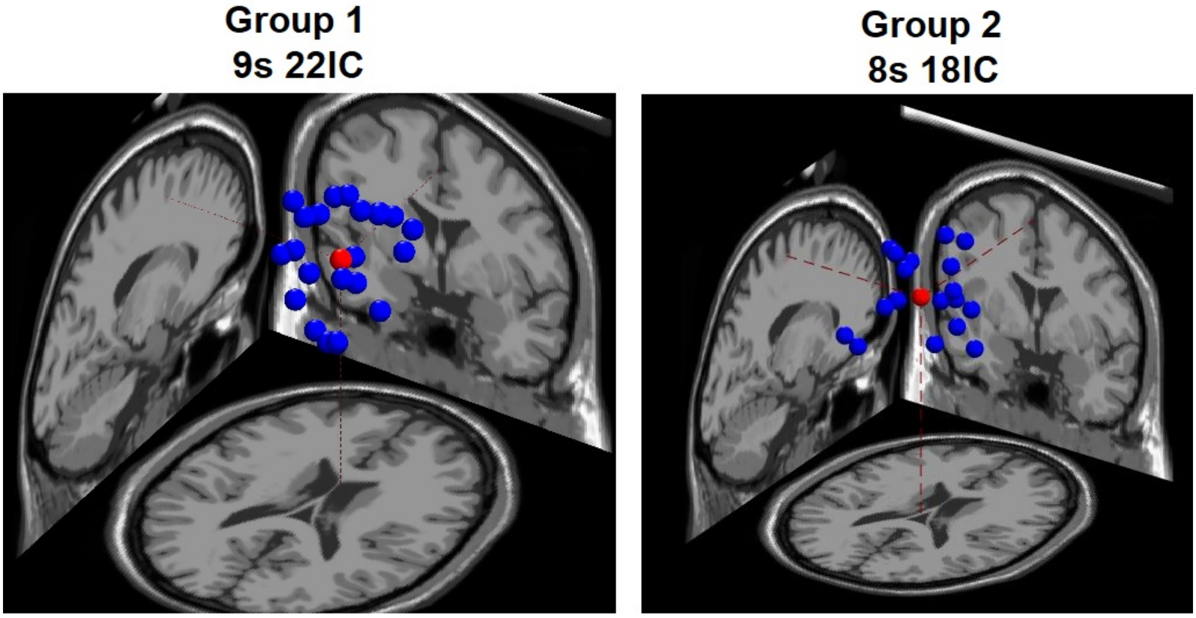

Figure 6.4: A comparison of Group 1 and Group 2 IC dipole clusters for the drone collision experiment

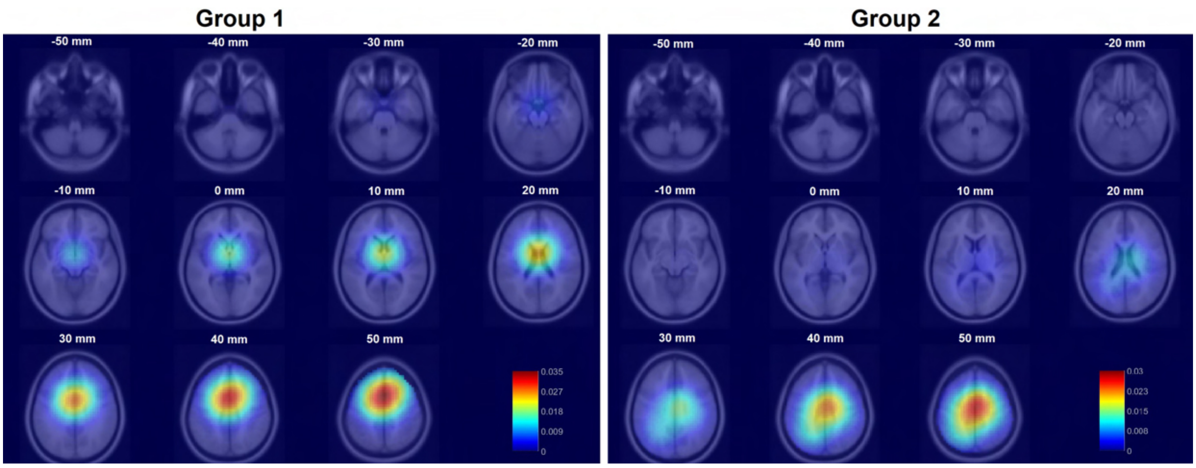

Figure 6.5: A comparison of Group 1 and Group 2 dipole density for the drone collision experiment

Group 2 exhibited a similar brain region of activity (frontal-central cortex).

6.3.2 Event-Related Potential

Figure 6.6 illustrates the ERP response for ClS3 and Cls8. The ERPs were baselined from -800 to 0 period of the epoch. This would assume the participant's neural state

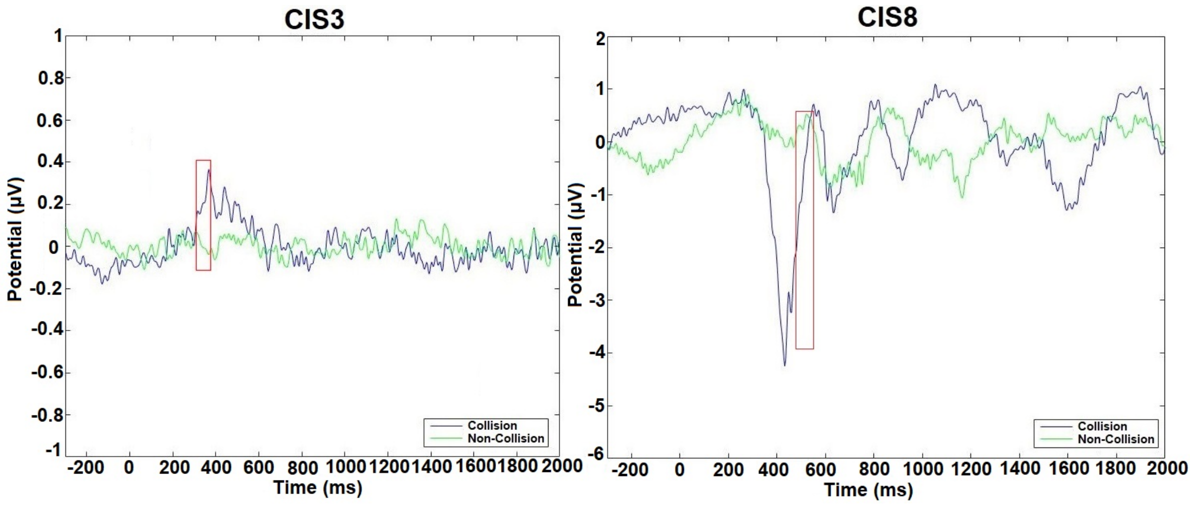

Figure 6.6: The ERP responses of the ClS3 and ClS8 clusters epoched around the 1m event for collision (blue) vs non-collision (green). The significance is indicated in the red box with $p<0.05$

was consistent between the collision and non-collision period prior to the 1m event. The ClS3 cluster ERP response shows a significant difference between the collision and non-collision condition around the 300-400ms period. There is a distinct signal peak during that period that is similar to a traditional P300 peak. The ClS8 clusters shows a distinct central region negativity around the 500-600ms period.

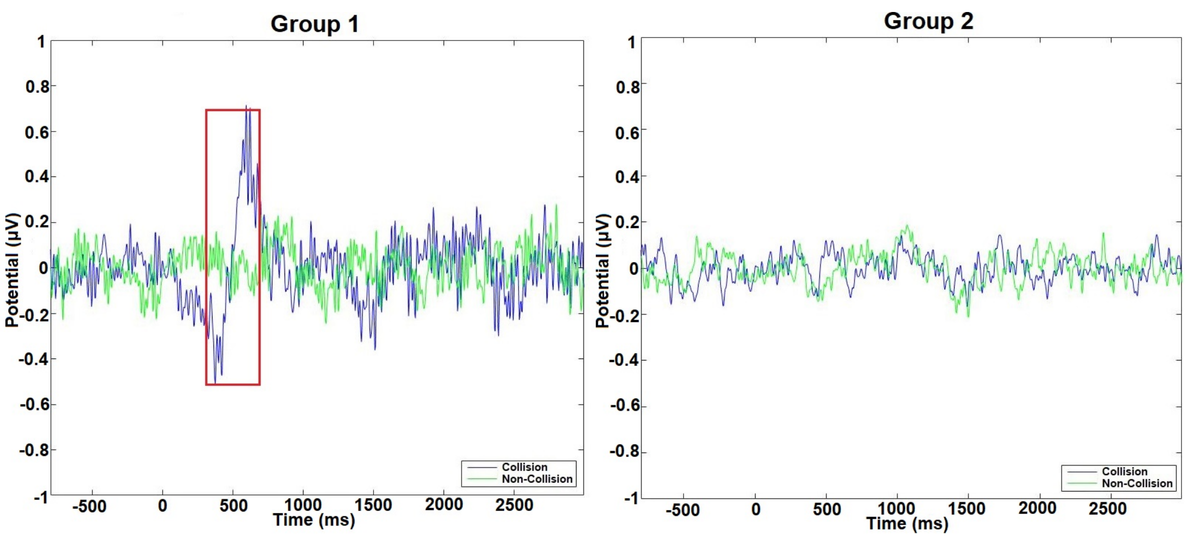

Figure 6.7: A comparison of Group 1 and Group 2 ERP response for the drone collision experiment. The significance is indicated in the red box with $p<0.05$.

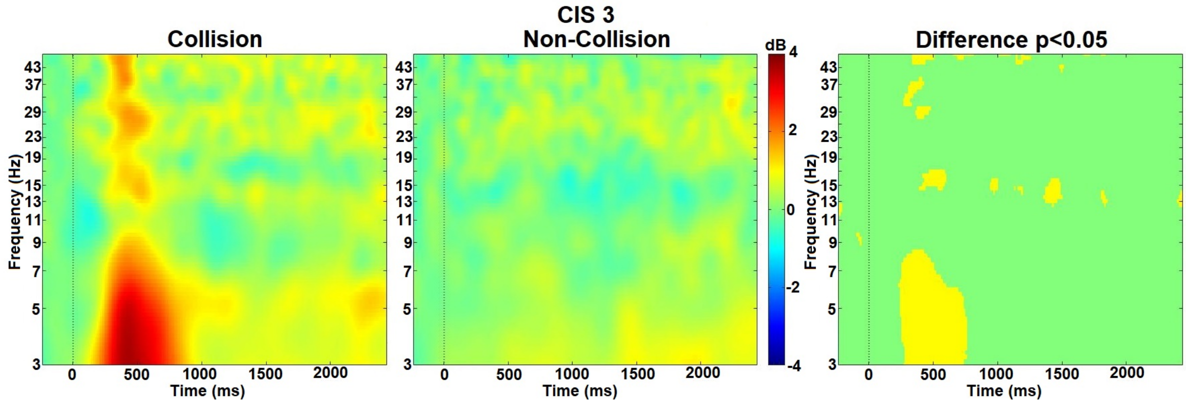

Figure 6.8: The ERSP response of the ClS3 cluster epoched around the 1m event. The significance is set at $p<0.05$, calculated through Permutation testing FDR correct

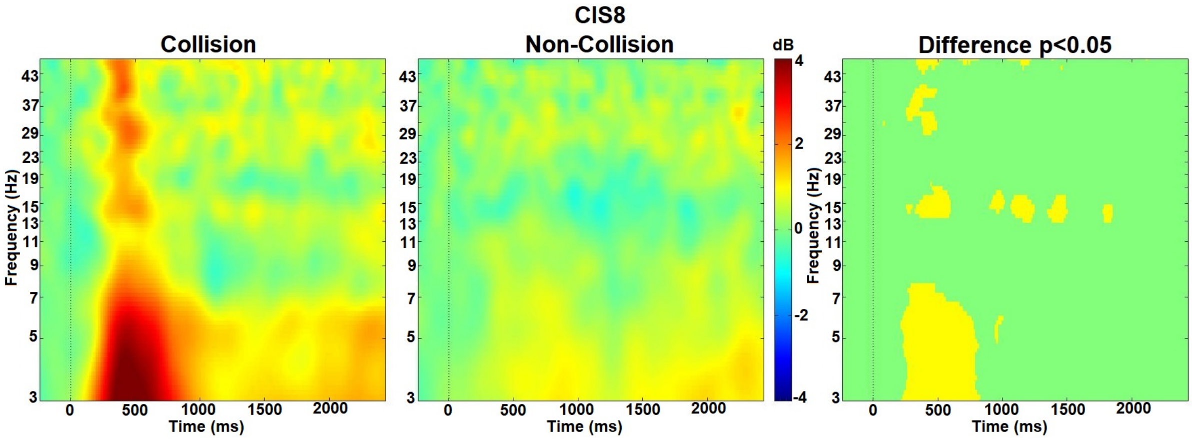

Figure 6.9: The ERSP response of the ClS6 cluster epoched around the 1m event. The significance is set at $p<0.05$, calculated through Permutation testing FDR correct

Figure 6.7 illustrates the ERP response of both groups. Interestingly, we found that Group 1 had both a significant negative and positive peak during drone collisions. On the other hand, we found no significant peaks or difference between collision and non-collision for Group 2.

6.3.3 Event-Related Spectral Perturbation

The ERSP response for the ClS3 and ClS8 clusters. The ClS3 (Figure 6.8) ERSP response shows that during the collision condition, there is activity around 300ms that attenuates at 1000ms. In comparison the non-collision condition does not exhibit activity. Comparing the statistical significance between the collision and non-collision conditions, a distinct delta (1-4Hz), theta (5-7Hz), low-alpha (8-9Hz), and low Beta (12-15Hz) power increase is observed around 300-700ms. This same behaviour can also be observed in ClS8 (Figure 6.9) which may suggest some connective behaviour linking the two regions of the brain (frontal - central).

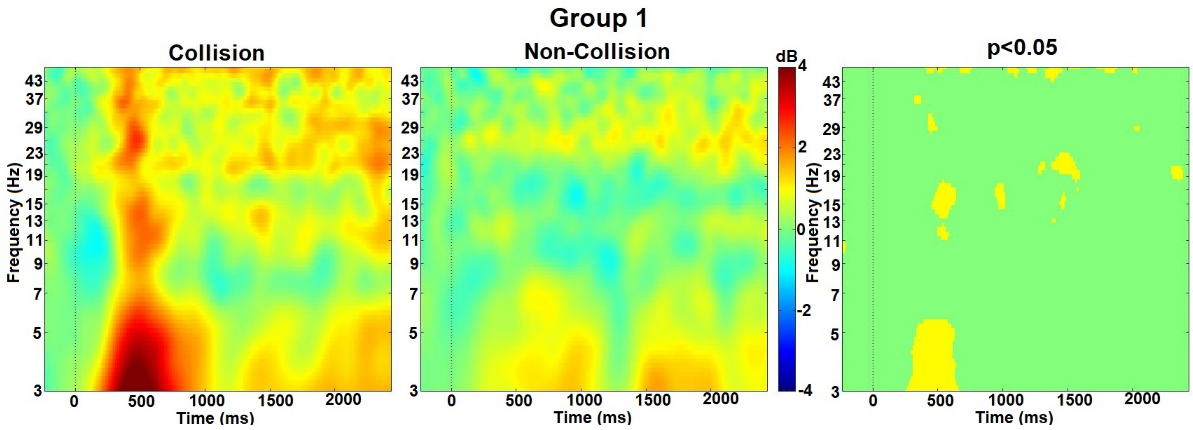

Figure 6.10: The Group 1 cluster ERSP response for the drone collision experiment. The significance is set at p<0.05, calculated through Permutation testing FDR correct.

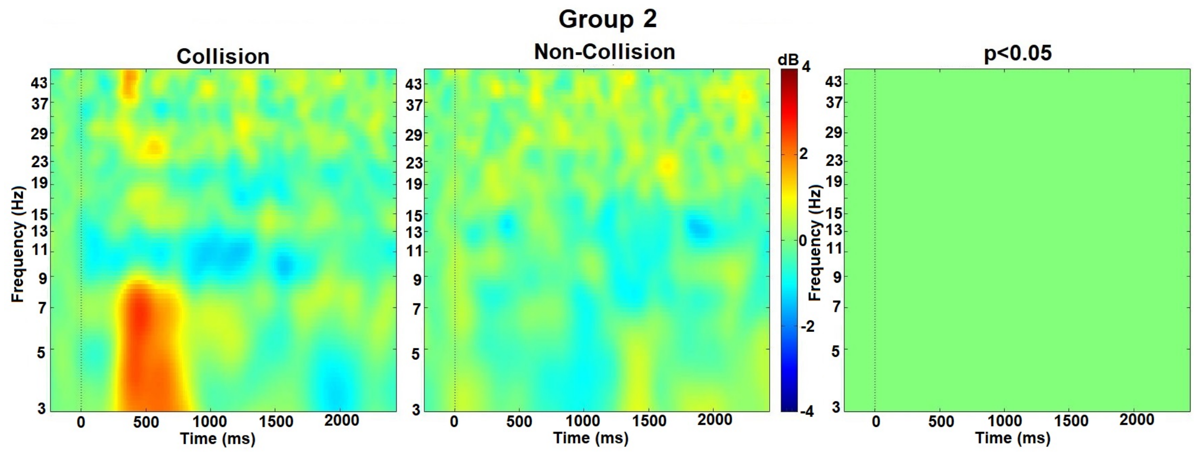

Figure 6.11: The Group 2 cluster ERSP response for the drone collision experiment. The significance is set at p<0.05, calculated through Permutation testing FDR correct.

Figure 6.10 and Figure 6.11 shows the ERSP responses for both Group 1 and Group 2. Both groups exhibited a theta synchronisation during the drone collision event. However, only Group 1 had a significant difference in theta power when comparing collision and non-collision events.

6.4 Discussion

6.4.1 Brain Dynamics during Drone Collision

Based on the exploratory analysis, the drone collision produces an apparent observable effect within the EEG signal. The ICA component clustering has found two clusters with many subjects and ICs within the cluster. Our ERP responses and activation regions are consistent with the findings of Muraskin et al. [139]. In particular, the ClS3 ERP is consistent with expected P300 behaviour in a Go/NoGo task [63]. The signal peak at 300-400 ms and the significance of that Go (collision) response compared to the NoGo (non-collision) response.

Interestingly, we observe significant negativity in the ClS8 ERP response in the central region of the brain. One reason could be due to the mixture of visual and auditory stimuli from the drone. Traditional Go/NoGo (and most P300 related paradigms) only uses auditory or visual stimuli sources. In the case of our oddball task in Chapter 3, the stimuli were visual. A study by Bennington and Polich [15] investigated the difference between visual and auditory P300 stimuli. They observed that auditory stimuli tended to have a slightly higher P300 peak amplitude but lower negativity (decrease in potential) before the P300 peak than a visual P300. Based on our prior threat score results (Figure 5.5), it could be reasoned that participants were more aware or attentive to the sound of the drone and reactive to the increase in volume rather than purely visual. In this case, CIS8 being an auditory response would correlate with the visual P300 results (from Chapter 4) that exhibit no negativity. However, a crucial flaw with this explanation is the timing of the negativity and the lack of a distinct P300 peak after the negativity. Given that CIS3 has a significant peak at around 300-400ms, it is unlikely that the CIS8 peak would be at such a different latency.

Another potential explanation for the ClS8 negativity may be related to the cognitive conflict response. Michael X Cohen [36] defines cognitive conflict as an event where one observes a mistake or incongruency within their external environment. The author further details that cognitive conflict can be detected by activity in the medial frontal cortex and theta-band synchronisation. Other studies [33, 123, 183, 197] have documented a distinct negativity in the ERP response when observing incongruent stimuli. Incongruent events are when a stimulus does not match a prior defined convention or intuitive expectation [209]. Tasks such as the Stroop task (mismatch words of colours and filled colours, see Figure 7.5) [36], Go/NoGo (curve vs fastball [139]), and Eriksen flanker task (detection of incongruent letters and arrows, [104]). The Cl8 ERSP response

(Figure 6.9) and the centroid source location (Figure 6.3) is consistent with the findings of cognitive conflict in both the theta synchronisation and region of activity [36]. Other studies such as Luu et al. [123], Trujillo et al. [197], and Cavanagh et al. [31] have all found a strong correlation between negativity and frontal-central theta activity.

Ultimately, the evidence for ClS8 being a cognitive conflict would be contingent on whether the drone collision is an incongruency compared to the non-collision events. Singh et al. [183] found a similar cognitive conflict response when a hand-controlled robotic arm produced unanticipated resistance during operation. This would suggest, at the very least, that robotic controls can produce a cognitive conflict event. Similarly, the drone collision event could be considered unanticipated. Naturally, one would assume a drone would not attempt to collide with them. So when the drone stops during the flight, it could be considered a congruent behaviour as the participant would remain relatively passive. A counterpoint to this could be that the participants were primed for drone collisions as the experiment was advertised as a drone collision experiment. The other argument could be that the participants still felt anxiety during non-collision trials due to the uncertainty of potential collision. In either case, the ERP (Figure 6.6) and ERSP (Figure 6.8 and Figure 6.9) is seemingly consistent with the cognitive conflict response. The factors of the non-collision (priming and anxiety of uncertainty) may have contributed to the lack of significance in some of the portions of the ERP response in ClS8.

We summarise that the drone collision event can produce a distinct cognitive conflict response observed in the ERP negativity and ERSP theta synchronisation in the front-central region. This would answer ***RQ2*** as the drone collision produces a detectable and discernible response that we can train into future classifiers for behaviour or drone collisions.

6.4.2 Dynamic and Prolonged Stimuli Stressors on Cognitive Behaviour

6.4.2.1 Cognitive Conflict Behaviour

As established in Chapter 5, DSS is more difficult to detect through physiological measures due to the short transient period of time that the participant produces a stress response. When comparing the ERP and ERSP response of the DSS experiment and the PSS results in Chapter 4, we observe two different kinds of EEG responses to stressor stimuli. The limitations of different experimental designs mean that we cannot directly compare the ERP response since the oddball task would produce a more consistent ERP P300 response. In the drone experiment, the independent variable is the drone collision with no other BCI-related paradigms. Thus we do not have a consistent P300 ERP for stress and non-stress compared with the PSS results. Nevertheless, we are able to draw some behaviour differences between the ESRP response drone collisions when compared with the *PH* condition of the PSS experiment.

In the *PH* conditions of our PSS experiment, we found an increase in beta and gamma power for the group whose performance was affected by stress. This suggested that the PSS potentially distracted the participants, causing a need for increased focus or attention resources (increase in frontal beta power) to be allocated during the P300 task. In contrast, DSS ERSP response shows mainly a theta and low beta synchronisation. The theta synchronisation is likely linked to the cognitive conflict response of the drone collision. The low beta synchronisation in DSS is similar to the PSS results. However, the ERSP response is missing the other markers (alpha synchronisation during the P300 period, see Figure 2.3) of a typical cognitive P300 behaviour.

6.4.2.2 Involuntary Disengagement

One rationale explanation we have for the ERSP behaviour of our DSS is that the drone collision event is a dominant threat that disengages other cognitive behaviours. In contrast, PSS height exposure is threatening, but it is not threatening enough for the participant to be disengaged in their oddball task. We derived this explanation from a study by Sherwin et al. [177], who investigated the cognitive process of a baseball swing decision process (similar to Muraskin et al. [139]). The authors reasoned that during the discernment of baseball pitch trajectory, the batter would have a period of cognitive decision making. Then, at a certain distance or time, the brain will disengage the decision making process in order to execute an action, whether it be a committed (experience leads to faster decision making [138]) or hurried (potentially incorrect) decision. The finding is unsurprising and corresponds to our intuitive understanding of human behaviour. The human brain can voluntarily and involuntarily force an action. In the case of the ball batting, if the ball is too close to the batter, then the person may have an involuntary rushed decision that is generated through the physiological stress response [144].

Similarly, for drone collisions, it is possible that the brain rapidly generates a physiological stress response to force an involuntary flinch or flight response for the drone collision. This response may dissipate shortly after due to the sense of safety from the net. From this perspective, the theta and low beta ERSP response and region (frontal and central) of the brain is consistent with the previous findings of the cognitive disengagement and the engagement of a reflex action (motor execution in motor cortex) [177].

Hypothetically, if a participant was performing an oddball task during the drone experiment, we suspect the drone collision would cause a complete disengagement to that task (potentially missing trials) due to an involuntary action such as flinching. This would differ from the distraction effect of a PSS which may hinder their attention and performance, but the participant would still maintain their ability for voluntary actions (such as walking) and produce distinct cognitive responses (oddball task). The only case that we observed PSS causing a disengagement (freeze response) is when they were first exposed to the virtual height. Over time participants were able to regain voluntary control of motor and cognitive functions. In contrast, our DSS stimuli would likely consistently cause an involuntary response causing the response seen in the ERP and ERSP, resulting in a cognitive disengagement.

6.4.2.3 Difference between Dynamic and Prolonged Stressors

From this idea, the crucial difference between PSS and DSS experiments is the neurological response and maintaining cognitive thought. PSS paradigm allows participants to adapt and acclimatise to the stress over time and regain their motor and cognitive abilities. This allows for consistent, measurable cognitive responses to assess various effects during the stress state as evident in the multitude of prior research [53, 100, 147]. On the other hand, DSS paradigms rely on the inherent unexpectedness of scary or threatening stimuli. The speed of the stimuli may not give participants enough time to react consciously, and thus the brain produces an instinctive physiological result that results in disengagement of cognition and engagement of involuntary motor response. This finding is essential as people may experience stress from various PSS and DSS sources in a real-world environment. The majority of prior research focuses on PSS sources that can translate into situations such as social anxiety, workplace stresses, and active tasks (such as walking at heights or driving). However, dynamic and unanticipated situations can induce a DSS response, such as reacting to dangers or threatening projectiles (such as drones, sports equipment, or magpies) or perceiving an incoming accident (driving). In order to improve the robustness of BCI, both these types of stimuli should be explored further and the findings applied to future BCI design.

6.4.3 The Effect of Group Splitting

The group EEG results indicate an observable difference between Group 1 and Group 2. While, both groups show a similar region of activity (see Figure 6.5, frontal-central), they differ greatly in the ERP (Figure 6.7) and ERSP (Figure 6.10 and Figure 6.11) response. Group 1's behaviour is consistent with our previous result in Chapter 6. The ERP response shows significant negativity at around 400ms and, surprisingly, a positive peak around 600ms. The ERSP response is similar to the cognitive conflict response with a significant theta synchronisation during drone collisions. On the other hand, Group 2 did not find any significant difference in the ERP and ERSP response between conditions. From the RT difference between the groups, we can infer that Group 1 was likely more surprised by the drone collision when compared to Group 2. Therefore, we argue that Group 1 may have a higher degree of cognitive conflict as the collision event was less anticipated than Group 2, which anticipated the collision by responding earlier. Group 2 did exhibit a theta synchronisation similar to an incongruent stimulus. However, it was not significantly different from the non-collision event. This lack of significance could be due to the participants in Group 2 being prepared for the drone collision or not perceiving the collision as an incongruent condition because it was anticipated.

These findings demonstrate the value of separating the participants into groups based on their RT performance. The separation of the participants yielded further insights into the different responses of drone collision and how the participant's anticipation can affect the response.

6.5 Limitations

6.5.1 Head and Body Reactionary Movement

One apparent limitation is the involuntary motor response during the drone collisions. Participants would instinctively flinch or attempt to avoid the drone. It would be unreasonable and impractical to inhibit this behaviour because the response is often involuntary. Asking participants to avoid flinching may introduce uncontrolled factors such as avoiding focusing on the drone or incorrectly performing the experiment to avoid flinching. We could not avoid the involuntary movement, so we excluded or filtered it through our data analysis. The first stage of filtering is during the removal of non-brain components. If the ICA was effective in localising various signal sources, then the muscle and head movement artefacts should have been separated into an IC and removed as part of the brain component criteria (remove components <85% brain component classification). The second stage is component clustering. If the movement components pass the classification removal, it would likely be clustered into outliers (if location spread erratically) or centralised (if consistent noise such as eye blink). After our 2000 repetition K-Mean clustering, we assessed the other non-chosen components to confirm that those clusters are likely noise sources that survived the first stage. Reintroducing ASR may remove some head movement artefact but could adjust stationary portions of the data. Another option is to use a more robust ICA algorithm such as AMICA (we found it took six times longer than RUNICA and more computationally taxing). Overall, our clusters were consistent with prior literature, and the response time would have placed the activity before the drone collision.

6.5.2 Impact of the Drone and Netting

A counterargument for the effectiveness of drone collision is that drone netting may introduce an unwanted sense of safety. We received similar feedback from individual participants during the post-experiment interview. Some participants reasoned during the interview that this was scientific research. Therefore we would not knowingly (maybe accidentally) put them in real danger. Ultimately, drone netting and experiment safety is a necessity in this day and age. We explored multiple participant standing positions (both further and closer to the net) and various drone speeds. Based on repeated testing, we chose the closest position that a participant could safely be on. The drone was flown towards the should and head, which increases the perceived threat and danger. We also considered alternate designs such as using VR or a safer (softer) tactile drone to remove the need for a net. However, both options seem less realistic and likely remove any sense of threat compared to the minor impact of the net.

Another perspective could be that the non-collision condition is already stressful. This could be reasoned as the participant could be anticipating a drone collision, and thus they might be in an alert state in a non-collision event. This reasoning could explain the slight activity in the ERSP response (for non-collision) and the lack of significance in some portions of the ERP. However, the collision event produces an overall more pronounced effect than the non-collision. This could be further explored by sampling the epochs into more stages, such as 3m to 2m, then 2m to 1m and 1m to collision, to better explore the sequential effect of the drone movement. Based on the 3m epoch result (Figure 6.1), we suspect the sequential result will not yield new results. A future experiment may better explore this effect by stopping the drone at a more comprehensive array of positions.

Instead of only collision vs non-collision, we could explore collision vs 1m stopping vs 2m stopping, which may provide further insight into whether the non-collision event affects the participants.

Due to the repetitive nature of the conditions, habituation is another limiting factor that should be considered with the drone netting. Over time, participants would likely be more comfortable with the drone netting become less stressed due to the decrease of threat. Future studies could involve more dynamic factors such as varying the drone's speed (high speed vs low speed), moving the participant closer to the net over time, or changing the shape of the drone (large drone vs smaller drone).

6.6 Key Points

Based on results in this chapter, we can address ***RQ1*** and ***RQ2***. The previous chapter (Chapter 5) partially achieved ***RQ1*** through the questionnaire results. From the EEG results and comparison with previous literature, we speculate that the stress response is more sudden and dissipates faster due to the brain's involuntary motor response (flinching) rather than a purely conscious recognition of the threat. This is unique to DSS type paradigms which generally require unexpected stimuli that pose a threat to the participant. We observed this in the dominance of the ERP peaks and ERSP theta and low beta power increase. Both the onset (Figure 6.1) at 3m and the 1m ERSP response (Figure 6.8, and Figure 6.9) presents indicates activity in the frontal-central region when the drone pass the 1m mark. We believe this response is the unexpected physiological acute stress response that causes a behaviour similar to cognitive conflict event.

RQ2 is achieved through the examination of the regional activity, ERP, and ERSP responses. Our ERP results suggest that the participants exhibit a P300 response along with a negativity response. This would suggest that the participants were alerted or reacted (P300) to the collision as an error or incongruent behaviour. This is support by the region of activity (frontal - central) and the theta synchronisation in the ERSP responses. These findings allude to the potential of using the cognitive conflict response to detect drone collisions or when the participants suspect a drone collision will occur. Potential future analysis could include investigating the functional connectivity between the cluster (and more if found) to find insightful brain dynamic behaviours.

Similar to Chapter 4, we found that a median RT split of the participants can yield two separate groups with differing behaviours. Group 1 struggled to anticipate the drone collision and exhibited consistent negativity behaviour as shown in the ERP (Figure 6.7) and ERSP (Figure 6.10). In contrast, Group 2 predicted the drone collision before the collision occurred. Likewise, the ERP and ERSP (Figure 6.11) responses did not show any significant difference between collision and non-collision events. This method of dividing the participants into groups based on performance could be reliable for future stress-related studies to apply as a screening tool or during data analysis.

CHAPTER 7

EXPLORATORY AND FUTURE WORKS

7.1 Overview

This chapter will cover the exploratory and potential future works that integrates the findings of previous chapters. We hope these works may significantly contribute to the improvement of BCI systems for real-world use. The BCI application projects explored in the chapter were completed in collaboration with undergraduate students as part of their capstone projects. The students, James Michell, Mena Basaly, and Vinayak Sharma, developed these projects with the collaboration and supervision of Dr Hsiang Ting (Tim) Chen, and Mr Howe Yuan Zhu.

7.2 Real-world BCI Applications

7.2.1 Motivation

The integration of BCI to a robotic system was initially one of the end product goals of this project. Towards this goal, we explored and created BCI integrated control system for various robotic systems (ground robot and drones) and investigated stress management mechanisms (biofeedback) to accompany our knowledge of stress detection.

7.2.2 Motor Imagery (MI)

Motor Imagery (MI) is another standard BCI paradigm used in cognitive neuroscience. Similar to how P300-related paradigms rely on the P300 wave, MI utilised the reliability of the neural mechanisms of the motor cortex [46, 120]. The motor cortex region of the brain primary handles the executive motor controls of the body [137]. The theory for the execution of motor functions is the translation of cognitive thought of moving into a motor output. Through the repetitive motor action (an example is open and closing the left and right hand) and central (motor cortex) electrode EEG data, one can train a MI classifier to recognise the cognitive thought of a motor action without the voluntary action itself [135].

The advantage of MI is that accurate classification can be achieved with a minimal amount of electrodes. A study by Munzert et al. [137] achieves around a MI accuracy of 87.6 % with only 1 channel, and 2 reference leads. The advantage of MI is the absence of required stimuli (e.g. P300 requires target sound or visual). Once the MI classifier is trained, it can be operated on executive thought alone without visual or auditory stimuli. One common pitfall of many BCI applications is the need for peripheral hardware such as display screens, speakers, and control interfaces; one can argue against replacing this equipment with traditional control mechanisms instead, thus eliminating the need for BCI. Without the need for additional displays or speaks, MI is portable and can be faster in operation.

The main disadvantage of MI is the dependence for individualised classification and susceptibility to environmental factors [182]. MI often require multiple repetitions of a task to train a classify accurately. While the MI can be highly accurate [87, 149], it disregards the fact that the accuracy is not transferable between subjects [182]. In comparison, the P300 wave is much more consistent within a population. Thus a P300 signal classifier does not deviate considerably from person to person [110]. From the first disadvantage, the second disadvantage is the real-world performance compared to the controlled laboratory environment of training. Studies [182, 211] have often found a significant decrease in accuracy for laboratory trained classifiers used in a real-world setting. This is likely due to the environmental distractions or the need for motor functions such as walking or moving their hand (confusing the MI classifier with real motor action).

Ultimately, after weighing the advantages and disadvantages of MI, we chose to use MI as the control paradigm. Previous studies such as Bauer et al. [14] and Irimia et al. [94] have found success in applying MI to the robotic system. We hope to incorporate our

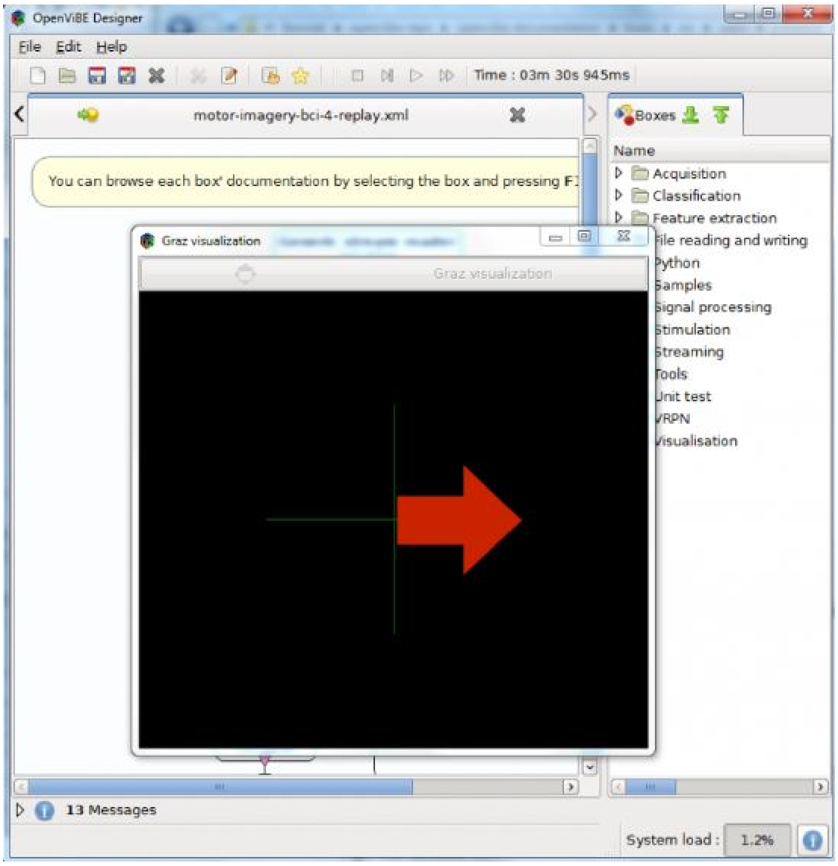

Figure 7.1: The MI classifer from the OpenVibe open-source package [22]

research on the real-world stress factors to improve the MI paradigm to function in a real-world setting.

7.2.3 Project 1: Swarm Drones and Ground-based Robots

7.2.3.1 Introduction

The basis of this project was to create an MI Robot prototype that we can use as a testing platform for future stress robotic experiments. We tested two robot platform types, a turtlebot3 (ground base robot) that operated on the Robot Operating System (ROS) package and a swarm of drones (three DJI sparks).

7.2.3.2 Design

The overall design of the system uses generic open sources packages. The challenge was the integration of these packages and tailoring them to the specific application.

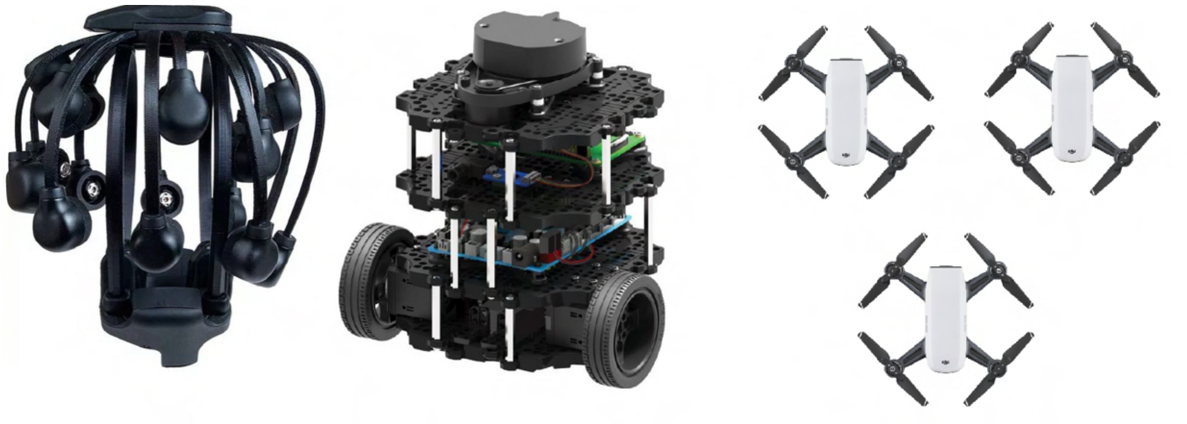

Figure 7.2: The Cognionics quick-30 dry EEG cap, a turtlebot3 robot [185], and three DJI sparks used

The MI paradigm and classifier was trained using the OpenVibe MI BCI classifier (see Figure 7.1) [22]. OpenVibe is natively compatible with our EEG system. The EEG signal was captured using the Cognionics Quick 30 dry EEG cap (see Figure 7.2) [**?**]. While dry electrodes are not as accurate as gel-based electrodes, the lack of needing to gel increases the general portability and reduces the time taken to set up the system. The MI classifier was trained using the dry electrodes, and then the output results were tailored to the intended application of either the ground robot or drone swarm.

The ground-based robot project used a turtlebot3 robot as the primary platform. The turtlebot3 is an educational robotics platform that is designed to be used with ROS packages [185]. We chose this platform because of the flexibility of integration and open-source availability of the ROS packages. The turtlebot3 moved through a maze (see Figure 7.3) using the ROS navigations package. The student Mena Balasy programmed a communications interface that streamed the MI classifier results to a ROS package. Then the package would translate the MI output into a goal position for the navigation package to direct the turtlebot3 to that location.

The drone swarm project used three DJI sparks to form a collective swarm. The concept of the drone swarm is to create a low level autonomous collaborative system that individually controls each drone, while the MI output controls the high-level formation of that swarm. The drones were controlled through TCP/IP with a master controller. The master controller would coordinate the drones to fly in three different formations. The MI left, and right output would signal the swarm to switch formations to the previous or next position (see Figure 7.4).

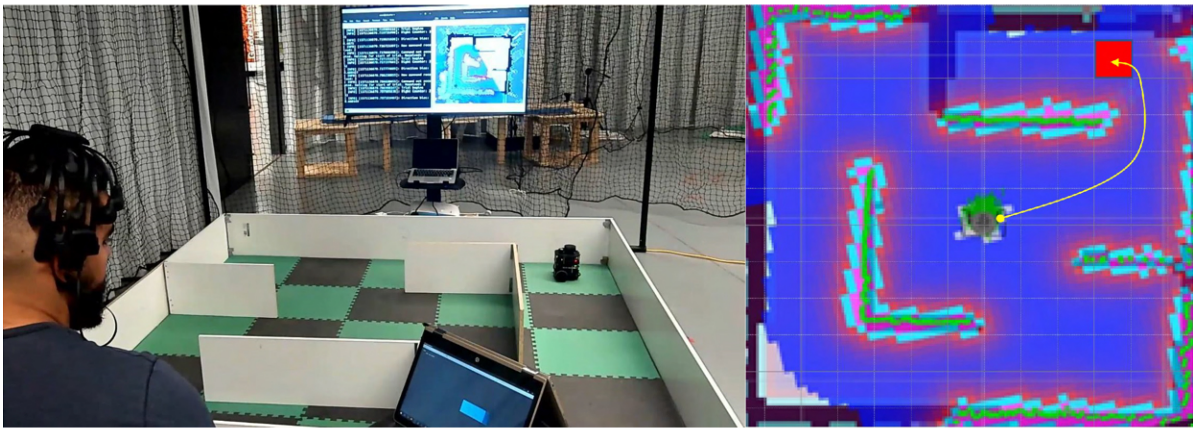

Figure 7.3: The Experimental set up for Mena Balasy's Turtlebot MI design.

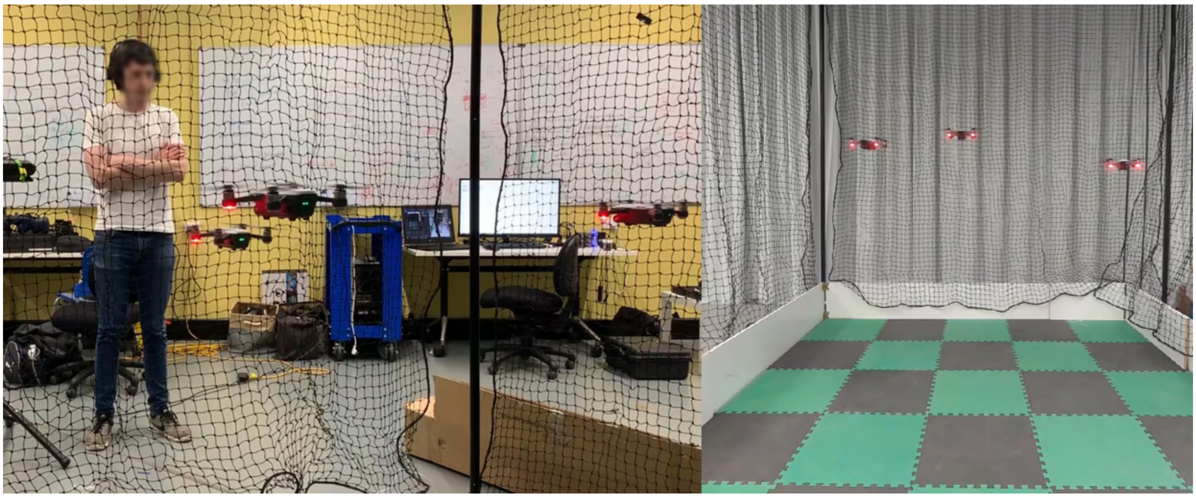

Figure 7.4: The Experimental set up for James Michell's swarm drones MI design.

7.2.3.3 Results and Findings

Both projects were successful in training an individualised classifier in a laboratory setting. The accuracy of the individualised classifier was up to 76.8%, which is accurate considering the BCI device was a dry sensor. The respective platforms of the turtlebot3 navigation package and the drone swarm high-level controller functioned well with a manual signal input (simulated MI left and right command). One observed issue was the synchronisation of the robotic systems and the MI classifier. The MI output samples at around 3-4 seconds faster than both the robotic platform communication and operating speed. We incorporated a communication protocol to prevent lag in communication. The protocol used a rolling average buffer that decreases the rate of communicated commands and ensures the robot has time to perform the task before receiving the following command. This method of down-sampling the rate of communication is a valuable finding for future iterations.

Another significant issue was the real-world performance of both systems. Both systems were displayed at a UTS capstone showcase with the working (in the laboratory) system for demonstrations. During the showcase, both systems suffered from heavy accuracy drops (incidentally, the drone caused an accident, inspiring our drone experiment in Chapter 5). The deterioration in the MI classifier is similar to previous studies [182, 211]. A mix of factors such as quality of hardware set up, the social anxiety of presenting in front of peers, and more complex visual and auditory noise, could have potentially contributed to the poor performance at the showcase. Incorporating our findings on stress factors in the BCI system could potentially improve the MI classifier to perform better in a real-world setting.

7.2.4 Project 2: Biofeedback

7.2.4.1 Introduction

The second exploratory project investigates the effectiveness of biofeedback techniques for stress modulation and improving cognitive performance. The motivation for this project was to investigate potential methods to regulate stress levels after being able to detect them in the EEG signal (Chapter 4, and Chapter 6). The relationship between stress and performance from the YD law suggests that the regulation of stress can contribute significantly to improving one's performance [194]. Biofeedback is a technique that involves using physiological measures such as HR, HRV, breathing rate (BR), or EDA to inform a person of their current physiological state [129]. By making individuals aware of their current stress state, they can implement other stress calming techniques to self-regulate their stress levels and improve cognitive performance. Studies by Prinsloo et al. [156, 157], Giggins et al. [71], and Nolan et al. [142] have all found success in the use of HR/HRV related biofeedback in reducing one's stress level. We reproduced the study by Prinsloo et al. [157] which used a Stroop task [161] to elicit a stress response.

7.2.4.2 Design

The overall experimental design is shown in Figure 7.5. We designed an essential Stroop task on Matlab with multiple levels of speed. Participants would do three sessions of the Stroop task switching between the colour and the word as targets. Each participant

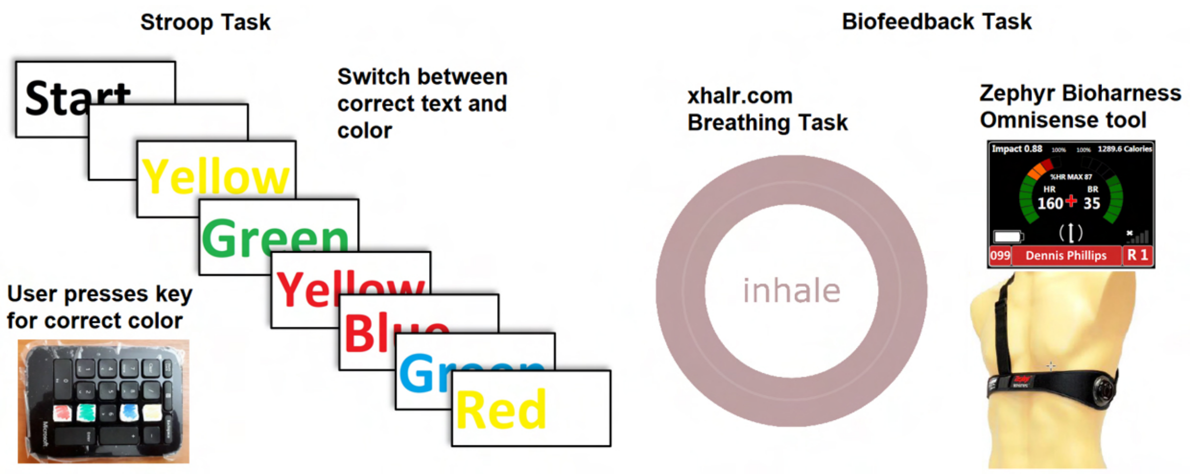

Figure 7.5: The Experimental set up for Vinayak Sharma's biofeedback experiment.

would experience a progressively more difficult (higher speed, more stressful) Stroop task; they would then either perform a breathing task, biofeedback (self-guided breathing), breathing task with biofeedback and breathing task with sham biofeedback. We chose to use an open-source online breathing task known as "xhalr" (https://xhalr.com/). The biofeedback displayed the participant's HR, HRV, and BR measured using the zephyr bioharness and its data display tool. We also incorporated sham feedback, which was the displayed tool with fake replayed values of Vinayak Sharma. Due to the pandemic's lockdown timing, the research student, Vinayak Sharma, could only collect performance and physiological data from 8 participants from neighbours and family members.

7.2.4.3 Results and Findings

Interestingly, the results of the 8 participants showed that the breathing task by itself outperformed (reaction time response and accuracy) the other conditions. We found little effect of biofeedback alone with similar results to the sham biofeedback. This would suggest that biofeedback without direction may not be an effective tool. One would likely have expected biofeedback with the breathing task as the most effective regulator. However, we believe it may be a case of increased workload with the biofeedback acting as a distraction from the breathing task instead of improving one's stress level.

Our results suggest that biofeedback is ineffective in our case, with the breathing task (reducing stress) being the more effective method of improving performance. The reason why biofeedback may not have been effective could be due to the precise nature of biofeedback. The Stroop task is quite a simple task, with the participants only needed to calm (lower their stress) to improve their performance. The breathing would seemingly accomplish this task with or without biofeedback. Biofeedback may be more beneficial in more complex tasks that require both up and down-regulation of stress levels. This limitation was highlighted by Tolin et al. [195]. The authors criticised that biofeedback should not be used in non-specific cases. Instead, biofeedback works best as a longitudinal method with direction therapy for the up and downregulation of stress.

Although, our study did have significant sample size limitations. We ultimately chose to abandon the incorporation of biofeedback as traditional methods of self stress regulation are seemingly superior.

7.2.5 Realistic BCI Training Environments

Figure 7.6 depicts a conceptualised image integrating our stress research into a drone control system (such as our MI control system). One of the primary goals of this research is the further innovate the experimental designs used for research. The previous works [73, 74] has shown that mobile BCI studies are possible, enabling researchers to investigate a more comprehensive array of research topics such as spatial navigation [55], VR based interactions [44], and balance [147]. We hope that the VR heights and drone collision experimental designs can progressively increase the realism of BCI experimental designs. We propose that the height exposure experiment be used for future BCI task training. Based on our findings, we know that stress will decrease the P300 peak amplitude at some level. A P300 based BCI classifier trained in a neutral controlled environment would likely have a distinct and optimal peak amplitude that is easy to classify with great accuracy. As shown in our experiment, this form of classifier would likely fail as the moment the individual experiences a stress response, the peak amplitude would decrease, and the classifier will likely decrease inaccuracy. Thus, by the progressively increased scope of real-world factors in BCI system testing, we can iteratively improve the robustness of the BCI classifier to be closer to real-world integration. Future BCI classifiers should incorporate noise factors within the training stage.

7.2.6 Drone Collision recognition

The reliable cognitive conflict response during drone collisions (Chapter 6) may prove immensely beneficial in designing a drone behaviour error correction system (Figure 7.7). Singh et al. [183] found success in using the ERP negativity to predict incorrect/undesired behaviours in a robotic arm. The negativity and theta synchronisation is a reliable

Figure 7.6: A conceptual image of integrating the height exposure experiment with drone BCI control.

indicator for incongruent behaviour. We have determined that the drone collision event is indeed an incongruent behaviour, we can design a classifier to detect the ERP and ERSP responses and appropriately adjust the drone's behaviour to prevent the collision. The main challenge of translating the findings of Chapter 6 into a real-time classifier is the source localisation method. Our data analysis identified the IC source after extensive preprocessing, running ICA to find ICs, and clustering the ICs. This form of data processing would not be practical in a real-time or online BCI system because of the computational load and the time required to complete the process. One option is it uses the inferred knowledge from our component analysis and perform channel based analysis using electrodes within our region of interest. However, this methodology is often problematic because IC clusters sources do not always directly translate into electrode positions [143]. A better option is to investigate tools such as Online Recursive Independent Component Analysis (ORICA) [91] and Real-time EEG Source-mapping Toolbox (REST) [151] for real-time source mapping of ICs.

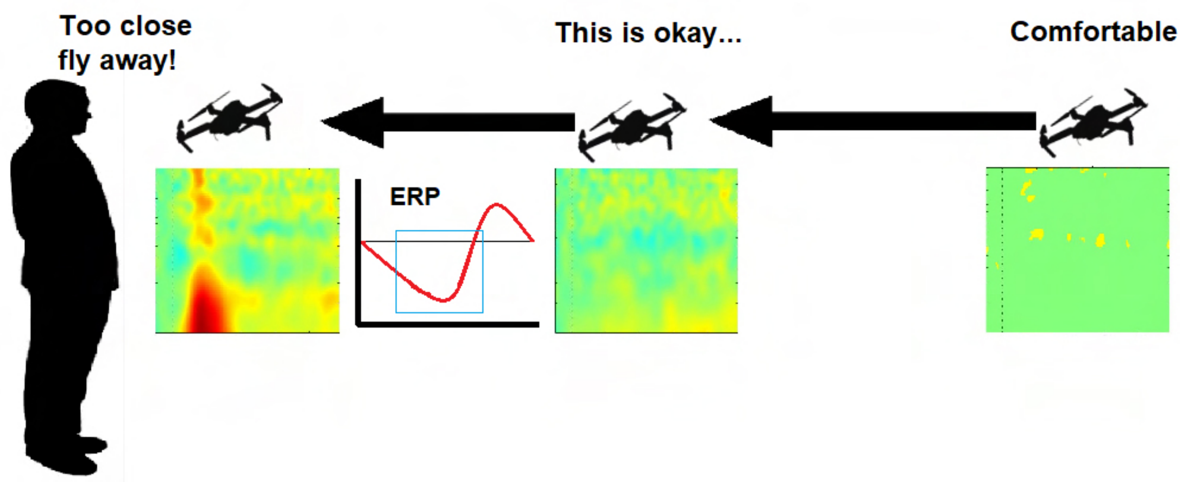

Figure 7.7: A conceptual image of drone EEG based collision detection.

7.3 Adaptive and Closed-Loop BCI

7.3.1 Adaptive BCI

An alternative method to improve a BCI classifier is to incorporate an adaptive BCI component within the system. Adaptive BCI is the concept of detecting a change in the state within the EEG and applying signal modulation to correct the EEG data before inputting it into a classifier [176]. The rationale behind adaptive BCI is to use computational intelligence and multi-modal data to improve the EEG raw data to minimise the need for the user to change their behaviour or cognitive state. We believe this method is an effective method of improving BCI performance compared to alternatives of using biofeedback to regulate stress (which we found to be ineffective). Studies such as Lotte et al. [119] and Vidaurre et al. [199] has found success in implement adaptive BCI techniques to improve the accuracy of signal classification.

7.3.2 Stressed-related Adaptive BCI

In the context of our findings, we could use RT and physiological measures to detect when a person is likely to exhibit decreased P300 amplitude due to stress (thus decreasing classifier performance), and create an adaptive BCI mechanism that would adjust (increase signal scale/resolution) the P300 response amplitude to maintain consistent classifier accuracy (see Figure 7.8). Another hypothetical implementation is drone collision detection. Any BCI drone control system would potentially have the issue of a

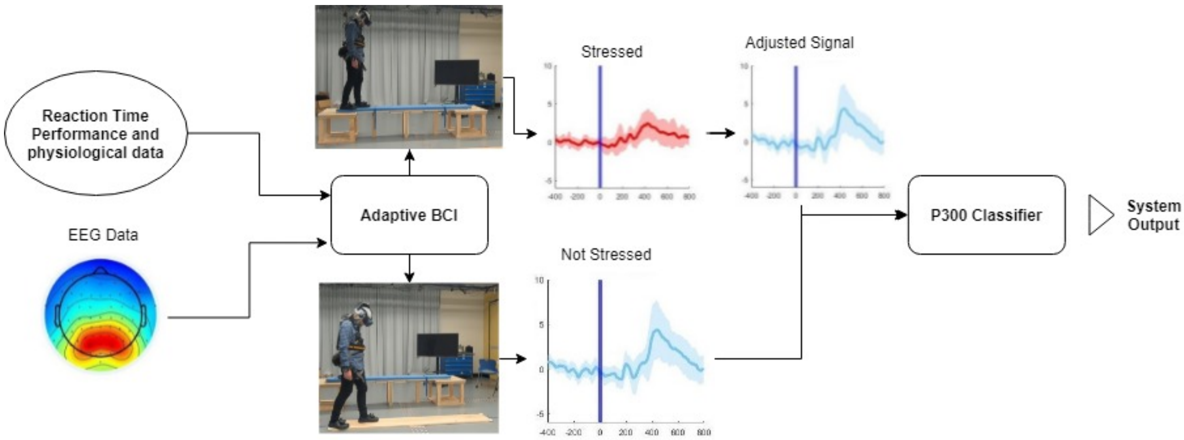

Figure 7.8: A conceptual image of a stress based adaptive BCI mechanism.

drone flying near the operator. If the drone flies too close, an operator might experience a stress response, further decreasing performance. An adaptive BCI system could detect the cognitive conflict response and adjust that portion of the signal to prevent inaccurate or dangerous behaviour.

7.3.3 Closed-Loop BCI

Closed-loop BCI is a system design that monitors an individual neurological state and produces regulatory feedback to enhance or improve the individual's cognitive performance [206]. Closed-loop BCI systems require an accurate detection (through machine learning or deep learning techniques) of a specific cognitive state and a known reliable technique to modulate that specific cognitive state [191]. A notable example is the musical closed-loop BCI by Ehrlich et al. [59], their system utilised various musical cues to modulate and improve the user's emotional state (positive emotions). Another example is the Friedrich et al. [67] closed-loop BCI that incorporated biofeedback elements to improve cognitive performance.

7.3.4 Closed-Loop Stress Regulation

In the context of physiological stress, a closed-loop BCI system could up and down-regulate (YD law) the user's stress levels to optimise cognitive performance. From our previous findings from the biofeedback experiment (Figure 7.5), we believe that biofeedback alone may not be sufficient in regulating the stress levels of an individual. One potential regulation method is to incorporate neurofeedback techniques to modulate the user's stress levels. Faller et al. [64] successfully employed this technique by using EEG, pupil dilation, and HR data to detect the subject's arousal level and use auditory cues (synthetic heartbeats) to up and down-regulate their stress level for optimised performance. In future experiments, we can test this approach in our height experimental design (or other PSS paradigms) to counteract the effects of stress and improve cognitive (P300) performance. We speculate that closed-loop BCI systems may only apply to PSS situations due to the sustained nature of the stress response. A DSS response would not need an external regulation due to the ephemeral and transient nature of the stress response. A shorter stress response implies that our body has naturally regulated and reduced the stress level to a healthy level.

7.4 Key Points

Our research findings from the previous chapters and our exploratory research provide a strong foundation for future projects and applications to improve the robustness of BCI systems. The exploratory projects found that MI is a valuable paradigm for controlling various robotic systems. On the other hand, our biofeedback experiment suggested that biofeedback alone may not be effective at regulating stress compared to traditional regulation techniques. In the future, we plan to use the reliable EEG response found in the drone experiment (Chapter 6) could be used as a drone behaviour error detection system to prevent drones from unstable behaviours. An alternative method could be using adaptive BCI and closed-loop BCI techniques to detect the biomarkers that we found (Chapter 4 and Chapter 6) to modulate the EEG signal or stress level to improve the BCI performance. Overall, there is a multitude of potential future research to be done by integrating and combining our research findings in this thesis with our exploratory projects to further our understanding of the real-world BCI system.

Chapter 8

Conclusions

8.1 Key Findings

In Chapter 3, we outlined the rationale and methodology of the experiments. These sections ensure that our experiment can (and should) be reproduced by other researchers to validate or offer counter-evidence to our findings. In that chapter, our key findings were that the combination of extreme virtual height and physical elevation was effective in producing a physical stress response shown through the significant increase in both questionnaire ratings, behavioural (gait) and physiological (HR, and EDA) measures. We found that a lower virtual height the physical elevation was a more negligible effect as the questionnaire and behavioural results showed an increase in stress, but the physiological data did not. Overall, our novel experiment design incorporating an elevated physical platform can produce a reliable stress response for BCI research.

In Chapter 4, we elaborate on our further investigation into how stress affects cognitive performance, specifically P300 behaviour. Our results may provide an explanation that resolves the conflicting findings of past literature. We categorised two groups within our participant that both exhibit behaviour consistent with either side of the previous results by splitting the participants using the RT measure. We found that one group (Group 1) exhibited behaviour, suggesting that stress decreases P300 performance and the other group (Group 2) showed no change in behaviour. By assessing the difference between the two groups, we found a relationship between a decrease in P300 performance and an increase in frontal-parietal beta power. We speculate that both groups are stressed, but Group 1 found the task more difficult when under stress, resulting in a deterioration in performance. Future assessments in stress and cognitive performance should consider RT performance and beta power as useful metrics when a participant is likely to experience a decrease in P300 and worse performance.

In Chapter 5, we explain the motivation, rationale, and methodology of our drone experiment. This experiment is unique as no other previous works have explored direct drone collision before. We have detailed the steps of the experiment in order for other researchers to reproduces the setup. Our questionnaire responses (SAM, DASS, and threat scale) indicates that the participants felt a stress response during drone collision. On the other hand, we did not detect any physiological change. Overall, we speculate the lack of physiological change is due to the short period of the stress response for the DSS paradigm when compared to the longer sustained response in PSS paradigms. Future iterations of this research should consider either using a measurement device with a higher sampling rate or a measurement that is more sensitive to a smaller temporal frame (such as EEG).

In Chapter 6, we further investigated and presented our EEG findings of the drone collision experiment. Our results found a distinct cognitive conflict behaviour in both the ERP and ERSP data during drone collisions. This is consistent with some similar studies and suggests that drone collisions are incongruent to the participants. This finding is valuable because it provides a clear marker that can detect and classify drone collisions or when a drone is flying too close to the participant.

For ***RQ1***, we explored two different forms of stress that we named PSS and DSS. Both experiments were able to reliable induce a stress response in different ways. The heights exposure (PSS) experiment was able to produce stress at extreme virtual heights reliably. On the other hand, we did not find change in our physiological (HR, HRV, and EDA) indicators of stress for the drone collision (DSS) experiment. This lack of physiological change could be due to the limitations of our measurement devices or the measurement being insensitivity to small and sudden changes. The questionnaire and EEG results indicated a heightened activity response from the drone collision, which we speculate signifies a shorter and more transient stress response. From the results, we believe that both experiments generate a stress response. However, the PSS paradigms produce a sustained stress response, and the DSS paradigms produce a short-term stress response.

For ***RQ2***, the height exposure experiments found that stress can modulate the difficulty of a task, decreasing the P300 peak amplitude. This provides an intermediary explanation that can account for the contradictory findings of previous research. We found that RT, physiological response, and frontal-parietal beta power indicates when the P300 peak amplitude will decrease. These indicators are consistent with being distracted and increased allocation to attention and memory resources.

The drone collision exhibited a clear cognitive conflict response suggesting drone collisions are an incongruent behaviour. We did not find traditional P300 or EEG based stress behavioural. However, this is likely due to the difference between DSS and PSS based responses. Most previous research on EEG-based stress markers utilise PSS-based paradigms that measure a sustained stress response. In the case of DSS, we only found the distinct cognitive conflict behaviour, which may infer that the participant did not anticipate or desire the drone collision behaviour. Other DSS paradigms may likely produce similar cognitive conflict responses as this experiment.

8.2 Key Contributions

- We presented two stress elicitation paradigms that can be used for future more translatable BCI research and classifier training
- Our height exposure experiment found that incorporating physical and virtual elevation creates a realistic and reliable method of stress elicitation.
- We determined that the participant's P300 behaviour when stressed is determined by their RT performance and perceived difficulty of the task. The P300 amplitude will decrease when participants perform worse when under stress.
- Our drone collision experiment found that drone collisions can produce a short-term stress level similar to other DSS paradigms.
- We found a reliable biomarker for drone collisions events that can be used to classify erroneous drone behaviour and improve future BCI drone-related technologies.

8.3 Summary and Future Works

This thesis has outlined the background to our research questions, our methodology in our investigations, and discussed our findings with potential future applications. We have found that height exposure and drone collisions are reliable methods of eliciting a stress response. Using the stress experiments, we explored the effects of stress on brain response and cognitive performances. We found that prolonged stress (PSS) can modulate the perceived difficulty and effort required for a task, therefore, lowering the P300 peak amplitude (distracted state) and increasing the frontal-parietal beta power (requiring more memory resources). In contrast, a short-term stress response does not exhibit these behaviours. This indicates that physiological stress can be elicited and responded to in various forms. Future BCI systems should incorporate these factors and multi-modal metrics during development to counteract the effects of stress and improve BCI system performance.

We believe that future research into the similarities and differences between PSS and DSS type paradigms will yield interesting findings into the effectiveness of the different stress paradigms. While individual paradigms have previously been extensively investigated, very few studies compare the various paradigms and assess the effectiveness of each paradigm. We will need to understand both PSS and DSS type stressors better to build more realistic stress elicitation paradigms that are more similar to the stresses of daily life.

We also propose that future stress-related research can apply the methodology of dividing participants based on RT performance during stressful conditions to differentiate the participants who were more susceptible to the stressors. This method may yield further insights into how stress affects cognitive performance on an individual level. These findings from the stress experiment can be further applied to adaptive or closed-loop BCI systems to provide real-time adjustments to the EEG signal by either modulating the signal or the user's cognitive state. Applying these techniques can potentially build a more robust BCI system that incorporates the potential changes in the user's emotional state when operating in the real world.

Appendix A

Appendix

A.1 Questionnaires used in the Experiments

A.1.1 SAM

Use the following diagram to rate how you felt during each trial.

Emotional Valence (Negative vs Positive)

- Assesses your pleasure during the trial
- First picture: distressed, panicked, irritated etc.
- Last picture: elated, fun, happiness, satisfaction etc.

Arousal (Low vs High)

- Assess your stress levels during the trial
- First picture: calm, relaxed, bored etc.
- Last picture: bursting with arousal, stressed, excited, agitated, anger etc.

Dominance (Low vs High)

- Assess your level of control during the trial
- First picture: dominant, in control of the situation, dominant, decisive etc.
- Last picture: lack of control, intimidated, submissive etc.

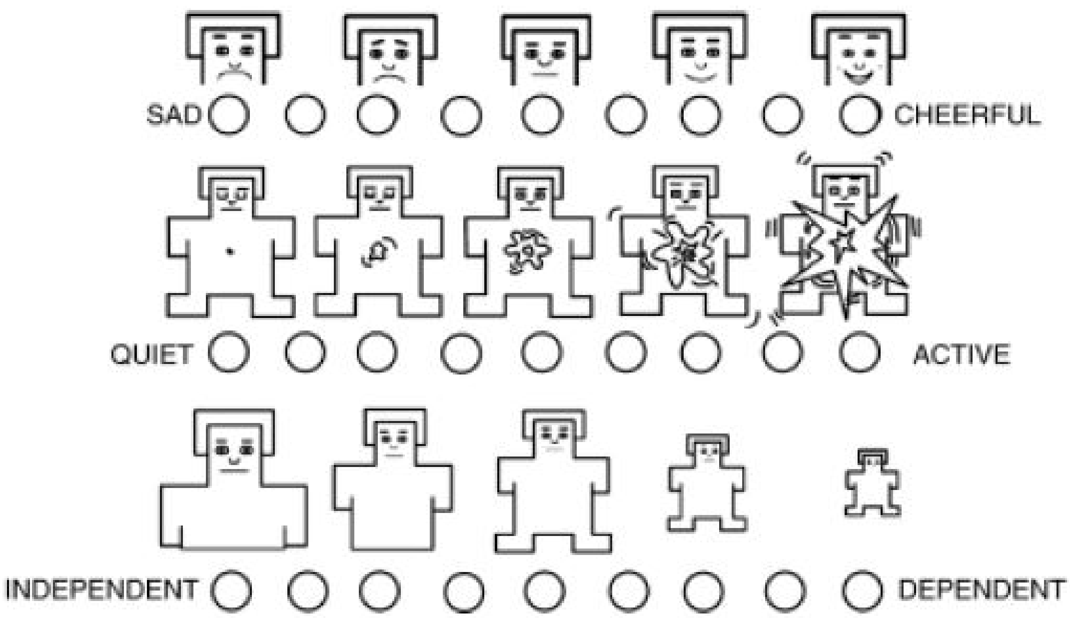

Figure A.1: The SAM questionnaire used in the Heights (Arousal Only) and Drone Experiment

A.1.2 Modified DASS Questionnaire

The following Questionnaire was used between conditions during the rest period.
Rating Scale:

- 0- Did not apply to me at all
- 1- Applied to me to some degree, or some of the time
- 2- Applied to me to a considerable degree, or a good part of time
- 3- Applied to me very much, or most of the time

DASS Question

- **A** - I was aware of dryness of my mouth
- **A** - I experienced breathing difficulty (eg, excessively rapid breathing, breathlessness in the absence of physical exertion)
- **A** - I had a feeling of shakiness or trembling
- **S** - I found it difficult to relax
- **S** - I was most relieved when the walks ended
- **S** - I felt that I was using a lot of nervous energy
- **A** - I found myself getting impatient when I was doing the tasks
- **S** - I had a feeling of faintness

Bibliography

[1] S. N. Abdulkader, A. Atia, and M.-S. M. Mostafa, *Brain computer interfacing: Applications and challenges*, Egyptian Informatics Journal, 16 (2015), pp. 213–230.

[2] P. Abtahi, D. Y. Zhao, J. L. E, and J. A. Landay, *Drone near me: Exploring touch-based human-drone interaction*, ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 1 (2017), pp. 1–8.

[3] F. A. Administration, *Recreational flyers and modeler community-based organizations*, Dec. 2020.

[4] S. P. Ahlfors, J. Han, J. W. Belliveau, and M. S. Hämäläinen, *Sensitivity of MEG and EEG to source orientation*, Brain Topography, 23 (2010), pp. 227–232.

[5] S. Alberto, *A technocritical review of drones crash risk probabilistic consequences and its societal acceptance.*, in RIMMA 2014, Berlin, Nov. 2015, Lecture notes in information sciences, pp. 27–38.

[6] A. P. Allen, P. J. Kennedy, S. Dockray, J. F. Cryan, T. G. Dinan, and G. Clarke, *The trier social stress test: Principles and practice*, Neurobiology of Stress, 6 (2017), pp. 113–126.
SI:Stressors in animals.

[7] N. A. Aminabadi, L. Erfanparast, A. Sohrabi, S. G. Oskouei, and A. Naghili, *The impact of virtual reality distraction on pain and anxiety during dental treatment in 4-6 year-old children: a randomized controlled clinical trial.*, Journal of Dental Research, Dental Clinics, Dental Prospects, 6 (2012).

[8] A. Arza, J. M. Garzón-Rey, J. Lázaro, E. Gil, R. Lopez-Anton, C. de la Camara, P. Laguna, R. Bailon, and J. Aguiló, *Measuring acute stress*

response through physiological signals: towards a quantitative assessment of stress, Medical & Biological Engineering & Computing, 57 (2018), pp. 271–287.

[9] N. S. Asjad, H. Adams, R. Paris, and B. Bodenheimer, *Perception of height in virtual reality: A study of climbing stairs*, in Proceedings of the 15th ACM Symposium on Applied Perception, SAP '18, New York, NY, USA, 2018, Association for Computing Machinery.

[10] L. Avila and M. Bailey, *Virtual reality for the masses*, IEEE Computer Graphics and Applications, 34 (2014), pp. 103–104.

[11] R. Avram, G. H. Tison, K. Aschbacher, P. Kuhar, E. Vittinghoff, M. Butzner, R. Runge, N. Wu, M. J. Pletcher, G. M. Marcus, and J. Olgin, *Real-world heart rate norms in the health eHeart study*, npj Digital Medicine, 2 (2019).

[12] B. L. Baker, D. C. Cohen, and J. T. Saunders, *Self-directed desensitization for acrophobia*, Behaviour Research and Therapy, 11 (1973), pp. 79–89.

[13] L. F. Barrett, Z. Khan, J. Dy, and D. Brooks, *Nature of emotion categories: Comment on cowen and keltner*, Trends in Cognitive Sciences, 22 (2018), pp. 97–99.

[14] R. Bauer, M. Fels, M. Vukelić, U. Ziemann, and A. Gharabaghi, *Bridging the gap between motor imagery and motor execution with a brain–robot interface*, NeuroImage, 108 (2015), pp. 319–327.

[15] J. Y. Bennington and J. Polich, *Comparison of p300 from passive and active tasks for auditory and visual stimuli*, International Journal of Psychophysiology, 34 (1999), pp. 171–177.

[16] M. Biran and G. T. Wilson, *Treatment of phobic disorders using cognitive and exposure methods: A self-efficacy analysis*, Journal of Consulting and Clinical Psychology, 49 (1981).

[17] D. P.-O. Bos, *Eeg-based emotion recognition the influence of visual and auditory stimuli*, 2007.

[18] C. Botella, R. M. Banos, H. Villa, C. Perpina, and A. Garcia-Palacios, *Virtual reality in the treatment of claustrophobic fear: A controlled, multiple-baseline design*, Behavior Therapy, 31 (2000), pp. 583–595.

[19] D. A. Bowman and R. P. McMahan, *Virtual reality: How much immersion is enough?*, Computer, 40 (2007), pp. 36–43.

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

[21] P. Broadhurst, *The interaction of task difficulty and motivation: The yerkes-dodson law revived*, Acta Psychologica, 16 (1959), pp. 321–338.

[22] N. Brodu, F. Lotte, and A. Lecuyer, *Comparative study of band-power extraction techniques for motor imagery classification*, in 2011 IEEE Symposium on Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB), IEEE, Apr. 2011.

[23] F. Brooks, *What's real about virtual reality?*, IEEE Computer Graphics and Applications, 19 (1999), pp. 16–27.

[24] A.-M. Brouwer and M. A. Hogervorst, *A new paradigm to induce mental stress: the sing-a-song stress test (ssst)*, Frontiers in Neuroscience, 8 (2014), p. 224.

[25] N. Browarska, A. Kawala-Sterniuk, P. Chechelski, and J. Zygarlicki, *Analysis of brain waves changes in stressful situations based on horror game with the implementation of virtual reality and brain-computer interface system: a case study*, Bio-Algorithms and Med-Systems, 16 (2020).

[26] L. A. Brown, M. A. Polych, and J. B. Doan, *The effect of anxiety on the regulation of upright standing among younger and older adults*, Gait & Posture, 24 (2006), pp. 397–405.

[27] G. Burdea, *Haptic feedback for virtual reality*, In Virtual reality and prototyping workshop, 2 (1999).

[28] R. Carabalona, F. Grossi, A. Tessadri, P. Castiglioni, A. Caracciolo, and I. de Munari, *Light on! real world evaluation of a p300-based brain–computer interface (BCI) for environment control in a smart home*, Ergonomics, 55 (2012), pp. 552–563.

[29] G. Cattan, A. Andreev, and E. Visinoni, *Recommendations for integrating a p300-based brain-computer interface in virtual reality environments for gaming: An update*, Computers, 9 (2020), p. 92.

[30] J. R. Cauchard, J. L. E, K. Y. Zhai, and J. A. Landay, *Drone & me: An exploration into natural human-drone interaction*, in Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '15, New York, NY, USA, 2015, Association for Computing Machinery, pp. 361–365.

[31] J. F. Cavanagh, M. J. Frank, T. J. Klein, and J. J. Allen, *Frontal theta links prediction errors to behavioral adaptation in reinforcement learning*, NeuroImage, 49 (2010), pp. 3198–3209.

[32] N. A. Ceballos, R. J. Giuliano, N. Y. Wicha, and R. Graham, *Acute stress and event-related potential correlates of attention to alcohol images in social drinkers*, Journal of Studies on Alcohol and Drugs, 73 (2012), pp. 761–771.

[33] T. Charman, S. Hepburn, M. Lewis, M. Lewis, A. Steiner, S. J. Rogers, A. Elburg, M. Lewis, F. L. Kohl, J. Danforth, E. Rotheram-Fuller, J. Robinson, M. H. Charlop, C. A. Miltenberger, A. L. Greenberg, P. Brucker, M. Lestrud, A. Cassidy, M. Goyette-Ewing, R. Jordan, J. Rohrer, J. Rohrer, F. R. Volkmar, B. Aaronson, B. Aaronson, K. Ruedel, A. Feinstein, J. W. Thomas, F. Happé, D. Vogler-Elias, S. Begeer, L. Pellicano, W. Gilliam, N. Yirmiya, I. Seidman, B. Reichow, P. Wehman, P. Wehman, P. Targett, I. Isasa, P. El-Fishawy, C. Mukerji, L. Mottron, J. C. McPartland, P. Cavanagh, E. Van-Bergeijk, F. R. Volkmar, C. Barthold, A. R. Gupta, L. Croen, T. Fernandez, G. Barnes, R. Sidhu, R. Tuchman, A. Bonnin, D. M. Lickenbrock, K. Wang, J. Pritchard, M. Malady, M. J. Crowley, J. Molteni, C. A. Doyle, C. McDougle, S. A. Mason, P. Walsh, T. J. Glennon, K. A. Pelphrey, B. Aaronson, B. Aaronson, M. Miklos, M. Solomon, D. M. Lickenbrock, H. Morris, H. Morris, H. Morris, M. Moyle, S. Long, B. Reichow, F. R. Volkmar, M. J. Weiss, B. Auyeung, M. Lombardo, R. Knickmeyer, S. Baron-Cohen, A. Senju, L. Sterling, and F. Shic, *Error-related negativity*, in Encyclopedia of Autism Spectrum Disorders, Springer New York, 2013, pp. 1159–1160.

[34] G. P. CHROUSOS, *Stress and disorders of the stress system*, Nature Reviews Endocrinology, 5 (2009), pp. 374–381.

[35] T. W. CLEWORTH, B. C. HORSLEN, AND M. G. CARPENTER, *Influence of real and virtual heights on standing balance*, Gait & Posture, 36 (2012), pp. 172–176.

[36] M. X. COHEN, *A neural microcircuit for cognitive conflict detection and signaling*, 37 (2014), pp. 480–490.

[37] S. COHEN, R. C. KESSLER, AND L. U. GORDON, *Measuring stress: a guide for health and social scientists*, Oxford University Press, 1998.

[38] J. L. COLLINGER, B. WODLINGER, J. E. DOWNEY, W. WANG, E. C. TYLER-KABARA, D. J. WEBER, A. J. MCMORLAND, M. VELLISTE, M. L. BONINGER, A. B. SCHWARTZ, AND ET AL., *High-performance neuroprosthetic control by an individual with tetraplegia*, The Lancet, 381 (2013), pp. 557–564.

[39] M. D. COMERCHERO AND J. POLICH, *P3a and p3b from typical auditory and visual stimuli*, Clinical Neurophysiology, 110 (1999), pp. 24–30.

[40] A. CONTE, *Human Rights in the Prevention and Punishment of Terrorism: Commonwealth Approaches: The United Kingdom, Canada, Australia and New Zealand*, Springer Berlin Heidelberg, New Zealand, 2010.

[41] A. B. CORPORATION, *Australian triathlete injured after drone crash*.
Article Available: https://www.abc.net.au/news/2018-08-27/drone-users-breaking-laws-as-prosecution-difficult/10153104/. [Accessed: 30- Dec- 2020].

[42] B. B. CORPORATION, *Australian triathlete injured after drone crash*.
Article Available: https://www.bbc.com/news/technology-26921504/. [Accessed: 30-Dec- 2020].

[43] ——, *Drone' hits british airways plane approaching heathrow airport*.
Article Available: https://www.bbc.com/news/uk-36067591/. [Accessed: 30- Dec-2020].

[44] C. A. T. CORTES, H.-T. CHEN, D. L. STURNIEKS, J. GARCIA, S. R. LORD, AND C.-T. LIN, *Evaluating balance recovery techniques for users wearing head-mounted display in vr*, IEEE Transactions on Visualization and Computer Graphics, 27 (2021), pp. 204–215.

[45] K. DEBATTISTA, T. BASHFORD-ROGERS, C. HARVEY, B. WATERFIELD, AND A. CHALMERS, *Subjective evaluation of high-fidelity virtual environments for driving simulations*, IEEE Transactions on Human-Machine Systems, 48 (2018), pp. 30–40.

[46] J. DECETY, *The neurophysiological basis of motor imagery*, Behavioural Brain Research, 77 (1996), pp. 45–52.

[47] J. DECOSTER, M. GALLUCCI, AND A.-M. R. ISELIN, *Best practices for using median splits, artificial categorization, and their continuous alternatives*, Journal of Experimental Psychopathology, 2 (2011), pp. 197–209.

[48] A. DELORME, T. MULLEN, C. KOTHE, Z. A. ACAR, N. BIGDELY-SHAMLO, A. VANKOV, AND S. MAKEIG, *EEGLAB, SIFT, NFT, BCILAB, and ERICA: New tools for advanced EEG processing*, Computational Intelligence and Neuroscience, 2011 (2011), pp. 1–12.

[49] T. DEMIRALP, A. ADEMOGLU, Y. ISTEFANOPULOS, C. BASAR-EROGLU, AND E. BASAR, *Wavelet analysis of oddball p300*, International Journal of Psychophysiology, 39 (2001), pp. 221–227.

[50] C. DI LORETO, J.-R. CHARDONNET, J. RYARD, AND A. ROUSSEAU, *Woah: A virtual reality work-at-height simulator*, in 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), 2018, pp. 281–288.

[51] D. M. DIAMOND, A. M. CAMPBELL, C. R. PARK, J. HALONEN, AND P. R. ZOLADZ, *The temporal dynamics model of emotional memory processing: A synthesis on the neurobiological basis of stress-induced amnesia, flashbulb and traumatic memories, and the yerkes-dodson law*, Neural Plasticity, 2007 (2007), pp. 1–33.

[52] J. DIEMER, N. LOHKAMP, A. MUHLBERGER, AND P. ZWANZGER, *Fear and physiological arousal during a virtual height challenge effects in patients with acrophobia and healthy controls*, Journal of Anxiety Disorders, 37 (2016), pp. 30–39.

[53] A. M. DIEROLF, J. FECHTNER, R. BOHNKE, O. T. WOLF, AND E. NAUMANN, *Influence of acute stress on response inhibition in healthy men: An erp study*, Psychophysiology, 54 (2017), pp. 684–695.

[54] A. M. Dierolf, D. Schoofs, E.-M. Hessas, M. Falkenstein, T. Otto, M. Paul, B. Suchan, and O. T. Wolf, *Good to be stressed? improved response inhibition and error processing after acute stress in young and older men*, Neuropsychologia, 119 (2018), pp. 434–447.

[55] T.-T. N. Do, C.-T. Lin, and K. Gramann, *Human brain dynamics in active spatial navigation*, Scientific Reports, 11 (2021).

[56] L. D.P, *Regulating government and private use of unmanned aerial vehicles: Drone policymaking, law enforcement deployment, and privacy concerns.*, Analysis and Metaphysics, 18 (2019), pp. 16–22.

[57] J. L. E, I. L. E, J. A. Landay, and J. R. Cauchard, *Drone & wo*, in Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, ACM, May 2017.

[58] G. Edlinger, B. Z. Allison, and C. Guger, *How many people can use a BCI system?*, in Clinical Systems Neuroscience, Springer Japan, Dec. 2014, pp. 33–66.

[59] S. K. Ehrlich, K. R. Agres, C. Guan, and G. Cheng, *A closed-loop, music-based brain-computer interface for emotion mediation*, 14 (2019), p. e0213516.

[60] P. Emmelkamp, M. Krijn, A. Hulsbosch, S. de Vries, M. Schuemie, and C. van der Mast, *Virtual reality treatment versus exposure in vivo: a comparative evaluation in acrophobia*, Behaviour Research and Therapy, 40 (2002), pp. 509–516.

[61] P. M. Emmelkamp and M. Felten, *The process of exposure in vivo : cognitive and physiological changes during treatment of acrophobia*, Behaviour Research and Therapy, 23 (1985), pp. 219–223.

[62] L. H. Epstein, R. A. Paluch, L. E. Kalakanis, G. S. Goldfield, F. J. Cerny, and J. N. Roemmich, *How much activity do youth get? a quantitative review of heart-rate measured activity*, Pediatrics, 108 (2001), pp. e44–e44.

[63] M. Falkenstein, J. Hoormann, and J. Hohnsbein, *ERP components in go/nogo tasks and their relation to inhibition*, Acta Psychologica, 101 (1999), pp. 267–291.

[64] J. FALLER, J. CUMMINGS, S. SAPROO, AND P. SAJDA, *Regulation of arousal via online neurofeedback improves human performance in a demanding sensory-motor task*, Proceedings of the National Academy of Sciences, 116 (2019), pp. 6482–6490.

[65] A. J. FALLGATTER, D. BRANDEIS, AND W. K. STRIK, *A robust assessment of the NoGo-anteriorisation of p300 microstates in a cued continuous performance test*, Brain Topography, 9 (1997), pp. 295–302.

[66] R. L. FINN AND D. WRIGHT, *Privacy, data protection and ethics for civil drone practice: A survey of industry, regulators and civil society organisations, computer law and security review.*, Computer Law and Security Review., 32 (2016), pp. 577–586.

[67] E. V. C. FRIEDRICH, N. SUTTIE, A. SIVANATHAN, T. LIM, S. LOUCHART, AND J. A. PINEDA, *Brainâ€“computer interface game applications for combined neurofeedback and biofeedback treatment for children on the autism spectrum*, 7 (2014).

[68] A. FURDEA, S. HALDER, D. KRUSIENSKI, D. BROSS, F. NIJBOER, N. BIRBAUMER, AND A. KÜBLER, *An auditory oddball (p300) spelling system for brain-computer interfaces*, Psychophysiology, 46 (2009), pp. 617–625.

[69] L. GARCIA, M. ZAK, C. GRENIER, S. HANRIO, D. HENRY, R. RANDRIAMANANTENA, C. SEMAL, J. M. ANDRE, V. LESPINET-NAJIB, R. RON-ANGEVIN, AND ET AL., *Is stress state an important factor in the bci-p300 speller performance?*, Advances in Computational Intelligence Lecture Notes in Computer Science, (2019), pp. 442–454.

[70] G. GIANNAKAKIS, D. GRIGORIADIS, AND M. TSIKNAKIS, *Detection of stress/anxiety state from eeg features during video watching*, 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), (2015), pp. 6034–6037.

[71] O. M. GIGGINS, U. PERSSON, AND B. CAULFIELD, *Biofeedback in rehabilitation*, Journal of NeuroEngineering and Rehabilitation, 10 (2013), p. 60.

[72] K. GRAMANN, J. T. GWIN, N. BIGDELY-SHAMLO, D. P. FERRIS, AND S. MAKEIG, *Visual evoked responses during standing and walking*, Frontiers in Human Neuroscience, 4 (2010).

[73] K. Gramann, J. T. Gwin, D. P. Ferris, K. Oie, T.-P. Jung, C.-T. Lin, L.-D. Liao, and S. Makeig, *Cognition in action: imaging brain/body dynamics in mobile humans*, Reviews in the Neurosciences, 22 (2011).

[74] K. Gramann, F. U. Hohlefeld, L. Gehrke, and M. Klug, *Human cortical dynamics during full-body heading changes*, (2018).

[75] Y. Granovsky, E. Sprecher, J. Hemli, and D. Yarnitsky, *P300 and stress in mild head injury patients*, Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section, 108 (1998), pp. 554–559.

[76] H. M. Gray, N. Ambady, W. T. Lowenthal, and P. Deldin, *P300 as an index of attention to self-relevant stimuli*, Journal of Experimental Social Psychology, 40 (2004), pp. 216–224.

[77] D. Gromer, M. Reinke, I. Christner, and P. Pauli, *Causal interactive links between presence and fear in virtual reality height exposure*, Frontiers in Psychology, 10 (2019), p. 141.

[78] S. S. G. M. Group, *The mystery of the gatwick drone*.
Article Available: https://www.theguardian.com/uk-news/2020/dec/01/the-mystery-of-the-gatwick-drone/. [Accessed: 30- Dec- 2020].

[79] C. Guger, G. Krausz, B. Allison, and G. Edlinger, *Comparison of dry and gel based electrodes for p300 brain computer interfaces*, Frontiers in Neuroscience, 6 (2012), p. 60.

[80] L. F. Haas, *Hans berger (1873-1941), richard caton (1842-1926), and electroencephalography*, Journal of Neurology, Neurosurgery & Psychiatry, 74 (2003), pp. 9–9.

[81] J. Healey and R. Picard, *Detecting stress during real-world driving tasks using physiological sensors*, IEEE Transactions on Intelligent Transportation Systems, 6 (2005), pp. 156–166.

[82] D. H. Hellhammer, S. Wast, and B. M. Kudielka, *Salivary cortisol as a biomarker in stress research*, Psychoneuroendocrinology, 34 (2009), pp. 163–171.

[83] C. HENDRIX AND W. BARFIELD, *Presence in virtual environments as a function of visual and auditory cues*, in Proceedings Virtual Reality Annual International Symposium '95, 1995, pp. 74–82.

[84] C. HENDRIX AND W. BARFIELD, *The sense of presence within auditory virtual environments*, Presence: Teleoper. Virtual Environ., 5 (1996), pp. 290–301.

[85] J. D. HENRY AND J. R. CRAWFORD, *The short-form version of the depression anxiety stress scales (dass-21): Construct validity and normative data in a large non-clinical sample*, British Journal of Clinical Psychology, 44 (2005), pp. 227–239.

[86] J. HERMAN, *Neural control of chronic stress adaptation*, Frontiers in Behavioral Neuroscience, 7 (2013), p. 61.

[87] P. HERMAN, G. PRASAD, T. M. MCGINNITY, AND D. COYLE, *Comparative analysis of spectral approaches to feature extraction for EEG-based motor imagery classification*, IEEE Transactions on Neural Systems and Rehabilitation Engineering, 16 (2008), pp. 317–326.

[88] C. S. HERRMANN, D. STRUBER, R. F. HELFRICH, AND A. K. ENGEL, *Eeg oscillations: From correlation to causality*, International Journal of Psychophysiology, 103 (2016), pp. 12–21.
Research on Brain Oscillations and Connectivity in A New Take-Off State.

[89] E. HINNERK AND V. VOGELPOHL, *Drone acceptance and noise concerns-some findings.*, in 53rd International Symposium on Aviation Psychology, Dayton, OH, USA, July 2019, Wright State University, pp. 199–204.

[90] L. HODGES, R. KOOPER, T. MEYER, B. ROTHBAUM, D. OPDYKE, J. DE GRAAFF, J. WILLIFORD, AND M. NORTH, *Virtual environments for treating the fear of heights*, Computer, 28 (1995).

[91] S.-H. HSU, T. MULLEN, T.-P. JUNG, AND G. CAUWENBERGHS, *Online recursive independent component analysis for real-time source separation of high-density EEG*, in 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, Aug. 2014.

[92] J. HVASS, O. LARSEN, K. VENDELBO, N. NILSSON, R. NORDAHL, AND S. SERAFIN, *Visual realism and presence in a virtual reality game*, in 2017 3DTV

Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), 2017, pp. 1–4.

[93] D. Iacobucci, S. S. Posavac, F. R. Kardes, M. J. Schneider, and D. L. Popovich, *Toward a more nuanced understanding of the statistical properties of a median split*, Journal of Consumer Psychology, 25 (2015), pp. 652–665.

[94] D. Irimia, R. Ortner, G. Krausz, C. Guger, and M. Poboroniuc, *BCI application in robotics control*, IFAC Proceedings Volumes, 45 (2012), pp. 1869–1874.

[95] N. Jake and T. Gorichanaz, *Trust as an ethical value in emerging technology governance: The case of drone regulation.*, Technology in Society., 59 (2019), p. 101131.

[96] R. Janapati, V. Dalal, N. Govardhan, and R. S. Gupta, *Review on EEG-BCI classification techniques advancements*, IOP Conference Series: Materials Science and Engineering, 981 (2020), p. 032019.

[97] C. Jiang, T. W. Buchanan, Z. Yao, K. Zhang, J. Wu, and L. Zhang, *Acute psychological stress disrupts attentional bias to threat-related stimuli*, Scientific Reports, 7 (2017).

[98] E. Jodo and K. Inoue, *Effects of practice on the p300 in a go/nogo task*, Electroencephalography and Clinical Neurophysiology, 76 (1990), pp. 249–257.

[99] L. R. Johnson, *Editorial: How fear and stress shape the mind*, Frontiers in Behavioral Neuroscience, 10 (2016).

[100] S.-M. Kamp, G. Forester, C. C. Vatheuer, and G. Domes, *Stress effects on the oddball p300 and n2 in males and females*, Biological Psychology, 162 (2021), p. 108095.

[101] T. B. L. KIRKWOOD, P. KAPAHI, and D. P. SHANLEY, *Evolution, stress, and longevity*, Journal of Anatomy, 197 (2000), pp. 587–590.

[102] C. Kirschbaum, K.-M. Pirke, and D. H. Hellhammer, *The 'trier social stress test' – a tool for investigating psychobiological stress responses in a laboratory setting*, Neuropsychobiology, 28 (1993), pp. 76–81.

[103] S. KISHORE, M. GONZÁLEZ-FRANCO, C. HINTEMÜLLER, C. KAPELLER, C. GUGER, M. SLATER, AND K. J. BLOM, *Comparison of SSVEP BCI and eye tracking for controlling a humanoid robot in a social environment*, Presence: Teleoperators and Virtual Environments, 23 (2014), pp. 242–252.

[104] J. KLAWOHN, N. J. SANTOPETRO, A. MEYER, AND G. HAJCAK, *Reduced p300 in depression: Evidence from a flanker task and impact on ERN, CRN, and pe*, Psychophysiology, 57 (2020).

[105] S. C. KLEIH, T. KAUFMANN, C. ZICKLER, S. HALDER, F. LEOTTA, F. CINCOTTI, F. ALOISE, A. RICCIO, C. HERBERT, D. MATTIA, AND A. KÜBLER, *Out of the frying pan into the fire—the p300-based BCI faces real-world challenges*, in Brain Machine Interfaces: Implications for Science, Clinical Practice and Society, Elsevier, 2011, pp. 27–46.

[106] M. KLUG AND K. GRAMANN, *Identifying key factors for improving ICA-based decomposition of EEG data in mobile and stationary experiments*, European Journal of Neuroscience, (2020).

[107] B. KOPP, F. RIST, AND U. MATTLER, *N200 in the flanker task as a neurobehavioral tool for investigating executive control*, Psychophysiology, 33 (1996), pp. 282–294.

[108] M. KRIJN, P. M. EMMELKAMP, R. BIEMOND, C. DE WILDE DE LIGNY, M. J. SCHUEMIE, AND C. A. VAN DER MAST, *Treatment of acrophobia in virtual reality: The role of immersion and presence*, Behaviour Research and Therapy, 42 (2004), pp. 229–239.

[109] D. KRUSIENSKI, E. SELLERS, D. MCFARLAND, T. VAUGHAN, AND J. WOLPAW, *Toward enhanced p300 speller performance*, Journal of Neuroscience Methods, 167 (2008), pp. 15–21.

[110] D. J. KRUSIENSKI, E. W. SELLERS, F. CABESTAING, S. BAYOUDH, D. J. MCFARLAND, T. M. VAUGHAN, AND J. R. WOLPAW, *A comparison of classification techniques for the p300 speller*, Journal of Neural Engineering, 3 (2006), pp. 299–305.

[111] A. KUBLER, V. MUSHAHWAR, L. HOCHBERG, AND J. DONOGHUE, *Bci meeting 2005-workshop on clinical issues and applications*, IEEE Transactions on Neural Systems and Rehabilitation Engineering, 14 (2006), pp. 131–134.

[112] P. LARSSON, D. VASTFJALL, P. OLSSON, AND M. KLEINER, *When what you hear is what you see: Presence and auditory-visual integration in virtual environments*, in Proceedings of the 10th Annual International Workshop on Presence, Barcelona, 2007, International Society for Presence Research, pp. 11–18.

[113] C. LAVALLEE, *The eu policy for civil drones: the challenge of governing emerging technologies.*, Institute for European Studies Policy Brief, 1 (2019).

[114] R. S. LAZARUS, J. DEESE, AND S. F. OSLER, *The effects of psychological stress upon performance.*, Psychological Bulletin, 49 (1952), pp. 293–317.

[115] R. S. LAZARUS AND S. FOLKMAN, *Stress, appraisal, and coping*, Springer, 1984.

[116] M. LEE, G. BRUDER, AND G. WELCH, *The virtual pole: Exploring human responses to fear of heights in immersive virtual environments*, J. Virtual Real. Broadcast., 14 (2017).

[117] Y.-C. LEE AND F. K. WINSTON, *Stress induction techniques in a driving simulator and reactions from newly licensed drivers*, Transportation Research Part F: Traffic Psychology and Behaviour, 42 (2016), pp. 44–55.

[118] F. LOTTE, L. BOUGRAIN, A. CICHOCKI, M. CLERC, M. CONGEDO, A. RAKOTOMAMONJY, AND F. YGER, *A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update*, Journal of Neural Engineering, 15 (2018), p. 031005.

[119] F. LOTTE AND C. JEUNET, *Defining and quantifying users' mental imagery-based BCI skills: a first step*, Journal of Neural Engineering, 15 (2018), p. 046030.

[120] M. LOTZE AND U. HALSBAND, *Motor imagery*, Journal of Physiology-Paris, 99 (2006), pp. 386–395.

[121] P. LOVIBOND AND S. LOVIBOND, *The structure of negative emotional states: Comparison of the depression anxiety stress scales (dass) with the beck depression and anxiety inventories*, Behaviour Research and Therapy, 33 (1995), pp. 335–343.

[122] S. J. LUPIEN, B. S. MCEWEN, M. R. GUNNAR, AND C. HEIM, *Effects of stress throughout the lifespan on the brain, behaviour and cognition*, Nature Reviews Neuroscience, 10 (2009), pp. 434–445.

[123] P. LUU, D. M. TUCKER, AND S. MAKEIG, *Frontal midline theta and the error-related negativity: neurophysiological mechanisms of action regulation*, Clinical Neurophysiology, 115 (2004), pp. 1821–1835.

[124] D. MARSHALL, D. COYLE, S. WILSON, AND M. CALLAGHAN, *Games, gameplay, and bci: The state of the art*, IEEE Transactions on Computational Intelligence and AI in Games, 5 (2013), pp. 82–99.

[125] G. MATTHEWS, L. DORN, T. W. HOYES, D. R. DAVIES, A. I. GLENDON, AND R. G. TAYLOR, *Driver stress and performance on a driving simulator*, Human Factors: The Journal of the Human Factors and Ergonomics Society, 40 (1998), pp. 136–149.

[126] C. MCCARTHY, N. PRADHAN, C. REDPATH, AND A. ADLER, *Validation of the empatica e4 wristband*, in 2016 IEEE EMBS International Student Conference (ISC), 2016, pp. 1–4.

[127] K. A. MCGONAGLE AND R. C. KESSLER, *Chronic stress, acute stress, and depressive symptoms*, American Journal of Community Psychology, 18 (1990), pp. 681–706.

[128] M. MEEHAN, B. INSKO, M. WHITTON, AND F. P. BROOKS, *Physiological measures of presence in stressful virtual environments*, ACM Trans. Graph., 21 (2002), pp. 645–652.

[129] D. MEICHENBAUM, *Cognitive factors in biofeedback therapy*, Biofeedback and Self-Regulation, 1 (1976), pp. 201–216.

[130] S. L. MILLER, S. CHELIAN, W. MCBURNETT, W. TSOU, AND A. A. KRUSE, *An investigation of computer-based brain training on the cognitive and eeg performance of employees*, in 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2019, pp. 518–521.

[131] N. MINKLEY, T. P. SCHRADER, O. T. WOLF, AND W. H. KIRCHNER, *The socially evaluated cold-pressor test (secpt) for groups: Effects of repeated administration of a combined physiological and psychological stressor*, Psychoneuroendocrinology, 45 (2014), pp. 119–127.

[132] R. A. MIRANDA, W. D. CASEBEER, A. M. HEIN, J. W. JUDY, E. P. KROTKOV, T. L. LAABS, J. E. MANZO, K. G. PANKRATZ, G. A. PRATT, J. C. SANCHEZ,

AND ET AL., *Darpa-funded efforts in the development of novel brain-computer interface technologies*, Journal of Neuroscience Methods, 244 (2015), pp. 52–67.

[133] M. MIYAKOSHI, M. NOMURA, AND H. OHIRA, *An ERP study on self-relevant object recognition*, Brain and Cognition, 63 (2007), pp. 182–189.

[134] M. MOORE, *Real-world applications for brain-computer interface technology*, IEEE Transactions on Neural Systems and Rehabilitation Engineering, 11 (2003), pp. 162–165.

[135] A. MORAN AND H. O'SHEA, *Motor imagery practice and cognitive processes*, Frontiers in Psychology, 11 (2020).

[136] T. MULLEN, C. KOTHE, Y. M. CHI, A. OJEDA, T. KERTH, S. MAKEIG, G. CAUWENBERGHS, AND T.-P. JUNG, *Real-time modeling and 3d visualization of source dynamics and connectivity using wearable EEG*, in 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, July 2013.

[137] J. MUNZERT, B. LOREY, AND K. ZENTGRAF, *Cognitive motor processes: The role of motor imagery in the study of motor representations*, Brain Research Reviews, 60 (2009), pp. 306–326.

[138] J. MURASKIN, S. DODHIA, G. LIEBERMAN, J. O. GARCIA, T. VERSTYNEN, J. M. VETTEL, J. SHERWIN, AND P. SAJDA, *Brain dynamics of post-task resting state are influenced by expertise: Insights from baseball players*, Human Brain Mapping, 37 (2016), pp. 4454–4471.

[139] J. MURASKIN, J. SHERWIN, AND P. SAJDA, *Knowing when not to swing: EEG evidence that enhanced perception–action coupling underlies baseball batter expertise*, NeuroImage, 123 (2015), pp. 1–10.

[140] R. NAGAO, K. MATSUMOTO, T. NARUMI, T. TANIKAWA, AND M. HIROSE, *Ascending and descending in virtual reality: Simple and safe system using passive haptics*, IEEE Transactions on Visualization and Computer Graphics, 24 (2018), pp. 1584–1593.

[141] G. NAZARI, P. BOBOS, J. C. MACDERMID, K. E. SINDEN, J. RICHARDSON, AND A. TANG, *Psychometric properties of the zephyr bioharness device: a systematic review*, BMC Sports Science,Medicine and Rehabilitation, 10 (2018).

[142] R. P. NOLAN, M. V. KAMATH, J. S. FLORAS, J. STANLEY, C. PANG, P. PICTON, AND Q. R. YOUNG, *Heart rate variability biofeedback as a behavioral neurocardiac intervention to enhance vagal heart rate control*, American Heart Journal, 149 (2005), pp. 1137.e1–1137.e7.

[143] J. ONTON, M. WESTERFIELD, J. TOWNSEND, AND S. MAKEIG, *Imaging human EEG dynamics using independent component analysis*, Neuroscience & Biobehavioral Reviews, 30 (2006), pp. 808–822.

[144] S. PABST, M. BRAND, AND O. T. WOLF, *Stress and decision making: A few minutes make all the difference*, Behavioural Brain Research, 250 (2013), pp. 39–45.

[145] F. PALLAVICINI, A. FERRARI, A. PEPE, G. GARCEA, A. ZANACCHI, AND F. MANTOVANI, *Effectiveness of virtual reality survival horror games for the emotional elicitation: Preliminary insights using resident evil 7: Biohazard*, in Universal Access in Human-Computer Interaction. Virtual, Augmented, and Intelligent Environments, Springer International Publishing, 2018, pp. 87–101.

[146] P. F. PASQUINA, M. EVANGELISTA, A. CARVALHO, J. LOCKHART, S. GRIFFIN, G. NANOS, P. MCKAY, M. HANSEN, D. IPSEN, J. VANDERSEA, AND ET AL., *First-in-man demonstration of a fully implanted myoelectric sensors system to control an advanced electromechanical prosthetic hand*, Journal of Neuroscience Methods, 244 (2015), pp. 85–93.

[147] S. M. PETERSON, E. FURUICHI, AND D. P. FERRIS, *Effects of virtual reality high heights exposure during beam-walking on physiological stress and cognitive loading*, Plos One, 13 (2018).

[148] G. PFURTSCHELLER, C. BRUNNER, AND C. NEUPER, *Eeg-based brain-computer interfaces*, Oxford Medicine Online, (2017).

[149] G. PFURTSCHELLER AND C. NEUPER, *Motor imagery and direct brain-computer communication*, Proceedings of the IEEE, 89 (2001), pp. 1123–1134.

[150] T. W. PICTON, *The p300 wave of the human event-related potential*, Journal of Clinical Neurophysiology, 9 (1992).

[151] L. PION-TONACHINI, S.-H. HSU, S. MAKEIG, T.-P. JUNG, AND G. CAUWENBERGHS, *Real-time EEG source-mapping toolbox (REST): Online ICA and*

source localization, in 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, Aug. 2015.

[152] J. POLICH, *Updating p300: An integrative theory of p3a and p3b*, Clinical Neurophysiology, 118 (2007), pp. 2128–2148.

[153] J. POLICH AND A. KOK, *Cognitive and biological determinants of p300: an integrative review*, Biological Psychology, 41 (1995), pp. 103–146.

[154] J. POLICH AND C. MARGALA, *P300 and probability: comparison of oddball and single-stimulus paradigms*, International Journal of Psychophysiology, 25 (1997), pp. 169–176.

[155] N. PRATT, A. WILLOUGHBY, AND D. SWICK, *Effects of working memory load on visual selective attention: Behavioral and electrophysiological evidence*, Frontiers in Human Neuroscience, 5 (2011), p. 57.

[156] G. E. PRINSLOO, W. E. DERMAN, M. I. LAMBERT, AND H. G. L. RAUCH, *The effect of a single session of short duration biofeedback-induced deep breathing on measures of heart rate variability during laboratory-induced cognitive stress: A pilot study*, Applied Psychophysiology and Biofeedback, 38 (2013), pp. 81–90.

[157] G. E. PRINSLOO, H. G. L. RAUCH, M. I. LAMBERT, F. MUENCH, T. D. NOAKES, AND W. E. DERMAN, *The effect of short duration heart rate variability (HRV) biofeedback on cognitive performance during laboratory induced cognitive stress*, Applied Cognitive Psychology, 25 (2010), pp. 792–801.

[158] M. QI AND H. GAO, *Acute psychological stress promotes general alertness and attentional control processes: An erp study*, Psychophysiology, 57 (2020).

[159] M. QI, H. GAO, AND G. LIU, *The effect of mild acute psychological stress on attention processing: an erp study*, Experimental Brain Research, 236 (2018), pp. 2061–2071.

[160] M. N. RASTGOO, B. NAKISA, A. RAKOTONIRAINY, V. CHANDRAN, AND D. TJONDRONEGORO, *A critical review of proactive detection of driver stress levels based on multimodal measurements*, ACM Comput. Surv., 51 (2018).

[161] P. RENAUD AND J.-P. BLONDIN, *The stress of stroop performance: physiological and emotional responses to color word interference, task pacing, and pacing speed*, International Journal of Psychophysiology, 27 (1997), pp. 87–97.

[162] B. REUDERINK, C. MÜHL, AND M. POEL, *Valence, arousal and dominance in the EEG during game play*, International Journal of Autonomous and Adaptive Communications Systems, 6 (2013), p. 45.

[163] ROOTMOTION, *Rootmotion*.
Home-RootMotion Available: http://root-motion.com/. [Accessed: 24- Aug- 2020].

[164] T. ROS, M. J. MOSELEY, P. A. BLOOM, L. BENJAMIN, L. A. PARKINSON, AND J. H. GRUZELIER, *Optimizing microsurgical skills with EEG neurofeedback*, BMC Neuroscience, 10 (2009).

[165] B. O. ROTHBAUM, L. F. HODGES, R. KOOPER, D. OPDYKE, J. S. WILLIFORD, AND M. NORTH, *Effectiveness of computer-generated (virtual reality) graded exposure in the treatment of acrophobia*, American Journal of Psychiatry, 152 (1995), pp. 626–628.
PMID: 7694917.

[166] J. A. RUSSELL, *A circumplex model of affect.*, Journal of Personality and Social Psychology, 39 (1980), pp. 1161–1178.

[167] M. V. SANCHEZ-VIVES AND M. SLATER, *From presence to consciousness through virtual reality*, Nature Reviews Neuroscience, 6 (2005).

[168] A. SCHLÖGL, C. VIDAURRE, AND K.-R. MÜLLER, *Adaptive methods in BCI research - an introductory tutorial*, in Brain-Computer Interfaces, Springer Berlin Heidelberg, 2009, pp. 331–355.

[169] R. SCHMIDT, M. H. RUIZ, B. E. KILAVIK, M. LUNDQVIST, P. A. STARR, AND A. R. ARON, *Beta oscillations in working memory, executive control of movement and thought, and sensorimotor function*, The Journal of Neuroscience, 39 (2019), pp. 8231–8238.

[170] R. SCHNIEPP, G. KUGLER, M. WUEHR, M. ECKL, D. HUPPERT, S. HUTH, C. PRADHAN, K. JAHN, AND T. BRANDT, *Quantification of gait changes in subjects with visual height intolerance when exposed to heights*, Frontiers in Human Neuroscience, 8 (2014), p. 963.

[171] L. SCHROTH, *Drone manufacturer market shares: Dji leads the way in the us*.
Article Available: https://droneii.com/drone-manufacturer-market-shares-dji-leads-the-way-in-the-us/. [Accessed: 03- Jan- 2021].

[172] C. SCHUBERT, M. LAMBERTZ, R. NELESEN, W. BARDWELL, J.-B. CHOI, AND J. DIMSDALE, *Effects of stress on heart rate complexity a comparison between short-term and chronic stress*, Biological Psychology, 80 (2009), pp. 325–332.

[173] S. B. SCOTT, M. J. SLIWINSKI, J. A. MOGLE, AND D. M. ALMEIDA, *Age, stress, and emotional complexity: Results from two studies of daily experiences.*, Psychology and Aging, 29 (2014), pp. 577–587.

[174] E. W. SELLERS, D. J. KRUSIENSKI, D. J. MCFARLAND, T. M. VAUGHAN, AND J. R. WOLPAW, *A p300 event-related potential brain-computer interface (bci): The effects of matrix size and inter stimulus interval on performance*, Biological Psychology, 73 (2006), pp. 242–252.

[175] C. SETZ, B. ARNRICH, J. SCHUMM, R. LA MARCA, G. TROSTER, AND U. EHLERT, *Discriminating stress from cognitive load using a wearable eda device*, IEEE Transactions on Information Technology in Biomedicine, 14 (2010), pp. 410–417.

[176] P. SHENOY, M. KRAULEDAT, B. BLANKERTZ, R. P. N. RAO, AND K.-R. MULLER, *Towards adaptive classification for BCI*, Journal of Neural Engineering, 3 (2006), pp. R13–R23.

[177] J. SHERWIN, J. MURASKIN, AND P. SAJDA, *You can't think and hit at the same time: Neural correlates of baseball pitch classification*, Frontiers in Neuroscience, 6 (2012).

[178] A. L. SHILTON, R. LAYCOCK, AND S. G. CREWTHER, *The maastricht acute stress test (mast): Physiological and subjective responses in anticipation, and post-stress*, Frontiers in Psychology, 8 (2017), p. 567.

[179] L. M. SHIN AND I. LIBERZON, *The neurocircuitry of fear, stress, and anxiety disorders*, Neuropsychopharmacology, 35 (2009), pp. 169–191.

[180] A. D. S. SIERRA, C. S. AVILA, J. G. CASANOVA, AND G. B. D. POZO, *A stress-detection system based on physiological signals and fuzzy logic*, IEEE Transactions on Industrial Electronics, 58 (2011), pp. 4857–4865.

[181] P. I. SIMEONOV, H. HSIAO, B. W. DOTSONM, AND D. E. AMMONS, *Height effects in real and virtual environments*, Human Factors, 47 (2005), pp. 430–438.
PMID: 16170948.

[182] A. Singh, A. A. Hussain, S. Lal, and H. W. Guesgen, *A comprehensive review on critical issues and possible solutions of motor imagery based electroencephalography brain-computer interface*, Sensors, 21 (2021), p. 2173.

[183] A. K. Singh, S. Aldini, D. Leong, Y.-K. Wang, M. G. Carmichael, D. Liu, and C.-T. Lin, *Prediction error negativity in physical human-robot collaboration*, in 2020 8th International Winter Conference on Brain-Computer Interface (BCI), 2020, pp. 1–6.

[184] A. K. Singh, K. Gramann, H.-T. Chen, and C.-T. Lin, *The impact of hand movement velocity on cognitive conflict processing in a 3d object selection task in virtual reality*, NeuroImage, 226 (2021), p. 117578.

[185] D. Singh, E. Trivedi, Y. Sharma, and V. Niranjan, *TurtleBot: Design and hardware component selection*, in 2018 International Conference on Computing, Power and Communication Technologies (GUCON), IEEE, Sept. 2018.

[186] M. Slater, *A note on presence terminology*, Presence-Connect, (2003).

[187] M. Slater, P. Khanna, J. Mortensen, and I. Yu, *Visual realism enhances realistic response in an immersive virtual environment*, IEEE Computer Graphics and Applications, 29 (2009), pp. 76–84.

[188] R. J. Stone, *Haptic feedback: a brief history from telepresence to virtual reality*, in Haptic Human-Computer Interaction, S. Brewster and R. Murray-Smith, eds., Berlin, Heidelberg, 2001, Springer Berlin Heidelberg, pp. 1–16.

[189] A. Subasi and M. I. Gursoy, *EEG signal classification using PCA, ICA, LDA and support vector machines*, Expert Systems with Applications, 37 (2010), pp. 8659–8666.

[190] F.-T. Sun, C. Kuo, H.-T. Cheng, S. Buthpitiya, P. Collins, and M. Griss, *Activity-aware mental stress detection using physiological sensors*, in Mobile Computing, Applications, and Services, M. Gris and G. Yang, eds., Berlin, Heidelberg, 2012, Springer Berlin Heidelberg, pp. 211–230.

[191] A. Sundaresan, B. Penchina, S. Cheong, V. Grace, A. Valero-Cabré, and A. Martel, *Evaluating deep learning EEG-based mental stress classification in adolescents with autism for breathing entrainment BCI*, 8 (2021).

[192] J. TAELMAN, S. VANDEPUT, A. SPAEPEN, AND S. VAN HUFFEL, *Influence of mental stress on heart rate and heart rate variability*, in 4th European Conference of the International Federation for Medical and Biological Engineering, J. Vander Sloten, P. Verdonck, M. Nyssen, and J. Haueisen, eds., Berlin, Heidelberg, 2009, Springer Berlin Heidelberg, pp. 1366–1369.

[193] M. TANGERMANN, K.-R. MULLER, A. AERTSEN, N. BIRBAUMER, C. BRAUN, C. BRUNNER, R. LEEB, C. MEHRING, K. MILLER, G. MUELLER-PUTZ, G. NOLTE, G. PFURTSCHELLER, H. PREISSL, G. SCHALK, A. SCHLAGL, C. VIDAURRE, S. WALDERT, AND B. BLANKERTZ, *Review of the bci competition iv*, Frontiers in Neuroscience, 6 (2012), p. 55.

[194] K. H. TEIGEN, *Yerkes-dodson: A law for all seasons*, Theory & Psychology, 4 (1994), pp. 525–547.

[195] D. F. TOLIN, C. D. DAVIES, D. M. MOSKOW, AND S. G. HOFMANN, *Biofeedback and neurofeedback for anxiety disorders: A quantitative and qualitative systematic review*, in Advances in Experimental Medicine and Biology, Springer Singapore, 2020, pp. 265–289.

[196] S. S. TOMKINS AND B. SMITH, *Exploring affect*, (1995).

[197] L. T. TRUJILLO AND J. J. ALLEN, *Theta EEG dynamics of the error-related negativity*, Clinical Neurophysiology, 118 (2007), pp. 645–668.

[198] R. VERMA, Y. P. S. BALHARA, AND C. S. GUPTA, *Gender differences in stress response: Role of developmental and biological determinants*, Industrial psychiatry journal, 20 (2011), pp. 4–10.
22969173[pmid].

[199] C. VIDAURRE, A. SCHLOOGL, R. CABEZA, R. SCHERER, AND G. PFURTSCHELLER, *A fully on-line adaptive BCI*, IEEE Transactions on Biomedical Engineering, 53 (2006), pp. 1214–1219.

[200] T. G. M. VRIJKOTTE, L. J. P. V. DOORNEN, AND E. J. C. D. GEUS, *Effects of work stress on ambulatory blood pressure, heart rate, and heart rate variability*, Hypertension, 35 (2000), pp. 880–886.

[201] S. WALDERT, *Invasive vs. non-invasive neuronal signals for brain-machine interfaces: Will one prevail?*, Frontiers in Neuroscience, 10 (2016), p. 295.

[202] R. WHELAN, *Effective analysis of reaction time data*, The Psychological Record, 58 (2008), pp. 475–482.

[203] G. WILD, M. JOHN, AND G. BAXTER, *Exploring civil drone accidents and incidents to help prevent potential air disasters.*, Aerospace, 3 (2016), p. 3.

[204] J. WOLPAW, N. BIRBAUMER, W. HEETDERKS, D. MCFARLAND, P. PECKHAM, G. SCHALK, E. DONCHIN, L. QUATRANO, C. ROBINSON, AND T. VAUGHAN, *Brain-computer interface technology: a review of the first international meeting*, IEEE Transactions on Rehabilitation Engineering, 8 (2000), pp. 164–173.

[205] M. WUEHR, K. BREITKOPF, J. DECKER, G. IBARRA, D. HUPPERT, AND T. BRANDT, *Fear of heights in virtual reality saturates 20 to 40 m above ground*, Journal of Neurology, 266 (2019).

[206] R. XU, N. JIANG, C. LIN, N. MRACHACZ-KERSTING, K. DREMSTRUP, AND D. FARINA, *Enhanced low-latency detection of motor intention from eeg for closed-loop brain-computer interface applications*, IEEE Transactions on Biomedical Engineering, 61 (2014), pp. 288–296.

[207] A. YEH, P. RATSAMEE, K. KIYOKAWA, Y. URANISHI, T. MASHITA, H. TAKEMURA, M. FJELD, AND M. OBAID, *Exploring proxemics for human-drone interaction.*, in 5th international conference on human agent interaction, Bielefeld Germany, Oct. 2017, ACM, pp. 81–88.

[208] R. M. YERKES AND J. D. DODSON, *The relation of strength of stimulus to rapidity of habit-formation*, 18 (1908), pp. 459–482.

[209] N. YEUNG, M. M. BOTVINICK, AND J. D. COHEN, *The neural basis of error detection: Conflict monitoring and the error-related negativity.*, Psychological Review, 111 (2004), pp. 931–959.

[210] M. YOUSEF, F. IQBAL, AND M. HUSSAIN, *Drone forensics: A detailed analysis of emerging DJI models*, in 2020 11th International Conference on Information and Communication Systems (ICICS), IEEE, Apr. 2020.

[211] Y. YU, Z. ZHOU, E. YIN, J. JIANG, J. TANG, Y. LIU, AND D. HU, *Toward brain-actuated car applications: Self-paced control with a motor imagery-based brain-computer interface*, Computers in Biology and Medicine, 77 (2016), pp. 148–155.

[212] B. ZAAIMI, R. RUIZ-TORRES, S. A. SOLLA, AND L. E. MILLER, *Multi-electrode stimulation in somatosensory cortex increases probability of detection*, Journal of Neural Engineering, 10 (2013), p. 056013.

[213] J. R. ZADRA AND G. L. CLORE, *Emotion and perception: the role of affective information*, WIREs Cognitive Science, 2 (2011), pp. 676–685.

[214] ZEPHYRANYWHERE, *Bioharness log data descriptions*.
Available: https://www.zephyranywhere.com/media/download/bioharness-log-data-descriptions-07-apr-2016.pdf/. [Accessed: 28- Aug- 2020]., April 2016.

[215] H. Y. ZHU, H.-T. CHEN, AND C.-T. LIN, *The effects of a stressful physical environment during virtual reality height exposure*, in 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), 2021, pp. 468–469.

[216] H. Y. ZHU, H.-T. CHEN, AND C.-T. LIN, *The effects of virtual and physical elevation on physiological stress during virtual reality height exposure*, IEEE Transactions on Visualization and Computer Graphics, (2021), pp. 1–1.

[217] H. Y. ZHU, E. M. MAGSINO, S. M. HAMIM, C.-T. LIN, AND H.-T. CHEN, *A drone nearly hit me! a reflection on the human factors of drone collisions*, in Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, ACM, May 2021.