Designing and testing of pilot control interfaces for vertical take-off and landing (VTOL) aircraft

by

Azwad Abid

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2023

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The advancements in vertical take-off and landing (VTOL) aircraft have rapidly increased in the past few years, and there are working prototypes with human pilots already tested. With the current state of progress, provided the battery technology and automation level catch up to the required standard, VTOL cars could come to the market soon. However, the regulatory bodies are still working on the policies for automated cars and are far from their end goals. Given the scenario, it would be extremely beneficial to have empirical data to inform engineers, designers, and policymakers about what could be an intuitive controller from the existing hardware widely available in the market. This study investigates the ease of use of flying a VTOL aircraft between three of the most widely used controllers, namely the driving wheelset, drone radio controller, and joystick, backed up by performance data, EEG data, and the NASA-TLX survey.

A case study was conducted for 30 participants with a G2/G license aged 18 to 64 years. Each participant tried all three controllers in a randomized order to fly through a standard track in Virtual Reality (VR). Performance data and EEG signals were recorded in real-time, and a NASA-TLX survey was conducted after the user tried each controller. After they tried all three controllers, an overall survey was given to rank the controllers from the most preferred to the least preferred and to reason their choices. Finally, the users were asked to fly through the track with the driving wheelset one more time, where the randomized wind was introduced to see if that could affect their performance and overall workload.

The results of the experiment are compared among the three controllers using the three different types of datasets. The result shows that the joystick controller was the most preferred controller among the three controllers, backed up by the user survey, EEG data and performance data. If we compare just the statistical performance and not the surveys, the result is not significant enough to be reported. It is also found that weather conditions can significantly affect performance for the users.

For future work, the experiment should be carried out with varying weather conditions. Also, to improve among the existing controllers, the joystick could be modified to include pedals for rotation instead of the twist, reducing the chances of unintentional input from the user.

Acknowledgements

All praise be to God, who blessed me with good health and understanding people to make my master's project come to a fruitful completion.

I would like to especially thank my supervisor, Assoc. Prof. Dr. Shi Cao, for being such a great mentor and guide throughout my journey. He knows how to bring the best of people, and I am thankful to work with such a remarkable supervisor. My hats off to you and your method of guidance.

Secondly, I would like to acknowledge the assistance from Dr. John Muñoz for his help with the EEG data collection procedures and explanations. Also, special remarks for Nischal Ghookal for being a keen research assistant and helping with some software development procedures.

My heartfelt thanks to Orion Bruckman, Marco Moran-Ledesma and Rana Abdelrahman for their constant support and understanding regarding my Teaching Assistant job, which helped me balance between study and work. I always had a great time with their presence!

Last but not least, my humble acknowledgement to my friends and colleagues from in and out of the lab; Wachirawit Umpaipant (Pete), Muhammed Patel, Rana Abdelrahman, Wendy Ding, Yovela Murzello and Kareem Mostafa, for being there for me when I needed the support to carry on with my research. Thank you for keeping me at the peak amidst the sinusoidal trajectory of life.

Dedication

This project is dedicated to my eldest sister, Tania Sultana, who helped and supported me throughout my master's journey in Canada. It is a blessing to have a sister like you, and I am richer than the worldly riches for having such a beautiful guardian, well-wisher, and forward-thinker like you. May Allah (God) bless you in this world and the hereafter, Ameen.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Vertical Takeoff and Landing (VTOL) aircraft's advent has revolutionised how we conceive of aviation. These versatile aircraft combine the capabilities of helicopters and fixed-wing aircraft, allowing them to take off and land vertically while transitioning to horizontal flight, thus eliminating the need for traditional runways. The development of VTOL technology has opened up new possibilities for transportation, logistics, search and rescue, and even urban air mobility, making them a focal point of innovation in the aviation industry.

As VTOL aircraft continue to evolve and find diverse applications, the critical aspect of user controls cannot be overstated. These aircraft's safe and efficient operation relies heavily on the interface between human operators and the sophisticated technology that powers them. The user controls for VTOL aircraft play a pivotal role in ensuring the pilot's ability to navigate and control the aircraft under various conditions, including vertical takeoffs, landings, and transitions between hover and forward flight.

This master's thesis seeks to embark on a comprehensive exploration of the existing user control systems available for VTOL aircraft. By analyzing the current state of user controls and conducting a user study on the three major types of controllers in a virtual simulation, this project aims to figure out the strengths and weaknesses of each of these major controllers and help in the advancement of a robust user control system for the commercialization of VTOL aircraft.

Chapter 2

Literature Review

2.1 Background and Context

With the advent of highly funded startups like Opener [83], VoloConnect [88], and Jetson [40], the concept of flying cars has been maturing rapidly. Recent breakthroughs include flying a manned vertical take-off and landing (VTOL) aircraft for public demonstration [57] and passing extensive cold-weather testing at -23°C. Opener, a startup company, has achieved both of these milestones with their VTOL named BlackFly. Initially, they planned to sell 25 of these eVTOL aircraft in the Fall of 2021 [59], but later they pushed the date further back. In July of 2023, the company confirmed an early access program is coming soon [43]. From the economic perspective, the flying car market is expected to grow from $34.41 million in 2020 to $330.94 million in 2025 at a compound annual growth rate (CAGR) of 58.7% [7]. As a result, famous companies like Uber, Hyundai Motor Company, The Boeing Company and Japan Airlines (through Volocopter) have invested in the flying cars market [13].

While the policymakers are trying to regulate self-driving cars [92], the progress of VTOL aircraft has been going in full force [33], and the current advancements in the field have been astonishing. Yet, the user control interface requires more research. Many current VTOL aircraft use cyclic controllers found on helicopters or aircraft yoke systems as prevalent on commercial and private aeroplanes [91]. They work great for pilots with enough hours in the air. But when commercial VTOL aircraft become available for the public, what user control should be standardized for them? Yes, these flying cars should generally be automated, and it is possible to do so [66, 21, 78, 65]. The current advancements in machine learning technologies have made such automation quite viable [17, 51, 32]. But

what about emergency situations? What about manual override? It is impossible to expect everyone who wants to ride the automated flying cars to have a full-fledged pilot license for such cases. Plus, most civil air transportation rarely considers unmanned vehicles as a suitable choice as most passengers would refuse to fly on a pilotless plane [21]. Therefore, a toned-down, easy-to-handle user controller should be available to ensure anyone can take control of the aircraft in an emergency.

Nowadays, vehicles such as UAVs are becoming highly sought after in industries. This has led to the rise of VTOL aircraft convertiplanes (or flying cars) that appear to be a compromise between planes, multi-copters and cars [16]. As these vehicles become more popular, there should be people who are eligible to pilot them. One of the goals of this thesis is to find an existing off-the-shelf controller that can allow people to use their experience in driving, gaming or drone racing to pilot VTOL aircraft. This can reduce the amount of training time necessary to get a license. Nevertheless, the focus is to have a manual controller in the vehicles, which people are mostly comfortable with. In the past, there have been a few works relating to this field. From these papers, we find that three commonly preferred manual controllers are steering wheels, drone remote controllers or joysticks.

2.1.1 Drone Radio Controller

In quite a few simulations, remote controllers are preferred, as shown in the paper by Alaez et al. [16], a digital twin of a VTOL aircraft convertiplane UAV is tested and modelled using a Gazebo robotics simulator. They use the Ardupilot controller, which is interfaced with Gazebo using MAVLink [56]. The manual controller used in the simulator is a Taranis X9D remote controller. An image of the setup is shown below:

In the paper by Sinha et al.[80], a remote controller is also preferred as a manual controller. In this paper, the researchers designed a Quadshot air-frame, which is an improvement of previous VTOL aircraft topologies by [82], [24] and [50]. This airframe uses Bluetooth XBee transceivers to communicate with the ground station (laptop running the main program). The actuators of the airframe are controlled by a drone remote controller. An overview of the control process in this paper is shown in Figure 2.2.

Similar to the above two examples, Reiss et al. [77], make use of remote controllers in a VTOL aircraft pilot simulation. In this paper, simulations are done for a larger and more complex VTOL aircraft, where multiple crew members may be needed for flight. Regarding the VTOL aircraft analyzed in [80], one pilot's human-machine interface (HMI) is designed to enable intuitive flight control for inexperienced pilots as described in [30]. The analyzed vehicle consisted of eight propellers which were to be test-flown remotely.

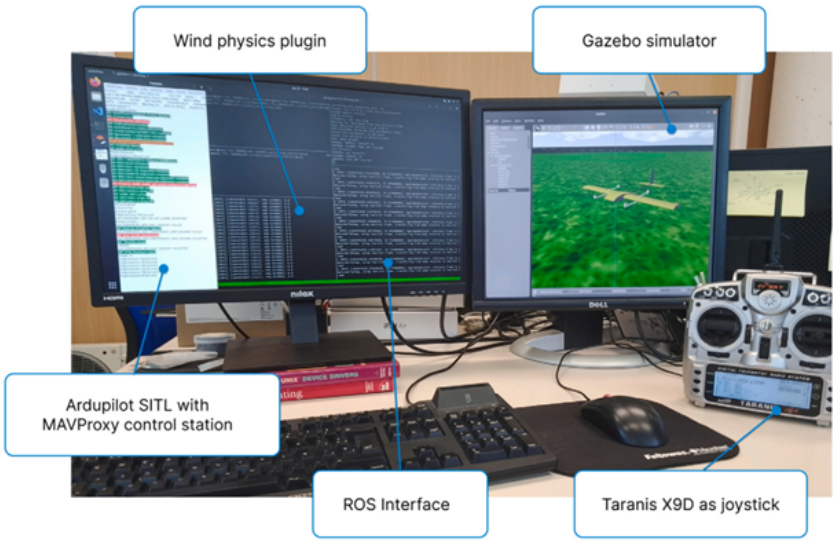

Figure 2.1: The simulation setup used by Alaez et al [16]

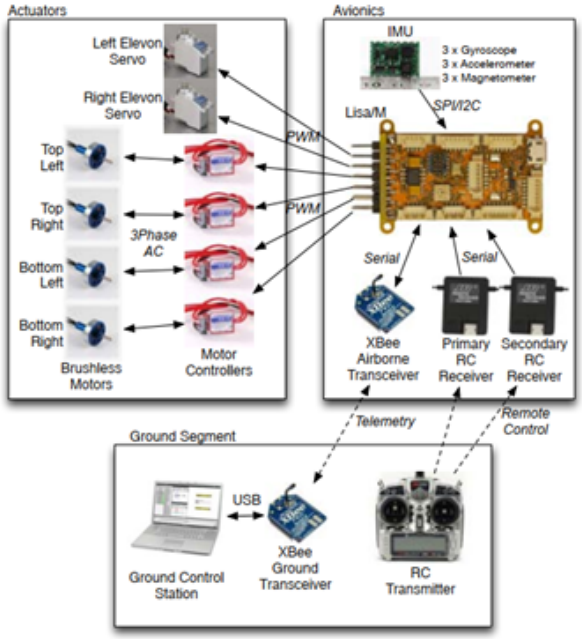

Figure 2.2: The control process employed by [80]

In the experimental setup, two pilots and a ground team monitor the VTOL aircraft's altitude. The main pilot of the VTOL aircraft is seated in a typical pilot station, where the control is done through hand-held joysticks. There is a backup pilot who makes use of the remote controller to manipulate the VTOL aircraft. In the paper, simulations are also done where both pilots use remote controllers, as shown in figure 2.3. The pilots can look at a display to see the current flight mode of the controller. Aural indications are introduced to increase the awareness of the pilots because both pilots' main task is to monitor the aircraft during the flight test. According to [77], the simulator was built up to be as close to the real flight tests as possible.

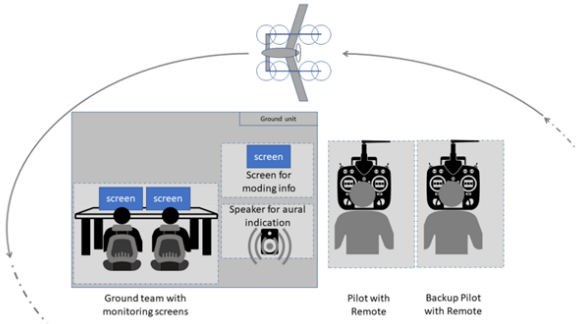

Figure 2.3: Schematic remote aircraft control for two pilots with equal remote control.

Besides the remote controller, steering wheel controllers are also favoured in some publications. For example, in the work of Gursky and Muller [45], analysis and simulation are done for personal aerial vehicles (PAVs), and in the simulations, the concept of a steering wheel with brakes (as in a car) is used to control the PAV.

2.1.2 Steering Wheel and pedals

It should be noted that the concept of steering wheels for VTOL aircraft is actually not new. As a matter of fact, it dates back to 1942. Many of these earlier concepts were related to the development of the helicopter. In the year 1942, Antoine Gazda, a Swiss aircraft manufacturer, employed engineer Harold Lemont to design a helicopter control system for him [4]. Due to Lemont's rather limited experience with helicopters, the draft he came up with was similar to the classic VS- 300 helicopter in many aspects. Still, there were some distinct differences. The control stick is of special interest here. It was mounted between the pilot's legs and worked like a conventional cyclic stick. In addition to that,

it had a steering wheel for yaw control on top and could be raised and lowered to control the main rotor's collective pitch angle. The vehicle designed by Lemont was named the Helicospeeder. The control concept of the vehicle is shown in Figure 2.4.

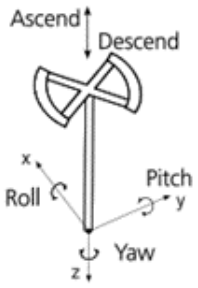

Figure 2.4: The helicopter control concept designed by Lemont [4]

In 1987, Jan M. Drees revisited the idea of small, low-cost helicopters which are easy to fly, safe and affordable [31]. While this idea had been in the minds of engineers since the 1950s, the introduction of Fly-by-Wire technology made it technically feasible to install innovative flight controls. He came up with a design sketch consisting of two devices similar to the steering wheel and brakes of a car. An image of the design is shown in Figure 2.5. What is remarkable about this design is the use of two thumb wheels, one to control lateral and the other to control vertical movement. Drees also suggested making inputs for acceleration and deceleration by using a slidable steering device, which would remind users of the steering wheel in a car.

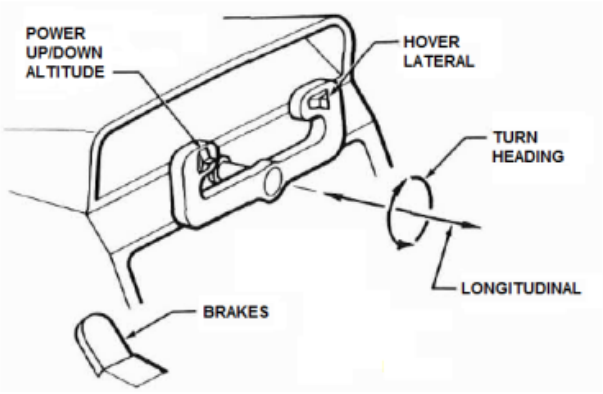

Figure 2.5: The car-like controller designed by Dress for an easy-to-fly helicopter [31]

For the next example, we can look into Flemisch's simulator study. Scientists of DLR

and the Technical Universities in Munich and Darmstadt used the "Horse-Metaphor" or in short, H-mode. It describes the idea of a vehicle acting autonomously like a well-trained horse [35]. Implemented in a car, the H-mode would be designed to control the vehicle using driver assistance functions. These functions would include highly advanced lane-keeping or obstacle avoidance. The driver is kept in the control loop with the help of active control elements that are configured for tactile cueing [35]. Following the metaphor, this behaviour is called "Loose Rein". "Tight Rein" means that the driver is given the majority of control over the vehicle, which can be initiated by the automation or the driver himself. In the opinion of the involved scientists, the H-Mode is not only limited to cars but can also be applied to any vehicle. A universal control concept was developed that could be used in both air and ground vehicles. Hence, training to control systems that differ in nature would reduce the training on one set and coordination between the two systems can be used to improve the driver's performance in both domains [35].

Under the direction of Frank Flemisch, simulation trials were conducted to determine if the H-Mode idea could be applied to such a universal control concept. The simulated vehicles were an automobile with the driving dynamics of DLR's FASCar prototype and an unmanned helicopter, both implemented with hardware-in-the-loop components and controlled from the same control station [35]. The control concept that is of interest here consisted of a steering wheel together with a sidestick. Its principle is shown in Figure 6. When in automobile mode, this stick commanded longitudinal movement and the wheel was used for steering tasks. In helicopter mode (in the simulator, the screen displayed the aircraft's ego perspective to simplify the task), the stick was additionally used for lateral movement control. A hat switch on top of it controlled the vertical movement. Control in the other two directions was the same as in automobile mode.

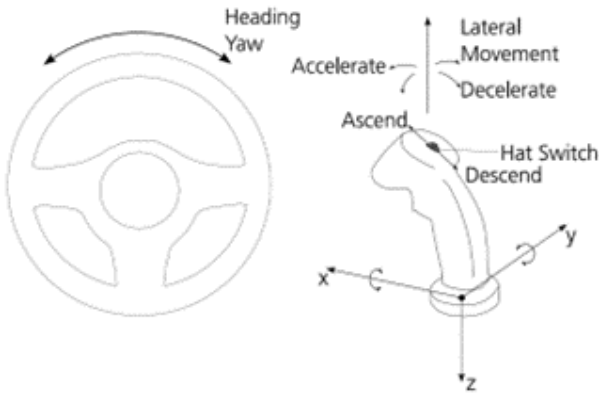

Figure 2.6: Control concept of Flemish's simulator study [35]

In the paper by Gursky and Muller [45], the steering wheel selected to control the PAV has only one primary axis. This is in contrast to the works of [4] and [31]. Based on the research by Landis [14], increasing the number of axes on one device can increase the likelihood of unintentional coupling between inputs in different axes. Nevertheless, Gursky and Muller [45] combined ideas from the works of [4], [31] and [35] to build their own control system for the VTOL aircraft simulator as shown in Figure 2.7. The steering wheel solely controls both roll and yaw motions. The transition between them depends on the forward airspeed. A collective lever is used to control vertical movements. The response due to the steering control in this paper is shown in Figure 2.8.

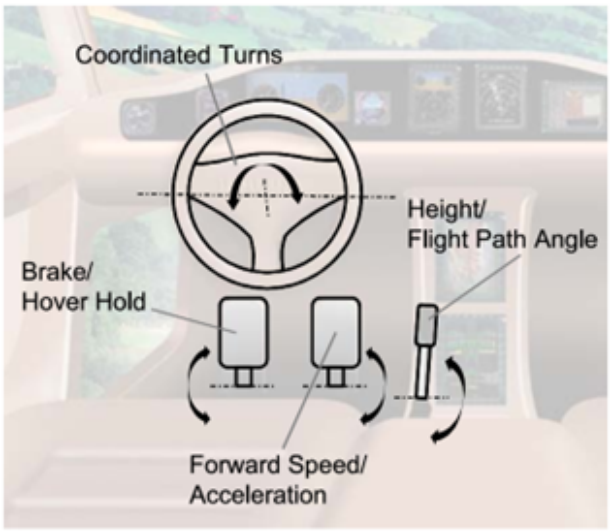

Figure 2.7: Primary control functions for PAV control as shown in [45]

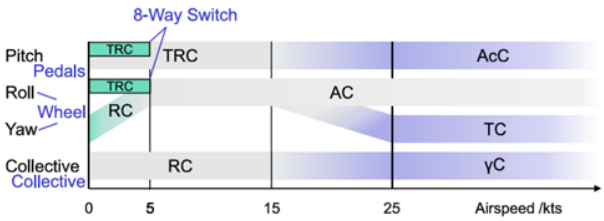

Figure 2.8: Response type modifications for control concept with the steering wheel. The steering wheel controls both the roll and yaw axis.

2.1.3 Joystick Controller

Joystick controller (also known as cyclic controller) is quite renowned in the VTOL industry for its intuitive design, ease of use, and familiarity. Most of the joysticks use the concept of hands-on throttle-and-stick (HOTAS). The idea behind HOTAS is to mount switches and buttons on the flight control stick and throttle lever in the cockpit of an aircraft. Looking at three of the advanced personal VTOL aircraft, all three are using the joystick as their controllers.

1. BlackFly by Opener: The aircraft, known as BlackFly, possesses unique characteristics that distinguish it from every other aircraft currently in existence. The electric vertical take-off and landing (eVTOL) aircraft, developed by Opener, a startup based in Palo Alto, California, features short wings positioned in front and behind the pilot, equipped with four motors and propellers on each wing. From a visual standpoint, it may be observed that the image resembles an airborne individual with remarkable velocity, reminiscent of a character from a science fiction narrative of the 1930s pulp era [94]. The pilot controls consist of a joystick equipped with a thumb control specifically designed for adjusting altitude. The flight controls of the aircraft consist of a triple-redundant fly-by-wire system that is responsible for operating the motors and dual elevons located on the outer edge of both wings. Control authority in pitch, roll, and yaw is achieved through the utilization of differential motor speeds [11].

2. The Jetson is a lightweight, fun-to-fly eVTOL aircraft with eight powerful motors and an intuitive flight computer. It features a race-car-inspired safety cell, auto land function, and multiple safety features to ensure the pilot's safety and maintain continuous flight even with one motor loss [10]. The operation of this vehicle also involves the utilization of a joystick (three-axis) and a throttle lever [41]. The cockpit view of Jetson One is shown on Figure 2.11.

3. The Volocopter 2X is a German-made electric vertical takeoff and landing (eVTOL) aircraft with a seating capacity of two individuals. It offers the option for either manned or autonomous operation and utilizes a multirotor configuration. The personal air vehicle was developed and manufactured by Volocopter GmbH, a company based in Bruchsal. It made its initial debut at the AERO Friedrichshafen airshow in 2017. The aeroplane is sold in a fully assembled and flight-ready condition. The previous designation of Volocopter was E-volo [84, 6]. The controls are operated using a network of mesh polymer fibre optics, which utilizes a fly-by-light technology. The primary flight control unit consists of triple redundancy, complemented by a dissimilar backup flight control unit and a joystick control [12].

Figure 2.9: On October 1, 2022, a BlackFly electric vertical takeoff and landing (eVTOL) aircraft successfully initiated its flight at the Pacific Airshow held in Huntington Beach, California [93].

Figure 2.10: The pilot inside the BlackFly is using the joysticks to pilot the aircraft [23].

Figure 2.11: The cockpit view of Jetson One [8].

Figure 2.12: The cockpit view of the Volocopter 2X [5].

Figure 2.13: Volocopter 2X during historic milestone flight at Gimpo airport, Republic of Korea [9].

2.1.4 EEG Signals and Processing

The field of human factors and ergonomics examines the dynamics between individuals, machinery, the surrounding environment, and technology, taking into account human capacities and constraints in order to establish secure and gratifying work settings [26, 48, 53, 54]. The evaluation of job activities in traditional approaches was characterized by subjectivity, employing qualitative methodologies [81, 60, 79]. These methodologies fail to provide sufficient means for a comprehensive examination of the intricate dynamics involved in the cognitive, perceptual, and physical dimensions of engaging with contemporary technology [53, 55, 69, 76, 46]. Furthermore, they do not facilitate the modelling and quantification of the intricate interplay between the human mind and technology [46].

The quantity of empirical investigations centred on neuroergonomics has experienced a significant rise in conjunction with the advent of neuroimaging methodologies [20]. These methodologies rely on the measurement of brain activity as opposed to alterations in cerebral blood flow or voltage variations caused by ionic current [87, 62, 39]. The electroencephalogram (EEG) possesses both merits and drawbacks when compared to alternative neuroimaging techniques, hence presenting both use and complexity in the context of neuroergonomics applications. The primary benefits encompass (1) a notable level of temporal resolution [89], (2) the ability to be easily transported for utilization in real-world settings, and (3) cost-effectiveness [18]. Nevertheless, EEG approaches are associated with three notable limitations: firstly, they suffer from low spatial resolution [86]; secondly, they are susceptible to the presence of unwanted nonbrain signals or artefacts [15, 85]; and thirdly, they necessitate a lengthy preparatory time for setup [39]. Notwithstanding these obstacles, recent progress in electroencephalography (EEG) technology has resulted in the creation of wireless EEG systems, which enable individuals to carry out continuous tasks without any disruption [64, 58]. Additionally, these systems employ dry electrodes instead of wet systems, thereby reducing the time required for preparation [25, 44, 71]. In addition, researchers have created software for automatic artefact detection [28] with the aim of enhancing the quality of signals.

The EEG-engagement index (beta/(alpha+theta), as introduced by Prinzel et al. [73], is characterized as "the proportion of beta power to the combined theta and alpha power observed in specific EEG measurement channels" [90]. This index has the potential to identify declines in task engagement [38]. The decline in the engagement index seen during vigilance tasks indicates a progressive decline in task involvement as time elapses [22, 52, 72]. A strong association has been seen between the EEG engagement index and task load, as indicated by previous research [22]. This finding highlights the efficacy of the index in accurately measuring the level of effort.

The utilization of the engagement index (beta/(alpha+theta) is strongly advocated for the purpose of constructing adaptive systems [19]. The primary objective of an adaptive system is to optimize cognitive engagement and situational awareness by effectively managing the cognitive workload within a modest range. When the engagement index value exhibits an increase, the system undergoes a transition into automated mode, subsequently leading to a reduction in the EEG-engagement index. On the other hand, in cases where the engagement index value declines, the task is transitioned to manual mode, resulting in an augmented workload and consequently leading to an increase in the EEG index [74, 73, 38, 19, 63, 37, 73].

2.1.5 NASA Task Load Index

The most extensively used and validated metric to gauge overall workload after completing a task is the National Aeronautics and Space Administration Task Load Index (NASA-TLX) [47, 42, 29]. It was initially developed for the aviation sector by the Human Performance Research Group at NASA Ames Research Centre. Since then, numerous additional domains, including computer science [70], psychophysiology [36], and transportation [75], have adopted its application. The NASA-TLX is a multidimensional tool that has six dimensions which are predefined. Three dimensions measure how the subject deals with the task at hand (self-rated performance, effort, and frustration level), and three dimensions focus on the demands imposed on the subject (mental, physical, and temporal demands). The six subcategories that make up the workload are intended to limit subject-to-subject variation and identify the workload's origin. There are two distinct approaches for employing the NASA-TLX tool. The weighted NASA-TLX score follows a two-step procedure, wherein the user initially assesses all six subcategories upon finishing a particular task and subsequently assigns predetermined importance to each factor. The objective is to gain a deeper insight into which possible factor primarily influences the perceived workload. Conversely, in the case of the raw NASA-TLX score, the user evaluates all six subcategories after completing a specific task, without assigning any weights. The final outcome is calculated as the average of all the subcategories. Studies have indicated that the raw NASA-TLX demonstrates a strong correlation with the weighted version [42] while being a more time-saving and straightforward assessment method [49, 67]. So, for this user study, raw NASA-TLX has been used.

2.2 Relevance and Importance

This research project aims to provide empirical data supporting the evaluation of three pilot control interfaces for VTOL aircraft. This data will be useful for engineers, designers, and policymakers to standardize user controllers for commercial flying cars. Furthermore, this project can encourage sustainable flying car designs as both the industry and the government can use the empirical data as a point of reference, reducing the need for redesigns in the industry after new regulations have been set.

2.3 Research Question

To summarize, the research question for this project is "What would be the best manual controller/interface for future flying cars (vertical take-off and landing aircraft) among current major design alternatives?"

2.4 Research Hypothesis

There are four major expectations from this study. They are:

1. The Driving wheelset is expected to perform better than the other controllers as people with a car driver's license are used to it.
2. RC controller is expected to perform better than cyclic controllers but worse than driving wheelset.
3. Cyclic controller is expected to perform the worst as people are not used to it.
4. People who have used a given controller before should perform better than people who have never tried that specific controller before coming to the experiment.

2.5 Objectives

1. Record previous experiences with the controllers using a survey.
2. Capture relevant driving performance data from the system on each user.
3. Compare the results to determine which controller would be suitable for industry standards based on empirical data. Prior experiences should be accounted for.

Chapter 3

Research Methodology

3.1 Experimental Design

In this experimental design, the manual controller to drive the VTOL aircraft is the independent variable, which will be manipulated to assess their impact on various dependent variables. Three controllers are considered for this user study; (1) Racing Wheel, (2) Drone Radio Controller, and (3) Joystick. The dependent variables encompass a comprehensive set of metrics, including NASA-TLX scores, Root Mean Square Error (RMSE) in task performance, time for completion of the assigned tasks, the number of wrong entries made during the tasks, the time spent outside the designated track or path, and the analysis of EEG data across five distinct frequency bands. This study adopts a within-subject experimental approach, allowing each participant to experience all three controllers to better capture individual differences and provide a more robust assessment of their performance and influence on the specified dependent variables.

3.1.1 Objective of the users

The user's objective is to fly the virtual VTOL aircraft from the start point to the end point while passing through checkpoints A to H. The users' primary goal is to ensure that they can pass through each checkpoint while staying as centred as possible. To help with that, each checkpoint has its own centre point, as shown in Figure 3.4. One concern regarding centring is knowing where the car's centre is. For the experiment, a reference ball has been put at the front of the car's dashboard, within the driver's peripheral vision. The driver

simply has to ensure that the reference ball of the car goes through the centre ball of each checkpoint.

There are three different phases of driving during the experiment. The user will initially start by taking off vertically upwards and keep on moving up until they reach an expected altitude. Once they reach that altitude, they will drive the aircraft just like a normal car on the road. After they reach their desired destination, the user drives vertically down to land the VTOL aircraft.

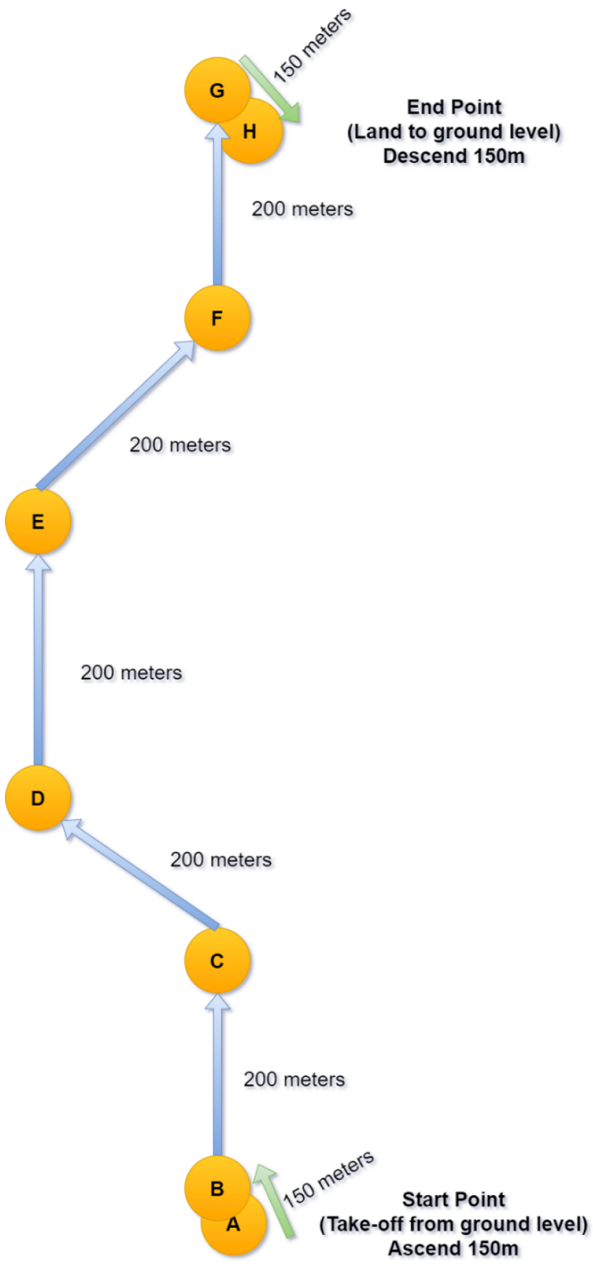

Figure 3.1: Experiment track route for all users

3.1.2 Track Designs (tutorial and actual)

There are two separate tracks available for this study. Before the participants go for the experiment, each controller is explained to them, and the tutorial track is launched for them to try to play around with the controller and get the hang of the manoeuvres. During this trial, the users do not wear the virtual reality headset but instead use the desktop monitor to see their movement from the inside cockpit of the VTOL aircraft. This had to be done to ensure that each participant is used to the expected movement of the vehicle based on the input of the controllers, ensuring they do not get startled, which can cause motion sickness.

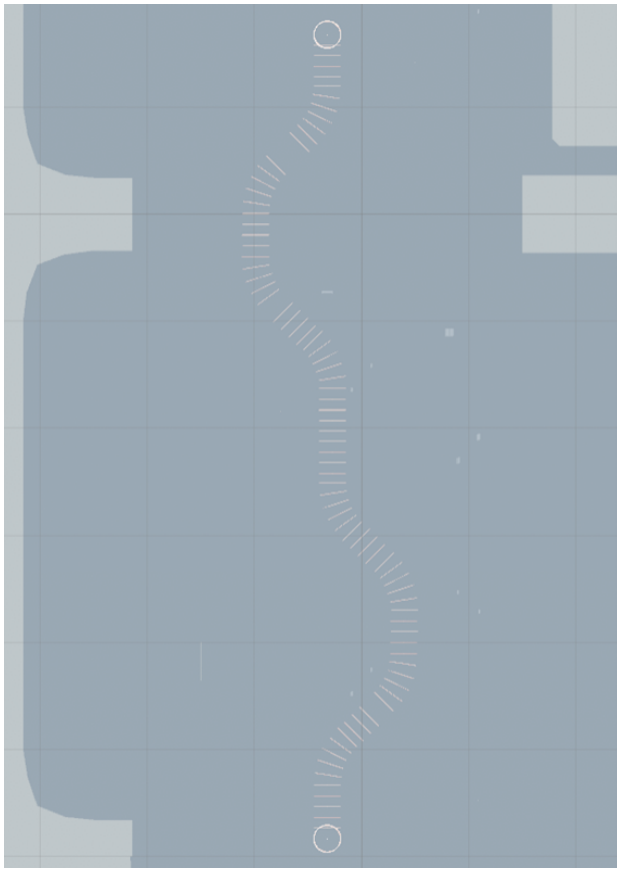

Figure 3.2: Top view of the overall track for the tutorial.

The tutorial track is shorter than the actual experiment track but has many left and right turns, letting the users get acquainted with the controllers. Each user completed the

tutorial track for each controller in the serial serial of their particular randomized order of controllers. For example, if a user's particular order is steering wheelset, drone radio controller and joystick, they tried the steering wheel first, drone radio controller second, and the joystick as the third controller inside the tutorial run.

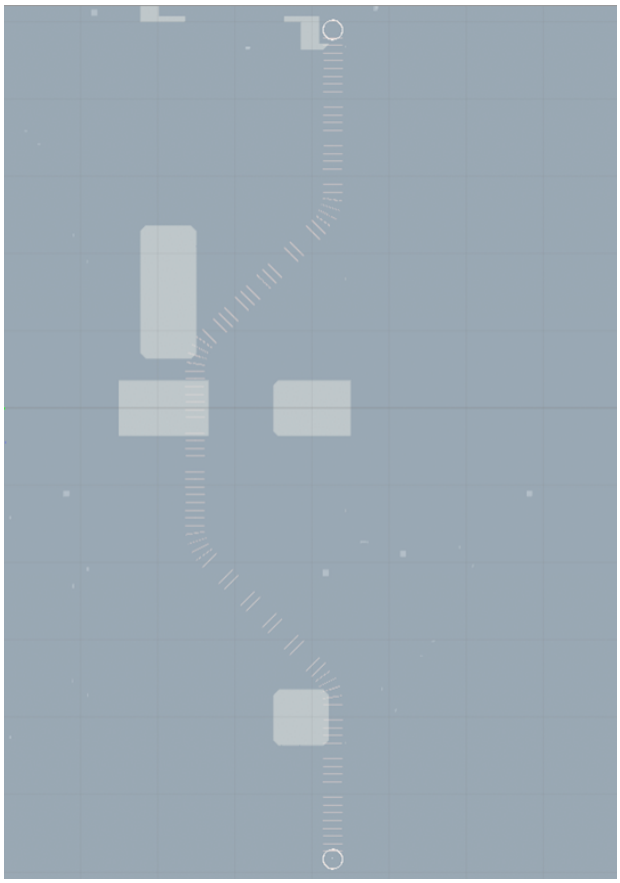

Figure 3.3: Top view of the overall track for the experiment.

Once the tutorial is completed, the actual experiment portion of the study begins, where each user goes through the same experiment track, trying all three controllers in randomized order. The experiment track is much more linear and realistic than the tutorial track but much longer in length. The user must take two left turns and two right turns, all at a 45-degree angle to the car's centre. The rest of their job is to ensure they are at the centre of the track and drive straight through the checkpoints. The track also includes vertical take-off and landing at the start and end points.

One important question is how to ensure that the user will go through each checkpoint while maintaining as centre as possible to each checkpoint. To do that, first of all, the checkpoints are given a centre reference bubble which helps to address where the centre point is located on each checkpoint. Regarding the VTOL aircraft, the user usually sits at the left of the car, which is not the centre of the car. If the exact centre of gravity of the car is used, it is very hard for the user to tell whether they're actually referencing the car's centre as they're going through each checkpoint. To solve this problem, a reference ball has been placed at the front centre of the car right above the dashboard, clearly visible within the peripheral vision of the participant. So, the user simply has to drive the car's reference ball through the centre of reference of each checkpoint. Once the user has driven through all 151 checkpoints, the student investigator ends the data collection and tells them they can take off their VR headset.

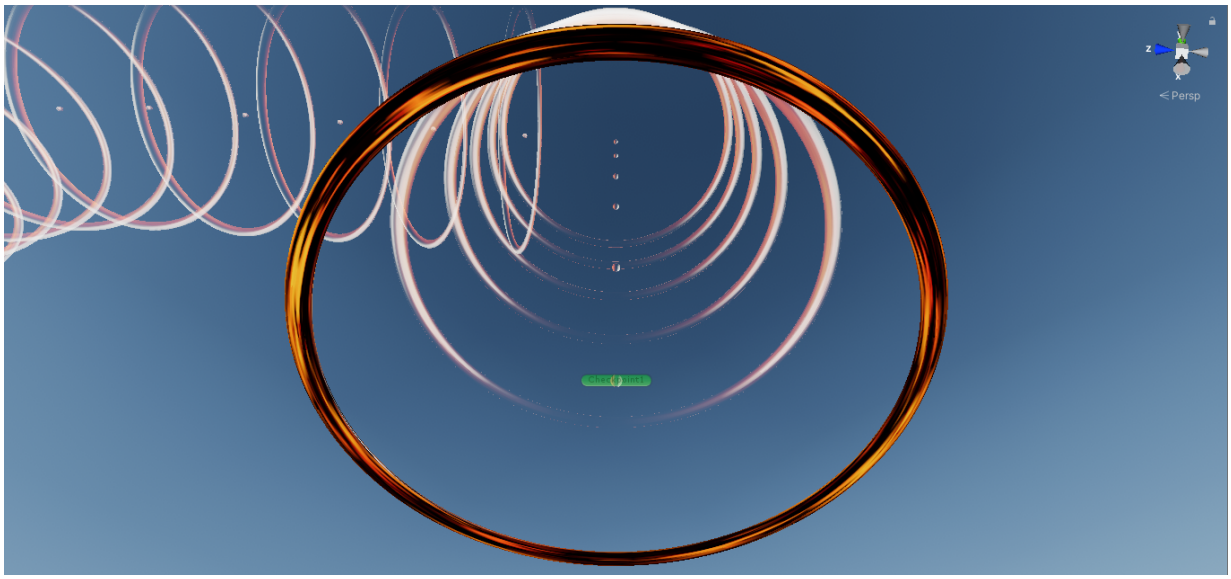

Figure 3.4: Bottom-up perspective view of the checkpoints from the start point. The active checkpoint can be seen with a darker colour. Each checkpoint is accompanied by a centre reference point.

Figure 3.5: Perspective view of the dashboard for the users in Virtual Reality. The white ball on the right represents the reference point of the car.

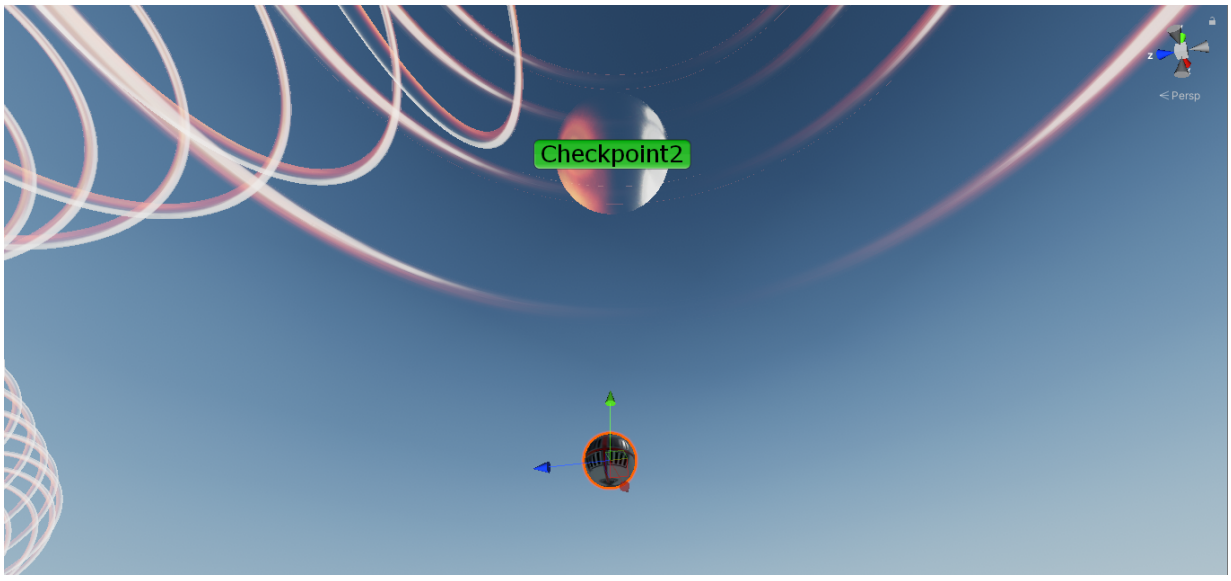

Figure 3.6: The white reference ball inside the car has to go through the reference bubble of each checkpoint for maximum accuracy.

3.1.3 The Four Conditions

Before we started collecting data for the experiment, each user was allowed to try each controller in the randomized order they were supposed to continue the experiment. During this trial run, each user was explained how each controller works then they were left to try the controllers without using VR by looking at the computer monitor. Data was not collected during the tutorial session, and the users were free to try the controllers without any restrictions. Once the user tried all three controllers and completed the tutorial track thrice (once for each controller), the users were allowed to begin the experiment.

Serial	UserID	Order1	Order2	Order3
1	9001	1	2	3
2	9002	3	1	2
3	9003	2	3	1
4	9004	1	2	3
5	9005	3	1	2
6	9006	2	3	1
7	9007	1	2	3
8	9008	3	1	2
9	9009	2	3	1
10	9010	1	2	3
11	9011	3	1	2
12	9012	2	3	1
13	9013	1	2	3
14	9014	3	1	2
15	9015	2	3	1
16	9016	1	2	3
17	9017	3	1	2
18	9018	2	3	1
19	9019	1	2	3
20	9020	3	1	2
21	9021	2	3	1
22	9022	1	2	3
23	9023	3	1	2
24	9024	2	3	1
25	9025	1	2	3
26	9026	3	1	2
27	9027	2	3	1
28	9028	1	2	3
29	9029	3	1	2
30	9030	2	3	1

Table 3.1: A balanced, randomized order for all the participants partaking in the experiment, where “1” stands for Racing Wheel, “2” stands for Drone Radio Controller, and “3” stands for Joystick.

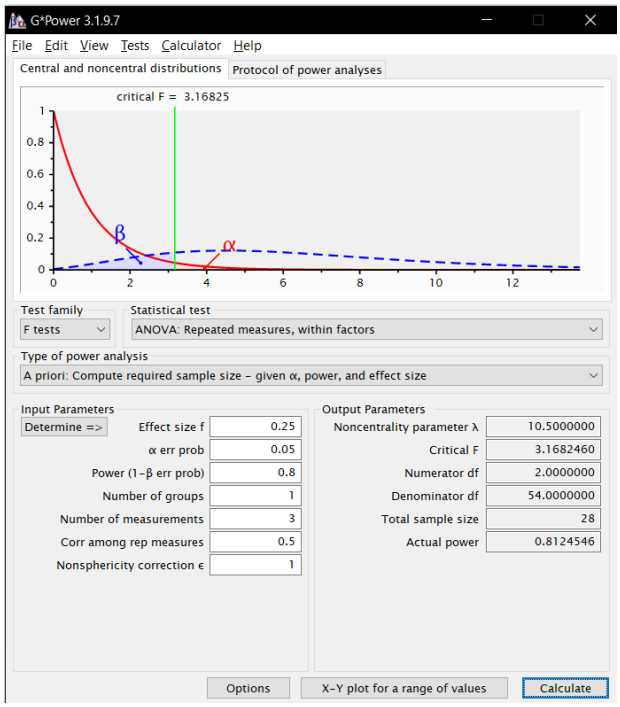

Figure 3.7: Power analysis calculation using the G-power software. Based on the estimated effect size of f = 0.25, the total number of samples required is 28, providing a power of 0.8.

As mentioned in the previous chapters, three different controllers are being tested. The order of the controllers is randomized to ensure that the data is not biased and there is a balance between all the data that are being collected. Even though the order of the controllers tried has been randomized, the controllers' serial remains the same. The normal order was the racing wheel first, the drone radio controller second, and the joystick as the third controller. Some users started with the racing will first and moved on to the drone controller and the joystick. Some users started with the drone controller first, moved onto the joystick, and ended with the racing wheel. The third group of participants started with the joystick first, the racing afterwards, and ended with the drone radio controller. So there are three possible combinations of the randomized order that have been used in this particular experiment. To balance this experiment, the overall number of participants should stay at a multiple of 3. Based on our power calculation in Figure 3.7, we found the number of participants needed for this experiment to have a P value of less than 0.05; we need a minimum of 28 participants. We collected data for 30 participants to balance out the overall randomized order of the given controllers.

After the user had tried all three controllers, a fourth condition was tested on them. This 4th condition included driving the same racing wheel that they initially tried as one of the first three conditions, except this time, there was randomized wind to restrict the movement of the users and see how much of a mental workload difference it made while having an external weather condition for the users.

The randomized wind was produced using Unity physics engine and a building over a technique used by Tate et al. [REF]. Every five seconds, the sideways direction and force of the external force on the car is randomized.

$$WindForce = (int)Mathf.Sign(UnityEngine.Random.Range(-100, 100)) \tag{3.1}$$

3.1.4 Types of Data Collected

Overview of all the different types of data collected for this study. For this study, three separate methods of measurement were used. Each of them is expected to complement each other or at least help to come to a reliable conclusion. The three types of data collected are:

1. Questionnaire: Survey data collected straight from the users to:
    - Understand their demographic information.
    - Understand their previous flying and controller experiences that might affect the performance of the study.
    - Self-assess their performance after each condition of the experiment.
    - Inform which manual controller they preferred to use and why.
2. Performance data from the simulator: Data logged straight from the unity system as the user passed through each checkpoint and finally completed the entire level. This performance data is quantitative and can help with empirical analysis.
3. Performance data from the EEG system: The EEG system recorded the brain signals in five of the most common frequency bands, namely delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (14-30 Hz), and gamma (¿ 30 Hz). The results can be compared by band and by engagement index (beta/ (alpha+theta)).

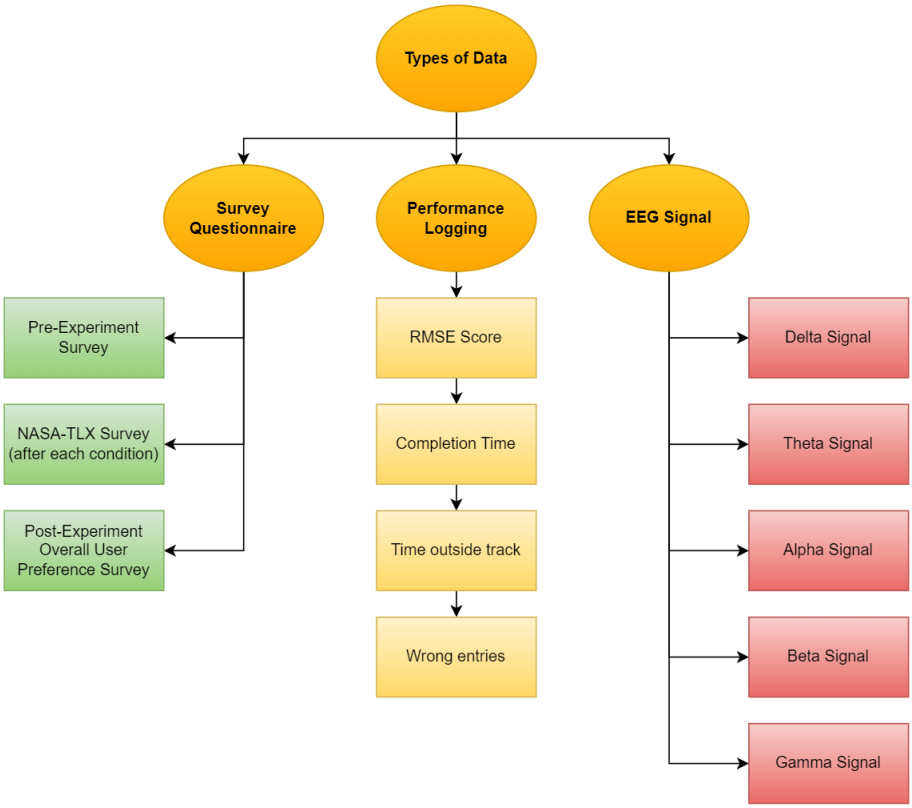

Figure 3.8: A diagram demonstrating the types of data collected for this user study

3.2 Hardware

3.2.1 Controllers

The objective of the research is to figure out what would be the best possible manual controller for people with no flying experience to start driving a VTOL aircraft. To do that, three separate controllers that are commonly available to the general population are considered. They are mentioned below, along with the picture of each of the controllers.

1. Logitech G27, Force Feedback Wheel and Pedal Set.
2. Radiomaster TX12, 2.4g Remote Control System.
3. Logitech Extreme 3D Pro, Precision Flightstick.

Figure 3.9: Logitech G27 Racing Wheel [2]

Figure 3.10: Radiomaster TX12 Drone RC [3]

Figure 3.11: Logitech Extreme 3D Pro Joystick [1]

This is a within-subject study; each of the users will try all of the three controllers that are shown in Figure 3.9, 3.10 and 3.11. The order in which they will try the controllers is randomized to ensure no biases in performance due to the order.

3.2.2 Tools and simulators

Experimenting on vehicles such as VTOL aircraft can be quite tedious if it is to be done on an actual aircraft. Instead of doing that, simulators can play an extremely safe and cost-effective way of experimenting. For this particular study, the experiment requires VTOL aircraft within the simulator. However, as VTOL is a relatively new field with not many customers in the market yet, there has not been any simulator to test out VTOL aircraft. To solve the problem, a new environment has been created inside the Unity game engine.

The physics of the simulated VTOL aircraft has been referenced with Microsoft AirSim's quadcopter physics. Basically, a VTOL aircraft can be considered as a bigger quadcopter with a bigger load. Therefore, the reference used should be reasonable.

Emersion is another important factor that must be considered while designing a system. That has been ensured using Oculus Quest 2, a Virtual Reality (VR) headset made by Meta (please refer to figure 3.2). The VR headset simulates an environment as if the user is actually sitting in a VTOL aircraft and is controlling the system real time. As far as the

Figure 3.12: Oculus Quest 2, the virtual reality headset from Meta.

controller is controller, the physical hardware will be available in front of them, and the user will use it to navigate the vehicle.

From the hardware side, the Muse BCI 2 EEG system will be used to record the brain functions of the users during the experiment. This will help to understand the mental workload of the users, which should complement the other performance metrics.

3.3 Software

This study had a very high usage of software. A virtual environment was created in unity to create the simulation where the users can try out the VTOL aircraft without any physical harm or risk. For this particular project, Unity version 2019.3.12f1 was used. It was initially selected to allow the use of the Microsoft AirSim library, which was taken out at the last phase of the software development Due to lack of freedom of customization. Inside the Unity, two levels were made, one for the tutorial and one for the actual experiment. The XR interaction tool (version 0.10.0) and XR plugin management (3.2.15) were used for the VR portion of the development. Fantastic city generator [61] was extremely useful for the large city creation (1500 meters by 1200 meters). The VTOL aircraft 3D model present in this project was created by Art Equilibrium [34].

Calendly was used for automatic scheduling and reminding systems for the participating users and the host. One important but tedious point when collecting user data is the scheduling. To make that process much smoother, Calendly premium was used. When people scan the QR code from the invitation poster (please refer to Figure B.1) posted all around the campus, it takes them to a registration page made in Qualtrics, where a

Figure 3.13: Muse BCI 2, an EEG system

Calendly link is given to book a particular slot. This simplifies the process because the system is synched with the student investigator's timetable. The user can choose any empty slot available on the Calendly link and register with their phone number. Once the registration is completed inside Calendly, an automated e-mail is sent to them with the option to save the event in their calendar as well. Not only that, but the system also reminds the user two hours before the time when their experiment is supposed to start through an SMS. This text message reminds the user that the user study will begin in two hours at the exact time and location, along with a link with video instructions on how to go to the particular location. After completing the experiment, Calendly sends another preset e-mail giving a thank you note to the participant. This software made the entire automated process extremely smooth and efficient, saving the participants and the student investigator a lot of time and effort.

The poster for the invitation to the user study was made using Canva Pro. The full view of the poster can be found in Figure B.1. It was distributed on the campus and on social media.

For all kinds of surveys, Qualtrics was used. This software is considered much more secure than Google Forms. The survey data that has been collected Are in a secure server within the University of Waterloo. Qualtrics also helped process the data collected from

the service using the automated system that had to be set up once and once only. The suggestions Qualtrics gave while making the forms were also quite helpful.

To collect data from the EEG device, the mind monitor app available on the Apple store provided an excellent service. The data could be recorded, stored directly on the server, and accessed immediately from the computer. The app also lets the student investigator observe the band signals in real time, allowing them to ensure the collected data is clean and unusable.

Once all the data had been collected for all the participants, the data was consolidated into a folder inside the lab computer and backed up to the secured OneDrive server of the university. Access to this OneDrive server was given only to the principal investigator and the co-investigator. The consolidated data was analyzed in Qualtrics, Microsoft Excel and SPSS.

3.4 Experimental Procedures

3.4.1 Description of Procedures

In the following table diagrams, the flow of the user study on the day of the experiment is explained in detail.

Step	Task
1	Clean the lens glasses with wipes.
2	Clean the body of the VR headset and controllers with anti-bacterial wipes.
3	Clean the silicon cover of the headset with anti-bacterial wipes.
4	Clean the Drone RC controller with wipes.
5	Clean the Joystick with wipes.
6	Clean the steering wheel with wipes.
7	Clean the Muse BCI 2 with wipes.
8	Take all the equipment out of charging and lay them down on the table.

Table 3.2: Steps to initialize and sanitize the equipment

Step	Task
1	Ask the user to read through the information letter.
2	Once the participant has read through the information letter, ask them to sign the sign the consent form before proceeding.
3	Inform the participants that they fill up the raffle draw form if they would like to be considered for the raffle draw of five $50 gift amazon gift cards.

Table 3.3: Steps to complete the legal documents before proceeding.

Step	Task
1	Let the participant sit on a comfortable chair and get the headset and the controllers (one controller will be practiced at a time but all will be prepared and connected beforehand).
2	Power on the Muse BCI 2 EEG system and set it on the forehead of the user.
3	Collect data for 3 minutes using the Muse BCI 2 to formulate the baseline of brain activity for the user.
4	Power on the VR system with the power button on the headset.
5	Adjust the position of the VR headset until the image is clear (e.g., not blurry).
6	For the joystick (cyclic controller), ensure that the subject is using their dominant hand.
7	For the steering wheel setup, ensure that the rig is set at a distance where the user can easily reach the pedals and the steering wheel while not being too cramped at the same time.
8	For the drone RC, the user should be able to rest their elbows on the handrests of the chair while putting the controller on the table comfortably.

Table 3.4: Steps to set up the environment and the equipment

Step	Task
1	Explain to the user how that controller works; point and describe which button in the controller does what.
2	Once the user is clear about the controls, the training scene is started on unity. The training track is shorter than the actual experimental track, but it has a lot more turns to let the user get acquainted to the controls and the feels of the controller.
3	During the tutorial session, the users will not be wearing the VR headset. Based on our preliminary tests on test users, it was found that putting the users in VR while they are not familiar with the controls tricks their brain and they feel motion sickness. That is why, during the tutorial session, the users will look through the computer monitor to drive.
4	The participant starts flying through the tutorial scene and drive from the start point till the end point. This includes taking-off vertically for 150 meters, then driving horizontally through all the turns, and landing at the end point.
5	Throughout the tutorial session, the student investigator ensures that the user understands how the controller works, and he gives verbal inputs to correct or remind the user, in case they forgot any of the important information regarding the controller.
6	Once the user has gone through all the checkpoints and completed the session, the entire process is repeated again for the second and third controller.

Table 3.5: Steps for the tutorial session

Step	Task
1	Explain the VR system to the user and let them wear the VR headset.
2	Start the actual experiment scene on unity.
3	Ask the user to sit comfortably and hold the controllers in ready position.
4	Reset the view inside VR to ensure that their current resting position is the desired sitting position inside the virtual VTOL aircraft.
5	Ensure that the logging option is turned on; unity should be logging the data automatically.
6	Start the EEG data recording and verbally tell the participants that they can start driving now.
7	The user will drive through all the checkpoints using one of the controllers in the randomized order. For each of the three controllers, the user will:
	a) Drive the VTOL aircraft from the start point to the endpoint, following the checkpoints one by one, trying to stay as close to the center of each checkpoint as possible. Accuracy is the objective of the user. b) Remove their VR headset. c) Take a NASA-TLX survey to express their experience and self-assess their performance.
8	After all of the three controllers have been tried and the surveys have been completed, a final survey will be given where the user has to rank the controllers based on their preference; "1" means the "most preferred", and "3" means the "least preferred". They also have to justify their ranking based on their experience by explaining their choices in the explanation box.
9	Next, the user will try the racing wheel controller one more time in the same track. But, this time, the track will have chaotic wind flowing sideways. The end goal for the user is still the same – race through the track as accurately as possible. The user will be wearing VR headset for this condition as well.
10	Take off VR headset and answer one last NASA-TLX survey to self-assess their performance.

Table 3.6: Steps for the main experiment

Step	Task
1	Remove VR headset first, and then the Muse BCI 2.
2	Power off all the systems.
3	The participants are thanked for joining this user study.
4	They are also informed when the raffle draw for remuneration and seminar will happen.
5	Participant leaves the room.
6	If there is another user session right after, the student investigator prepares for it. Otherwise, the Muse BCI EEG headset, ipad and the Oculus Quest 2 headset are put to charge for later use.

Table 3.7: Steps to wrap up the session

Step	Action	Time (HH:MM)
1	Arrive at the lab.	00:00
2	Confirm eligibility, read the information letter and sign the consent form.	00:00 – 00:05
3	Fill up the raffle draw entry details. (Optional)	00:05 – 00:06
4	Complete the pre-experiment survey (Demographic information)	00:06 – 00:10
5	Record 3 minutes of baseline EEG brain signal data.	00:10 – 00:15
6	Get acquainted with the three types of controllers and try the tutorial for each controller.	00:15 – 00:40
7	Use the controller to drive the VTOL aircraft through the checkpoints.	00:40 – 00:50
8	Complete the NASA-TLX survey to self-evaluate your performance.	00:50 – 00:53
9	Repeat steps 7 and 8 for the 2nd and 3rd controllers.	00:53 – 01:20
10	Complete the user preference survey to compare and evaluate the three controllers.	01:20 – 01:25
11	Try driving the VTOL aircraft through random wind resistance on the track.	01:25 – 01:35
12	Complete the NASA-TLX survey to self-evaluate your performance.	01:35 – 01:38
13	The participants are thanked and explained how they will be informed about the raffle draw.	01:38 – 01:40

Table 3.8: Overall flow of the session and their approximate timings based on all 30 users

3.4.2 Risks

- Fatigue or exhaustion: Participants may feel tired after playing the game because they have to control a vehicle with new controls and movements that they never tried before. In case of fatigue or exhaustion, participants will have the option to stop the interaction and remove the headset. Participants will also be asked to perform a series of relaxing exercises recommended by the research team. Additional resources and steps are addressed in the Safety Procedures section in case the symptom persists for a longer time.

- Emotional stress: Participants may experience short-term and low-severity emotional stress due to the game's complexity. If participants manifest a prolonged, constant distress symptom during gameplay, the participants will be suggested to stop, postpone or withdraw from/the study. Additional resources and steps are addressed in the Safety Procedures section in case the symptom persists for a longer time.

- Disconnection or confusion: Participants may be confused between the real world and virtual reality and thus feel disconnected from the real world. The research team will introduce the game environment to the participants before starting the main experiment. Participants should not start the training program until after researchers have introduced the gaming environment and allowed them to get familiar with the system. Additional resources and steps are addressed in the Safety Procedures section in case the symptom persists for a longer time.

- Cybersickness: Participants may experience dizziness, headaches, or nausea because of wearing the head-mounted display device. In persistent motion sickness, the participants will be suggested to stop, postpone, or withdraw from/the study. Additional resources and steps are addressed in the Safety Procedures section in case the symptom persists for a longer time.

3.4.3 Safety

- One researcher will always be available to assist and accompany the participants in their interaction with the VR game in case of any issues.

- The participants will be instructed to sit on a firm but comfortable chair to prevent any kind of falls.

- Participants will be verbally introduced into and out of the scenarios in virtual reality to reduce the emotional stress or confusion about the game and transitions into/out of the immersive virtual environment.
- The study will end in case of any persisting discomfort (fatigue, stress, anxiety, motion sickness) explicitly expressed by the participants.
- In case of reporting fatigue, distress, or exhaustion, participants will be asked to perform a series of relaxing exercises, such as stretching or respiration, suggested by the researcher right after interacting with the virtual environment.
- Should any participants experience extreme symptoms of cybersickness at any point, the experiment will be immediately ended, and participants will be given a waste container should they need to vomit. The participants can remove and place all the equipment on the floor and proceed to a rest station where they can sit or lie down to rest. Water and granola bars will be provided to the participants as well. This will be outlined to all participants at the beginning of the experiment, along with being reminded that they can choose to end the experiment at any time they wish. In case of vomiting, researchers will handle the waste container wearing gloves and contact the custodial staff if needed.

Chapter 4

Result

4.1 Survey Results

4.1.1 User Choice Ranking

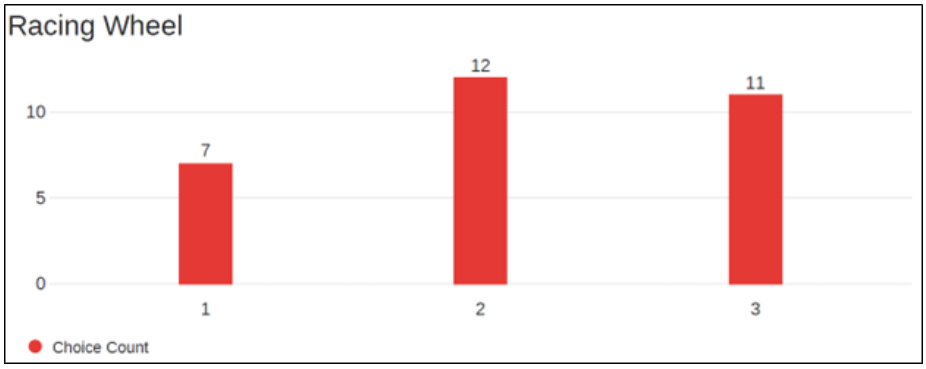

Figure 4.1: Vote of preference from the participants for the racing wheel controller condition, where "1" is the most preferred, and "3" is the least preferred.

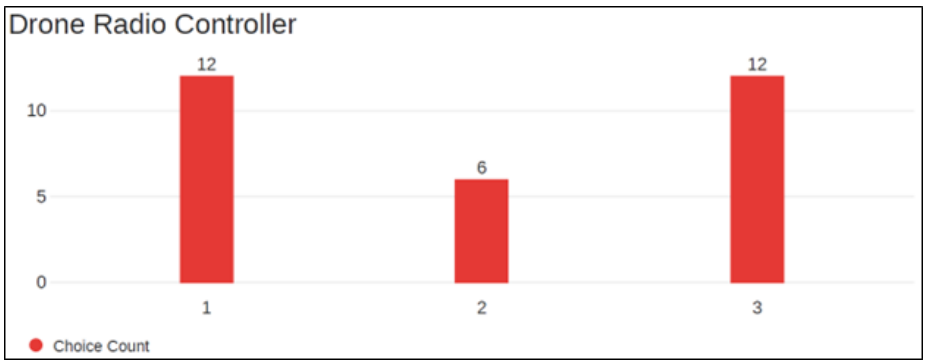

Figure 4.2: Vote of preference from the participants for the drone radio controller condition, where "1" is the most preferred, and "3" is the least preferred.

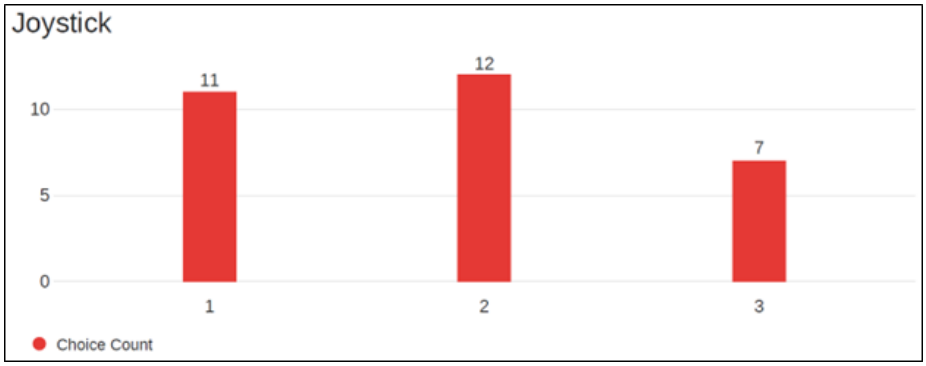

Figure 4.3: Vote of preference from the participants for the joystick controller condition, where "1" is the most preferred, and "3" is the least preferred.

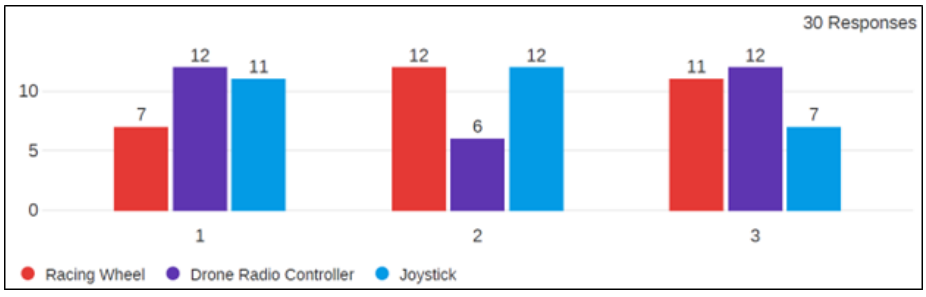

Figure 4.4: Overall ranking from the users, where rank "1" is the most preferred, and rank "3" is the least preferred.

4.1.2 NASA-TLX Score

For the NASA-TLX score, all six dimensions were equally weighted and normalized to a scale of 100. For example, if the scores are 4,3,2,2,4, and 1, the NASA-TLX score is 44.44 = ((4+3+2+2+4+1)/36) * 100.

UID User ID	Categorize Which controller did you try just now?	Q1_1 Mental Demand	Q2_1 Physical Demand	Q3_1 Temporal Demand	Q4_1 Performance	Q5_1 Effort	Q6_1 Frustration	Score NASA_TLX Score
9005	3	4	3	2	2	4	1	44.44
9005	1	3	2	2	2	2	1	33.33
9005	2	3	2	2	2	3	1	36.11
9005	4	3	3	3	2	3	1	41.67

Table 4.1: Sample of NASA-TLX data and the corresponding NASA-TLX scores with 95% CI.

	1 (Racing Wheel)	2 (Drone RC)	3 (Joystick)
Sample Size	30	30	30
Mean	**43.61**	**37.78**	**43.52**
Standard Deviation	15.44	16.39	12.87
Alpha	0.05	0.05	0.05
Confidence Interval	5.77	6.12	4.81

Table 4.2: General statistical data for NASA-TLX scores of the three controllers

A repeated measures ANOVA was performed to compare the effect of three controller conditions on the NASA-TLX score.

There was no significant difference in NASA-TLX score across the three controller conditions ($F(2, 58) = 2.298$, $p = .110$, $\eta_p^2 = .073$).

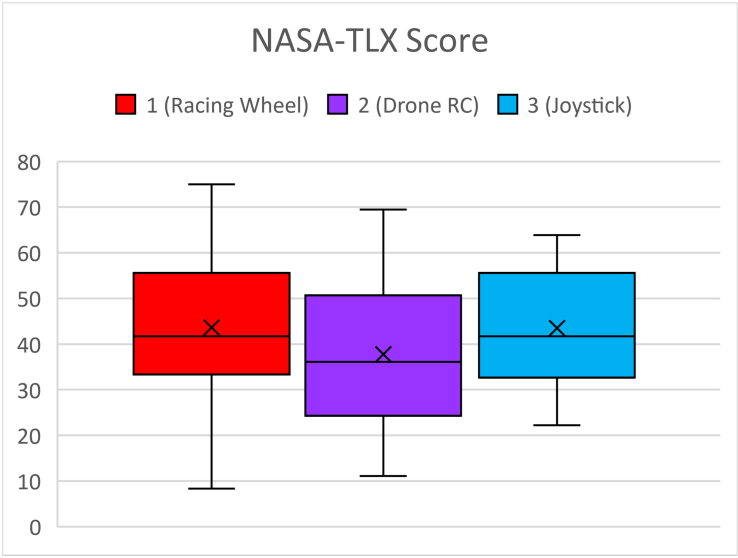

Figure 4.5: NASA-TLX box plot for the three controller conditions with 95% CI.

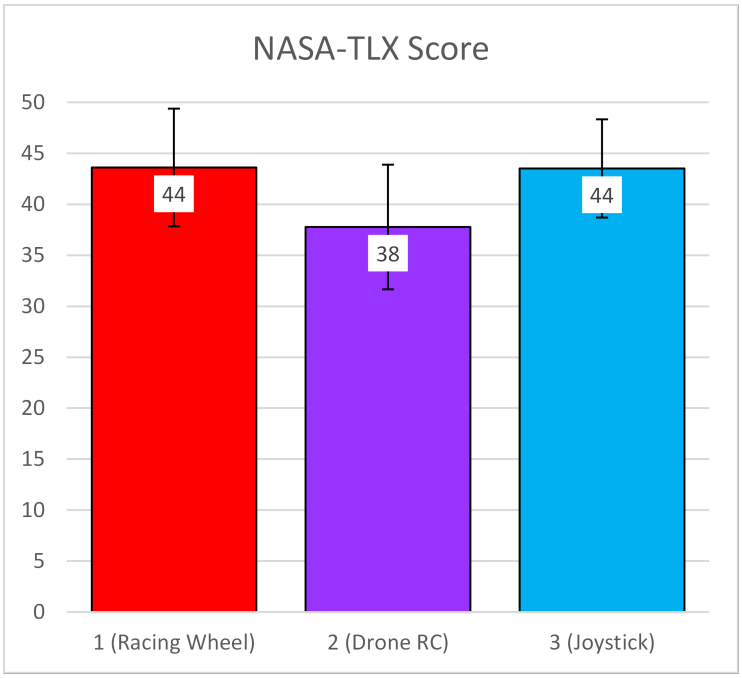

Figure 4.6: NASA-TLX bar-chart for the three controller conditions with 95% CI.

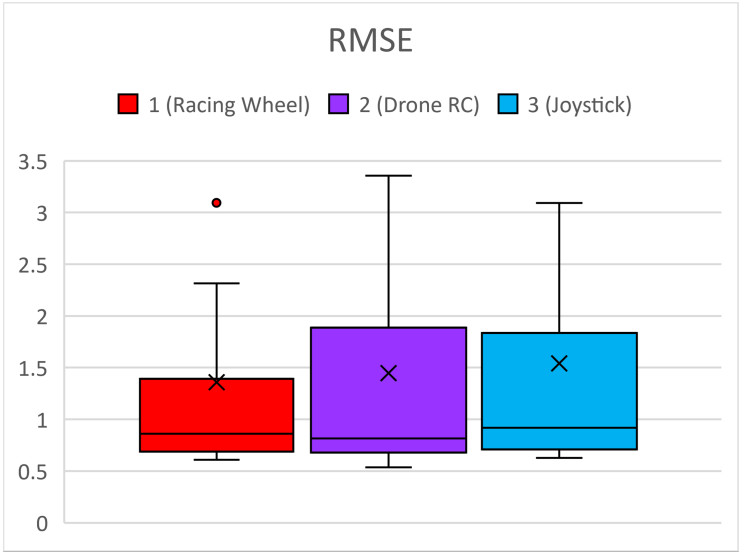

Figure 4.7: Root Mean Square Error (RMSE) box plot for the three controller conditions with 95% CI.

4.2 Performance Results

4.2.1 Root Mean Square Error

	1 (Racing Wheel)	2 (Drone RC)	3 (Joystick)
Sample Size	30	30	30
Mean	1.36	1.45	1.54
Standard Deviation	1.23	1.28	1.35
Alpha	0.05	0.05	0.05
Confidence Interval	0.46	0.48	0.50

Table 4.3: General statistical data for RMSE scores of the three controllers

A repeated measures ANOVA was performed to compare the effect of three controller conditions on the Root Mean Squared Error (RMSE). There was no significant difference in RMSE score across the three controller conditions ($F(2, 58) = 1.217$, $p = .303$, $\eta_p^2 = .040$).

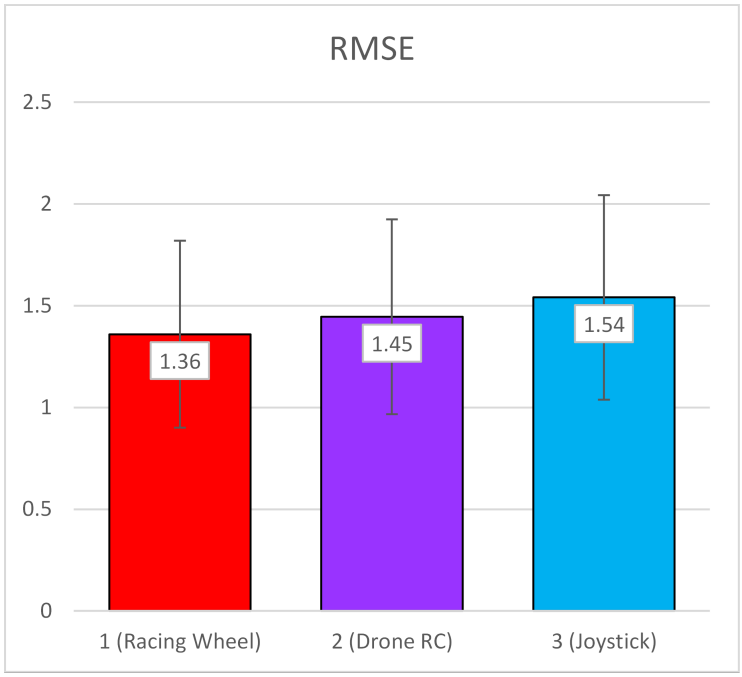

Figure 4.8: RMSE bar-chart for the three controller conditions with 95% CI.

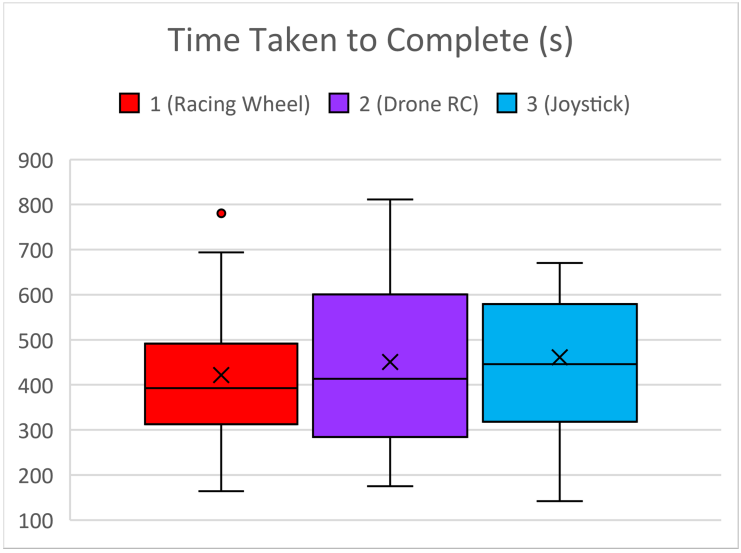

Figure 4.9: Box plot of the completion times for the three controller conditions with 95% CI.

4.2.2 Completion Time

	1 (Racing Wheel)	2 (Drone RC)	3 (Joystick)
Sample Size	30	30	30
Mean	**421.43**	**450.66**	**460.99**
Standard Deviation	154.22	222.85	192.85
Alpha	0.05	0.05	0.05
Confidence Interval	57.59	83.21	72.01

Table 4.4: General statistical data for completion time scores of the three controllers

A repeated measures ANOVA was performed to compare the effect of three controller conditions on the completion time. There was no significant difference in completion time across the three controller conditions ($F(2, 58) = 2.078$, $p = .134$, $\eta_p^2 = .067$).

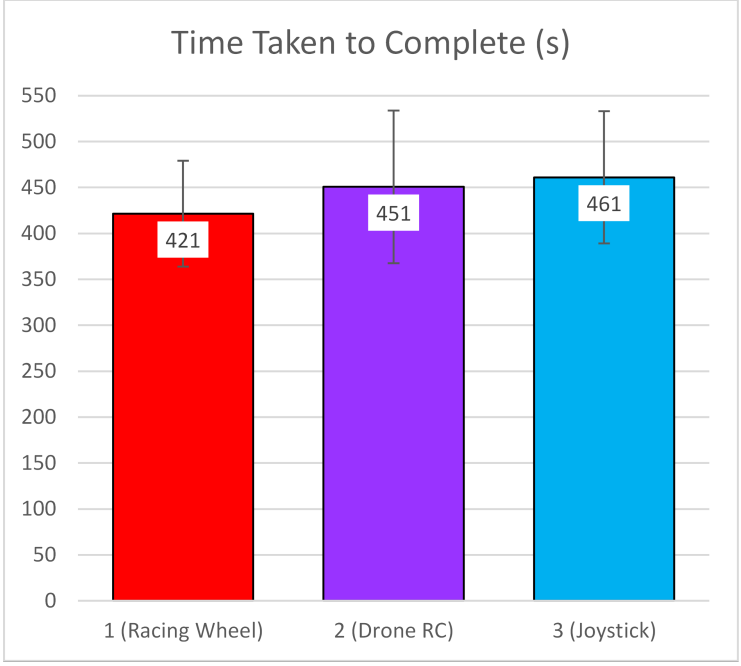

Figure 4.10: Completion time bar-chart for the three controller conditions with 95% CI.

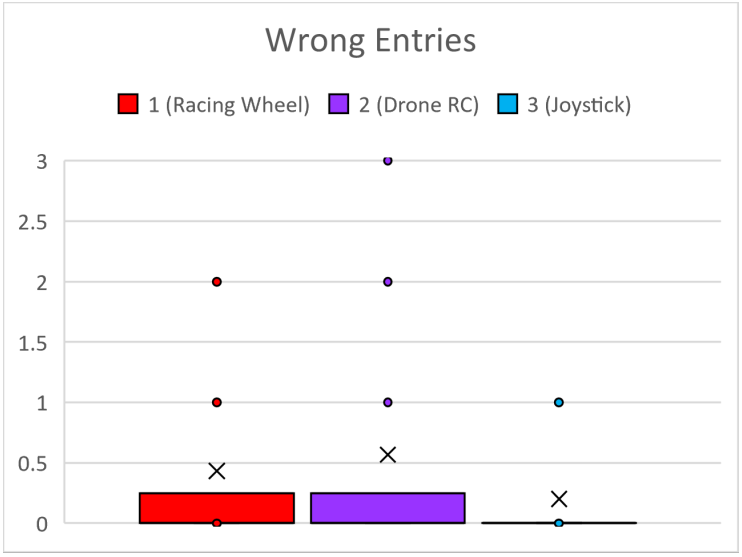

Figure 4.11: Box plot of the number of wrong entries for the three controller conditions with 95% CI.

4.2.3 Number of Wrong Entries

	1 (Racing Wheel)	2 (Drone RC)	3 (Joystick)
Sample Size	30	30	30
Mean	**0.43**	**0.57**	**0.20**
Standard Deviation	1.04	1.28	0.41
Alpha	0.05	0.05	0.05
Confidence Interval	0.39	0.48	0.15

Table 4.5: General statistical data for the number of wrong entries of the three controllers

A repeated measures ANOVA was performed to compare the effect of three controller conditions on the number of wrong entries. There was no significant difference in the number of wrong entries across the three controller conditions ($F(2, 58) = 1.959$, $p = .150$, $\eta_p^2 = .063$).

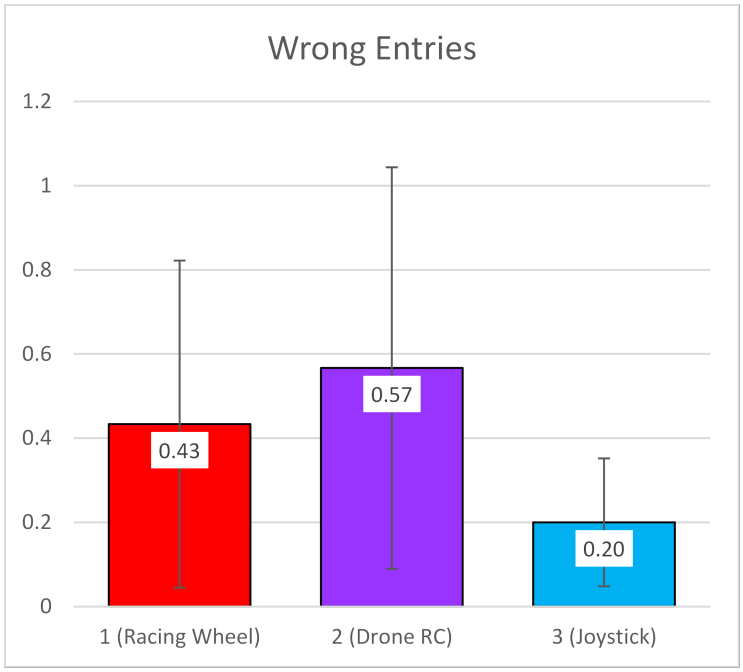

Figure 4.12: Bar-chart of the number of wrong entries for the three controller conditions with 95% CI.

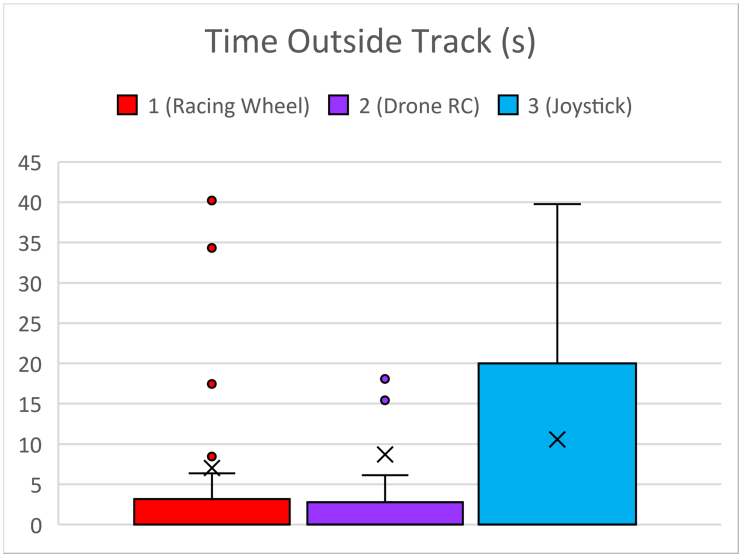

Figure 4.13: Box plot of the time spent outside track for the three controller conditions with 95% CI.

4.2.4 Time Outside Track

	1 (Racing Wheel)	2 (Drone RC)	3 (Joystick)
Sample Size	30	30	30
Mean	**7.01**	**8.72**	**10.59**
Standard Deviation	15.42	21.80	18.72
Alpha	0.05	0.05	0.05
Confidence Interval	5.76	8.14	6.99

Table 4.6: General statistical data for the time spent outside of track for the three controllers

A repeated measures ANOVA was performed to compare the effect of three controller conditions on the amount of time spent outside the track route. There was no significant difference in time spent outside track across the three controller conditions ($F(2, 58) = .352$, $p = .704$, $\eta_p^2 = .012$).

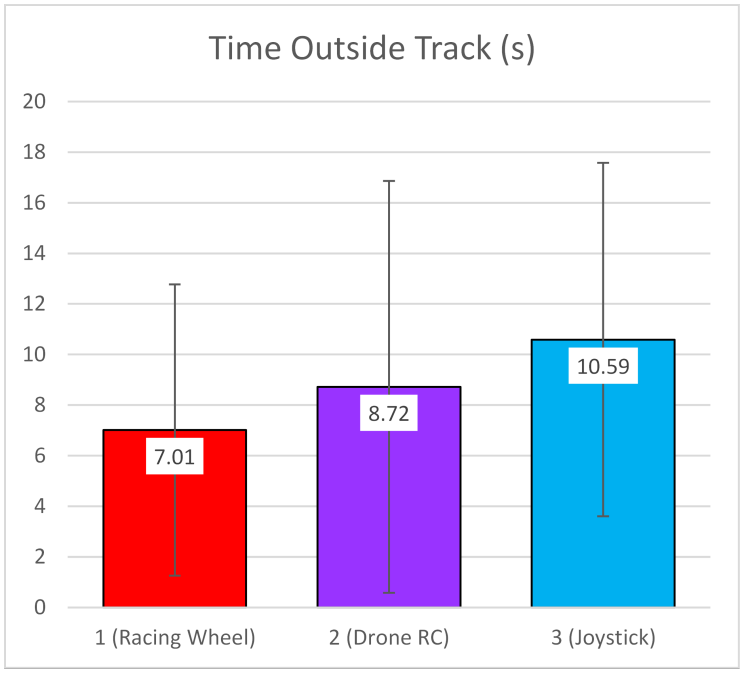

Figure 4.14: Bar-chart of the time spent outside track for the three controller conditions with 95% CI.

4.2.5 Special Condition (Racing Wheel with and without randomized wind)

Four paired samples t-tests were performed to compare performance metrics of driving using racing wheels between wind and no-wind conditions.

1. There was a significant difference in NASA-TLX score between no-wind condition (M = 41.41, SD = 14.97) and wind condition (M = 67.42, SD = 13.48); t(21) = -8.573, p = .000.

	1 (Racing Wheel)	4 (RW-Wind)
Sample Size	22	22
Mean	**41.41**	**67.42**
Standard Deviation	14.97	13.48
Alpha	0.05	0.05
Confidence Interval	6.64	5.98

Table 4.7: General statistical data for NASA-TLX scores of wind and no-wind conditions with racing wheel controller.

2. There was no significant difference in completion time between no-wind condition (M = 432.29, SD = 122.90) and wind condition (M = 438.68, SD = 130.24); t(21) = -.242, p = .811.

	1 (Racing Wheel)	4 (RW-Wind)
Sample Size	22	22
Mean	**432.29**	**438.68**
Standard Deviation	122.90	130.24
Alpha	0.05	0.05
Confidence Interval	54.49	57.75

Table 4.8: General statistical data for completion time scores of wind and no-wind conditions with racing wheel controller.

3. There was a significant difference in RMSE between the no-wind condition (M = 1.00, SD = .61) and the wind condition (M = 1.77, SD = .86); t(21) = -4.509, p = .000.

	1 (Racing Wheel)	4 (RW-Wind)
Sample Size	22	22
Mean	1.00	1.77
Standard Deviation	0.61	0.86
Alpha	0.05	0.05
Confidence Interval	0.27	0.38

Table 4.9: General statistical data for RMSE scores of wind and no-wind conditions with racing wheel controller.

4. There was no significant difference in time spent outside the track between no-wind condition (M = 3.94, SD = 12.91) and wind condition (M = 12.07, SD = 18.65); t(21) = -1.587, p = .127.

	1 (Racing Wheel)	4 (RW-Wind)
Sample Size	22	22
Mean	**3.94**	**12.07**
Standard Deviation	12.91	18.65
Alpha	0.05	0.05
Confidence Interval	5.73	8.27

Table 4.10: General statistical data for time spent outside track scores of wind and no-wind conditions with racing wheel controller.

4.3 EEG Results

4.3.1 Three main controller conditions versus baseline.

	RW_Delta_AF	DRC_Delta_AF	Joystick_Delta_AF	Baseline_Delta_AF
Sample Size	30	30	30	30
Mean	**0.4809**	**0.4554**	**0.4616**	**0.4902**
Standard Deviation	0.0669	0.0828	0.0609	0.0689
Alpha	0.0500	0.0500	0.0500	0.0500
Confidence Interval	0.0250	0.0309	0.0228	0.0257

Table 4.11: General statistical data for the EEG data of Delta band for the three controllers and the baseline

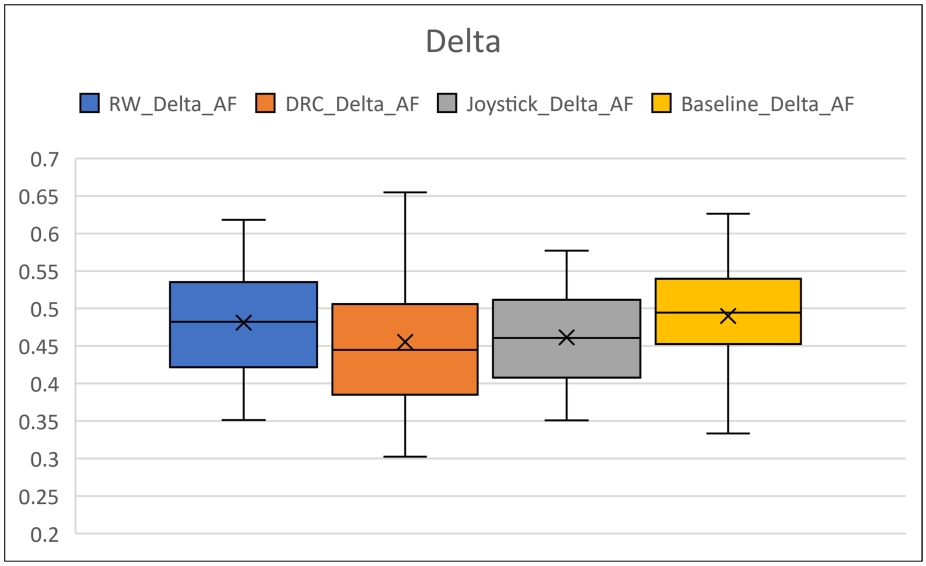

Figure 4.15: Box plot of mean-normalized values of Delta band for the three controller conditions and the baseline with 95% CI.

Friedman test was conducted to determine whether the mean-normalized score of the delta band differs between the three controller conditions compared to the baseline. The results show non-significant difference, $\chi^2(3) = 4.120$, p = .249. Therefore, there is no significant difference in the mean-normalized score of the delta band among the three controllers and the baseline.

	RW_Theta_AF	DRC_Theta_AF	Joystick_Theta_AF	Baseline_Theta_AF
Sample Size	30	30	30	30
Mean	**0.4176**	**0.4274**	**0.4246**	**0.4182**
Standard Deviation	0.0702	0.0709	0.0776	0.0794
Alpha	0.0500	0.0500	0.0500	0.0500
Confidence Interval	0.0262	0.0265	0.0290	0.0296

Table 4.12: General statistical data for the EEG data of Theta band for the three controllers and the baseline

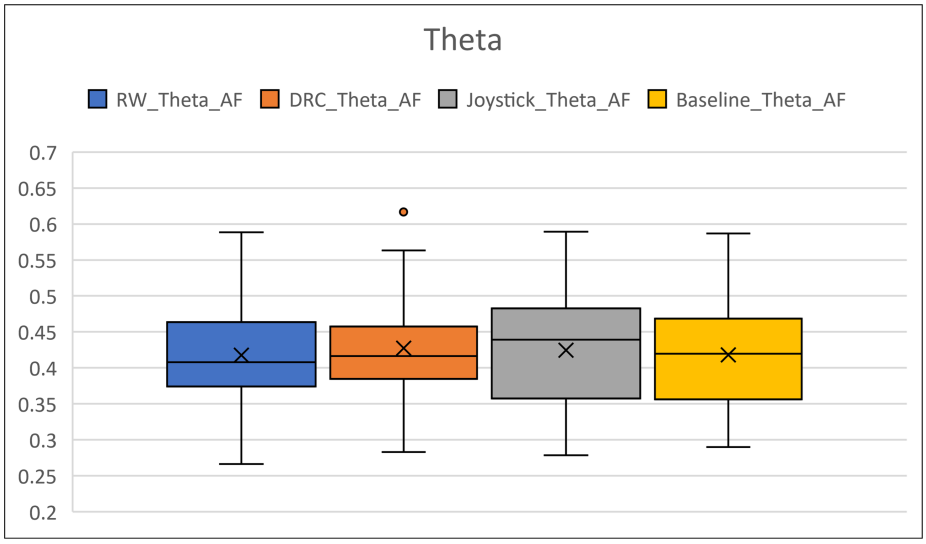

Figure 4.16: Box plot of mean-normalized values of Theta band for the three controller conditions and the baseline with 95% CI.

Friedman test was conducted to determine whether the mean-normalized score of theta band differs between the three controller conditions compared to the baseline. The results show non-significant difference, $\chi^2(3) = .280$, p = .964. Therefore, there is no significant difference in the mean-normalized score of the theta band among the three controllers and the baseline.

	RW_Alpha_AF	DRC_Alpha_AF	Joystick_Alpha_AF	Baseline_Alpha_AF
Sample Size	30	30	30	30
Mean	**0.4660**	**0.4551**	**0.4677**	**0.4721**
Standard Deviation	0.0973	0.0649	0.0558	0.0808
Alpha	0.0500	0.0500	0.0500	0.0500
Confidence Interval	0.0363	0.0242	0.0208	0.0302

Table 4.13: General statistical data for the EEG data of Alpha band for the three controllers and the baseline

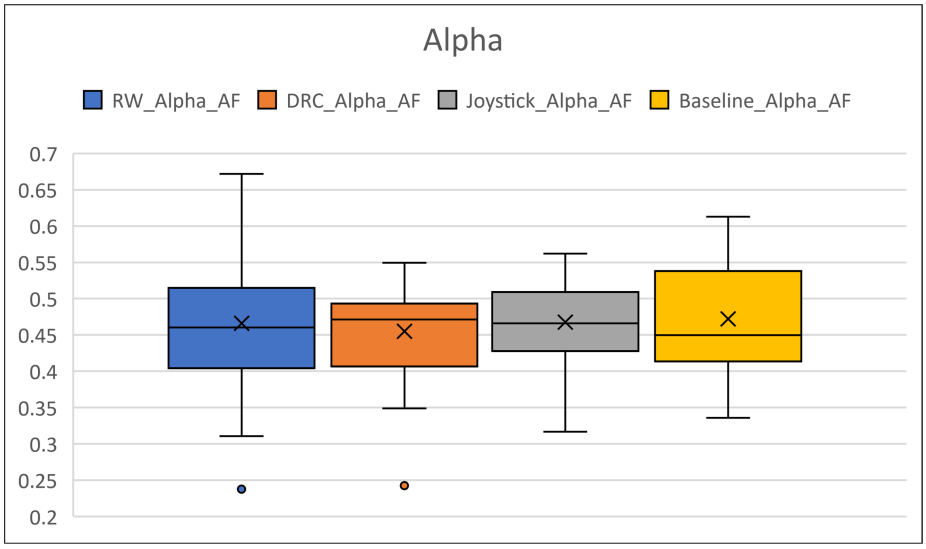

Figure 4.17: Box plot of mean-normalized values of Alpha band for the three controller conditions and the baseline with 95% CI.

Friedman test was conducted to determine whether the mean-normalized score of the alpha band differs between the three controller conditions compared to the baseline. The results show non-significant difference, $\chi^2(3) = .280$, $p = .964$. Therefore, there is no significant difference in the mean-normalized score of the alpha band among the three controllers and the baseline.

	RW_Beta_AF	DRC_Beta_AF	Joystick_Beta_AF	Baseline_Beta_AF
Sample Size	30	30	30	30
Mean	**0.4811**	**0.4758**	**0.4908**	**0.4070**
Standard Deviation	0.1215	0.1375	0.1251	0.0895
Alpha	0.0500	0.0500	0.0500	0.0500
Confidence Interval	0.0454	0.0514	0.0467	0.0334

Table 4.14: General statistical data for the EEG data of Beta band for the three controllers and the baseline

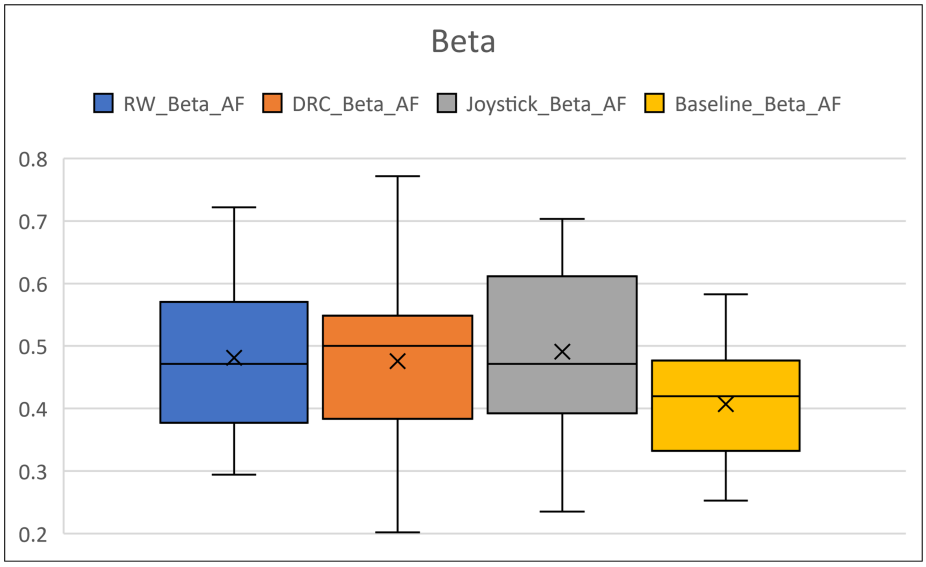

Figure 4.18: Box plot of mean-normalized values of Beta band for the three controller conditions and the baseline with 95% CI.

Friedman test was conducted to determine whether the mean-normalized score of the beta band differs between the three controller conditions compared to the baseline. The results show a significant difference, $\chi^2(3) = 9.20$, $p = .027$. Therefore, there is a significant difference in the mean-normalized score of the beta band among the three controllers and the baseline.

Three Wilcoxon Signed Rank Tests were performed to determine if there was a statistically significant difference in Beta band values between the three controllers and the baseline. A total of 30 participants were used in the analysis. Bonferonni's corrected p-value for the three tests is 0.0167.

The first test revealed that there was no statistically significant difference in Beta band

values between the racing wheel condition and the baseline (z = -2.088, p = .037).

The second test revealed that there was no statistically significant difference in Beta band values between the drone radio controller condition and the baseline (z = -2.293, p = .022).

The third test revealed that there was no statistically significant difference in Beta band values between the racing wheel condition and the baseline (z = -2.396, p = .017).

These results indicate that none of the three controller conditions significantly affected the Beta band of users' brain activity compared to the baseline of no activity.

	RW_Gamma_AF	DRC_Gamma_AF	Joystick_Gamma_AF	Baseline_Gamma_AF
Sample Size	30	30	30	30
Mean	**0.4806**	**0.4530**	**0.4692**	**0.3945**
Standard Deviation	0.1237	0.1444	0.1339	0.0881
Alpha	0.0500	0.0500	0.0500	0.0500
Confidence Interval	0.0462	0.0539	0.0500	0.0329

Table 4.15: General statistical data for the EEG data of Gamma band for the three controllers and the baseline

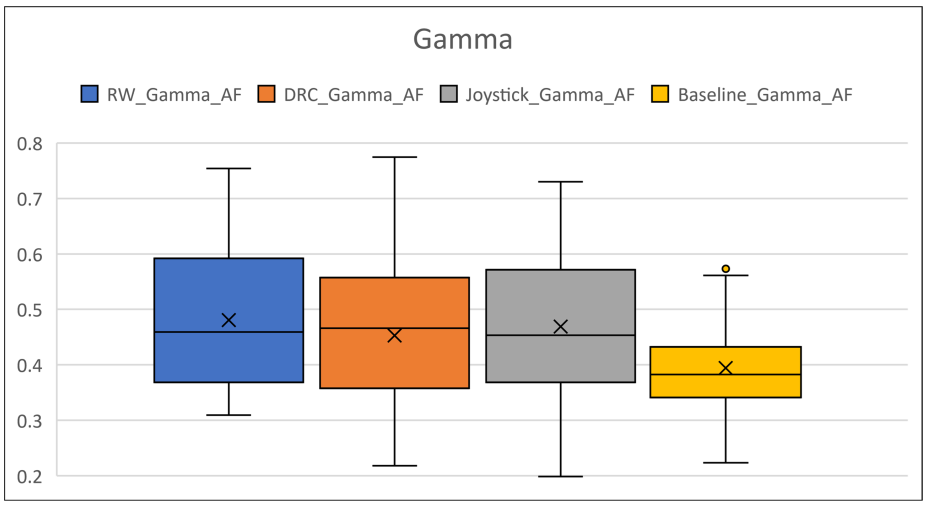

Figure 4.19: Box plot of mean-normalized values of Gamma band for the three controller conditions and the baseline with 95% CI.

Friedman test was conducted to determine whether the mean-normalized score of the gamma band differs between the three controller conditions compared to the baseline. The results show non-significant difference, $\chi^2(3) = 6.120$, p = .106. Therefore, there is no significant difference in the mean-normalized score of the gamma band among the three controllers and the baseline.

	RW_Engagement	DRC_Engagement	Joystick_Engagement	Baseline_Engagement
Sample Size	30	30	30	30
Mean	**0.5431**	**0.5405**	**0.5589**	**0.4601**
Standard Deviation	0.1035	0.1502	0.1615	0.0933
Alpha	0.0500	0.0500	0.0500	0.0500
Confidence Interval	0.0387	0.0561	0.0603	0.0348

Table 4.16: General statistical data for the EEG data of the Engagement Index for the three controllers and the baseline

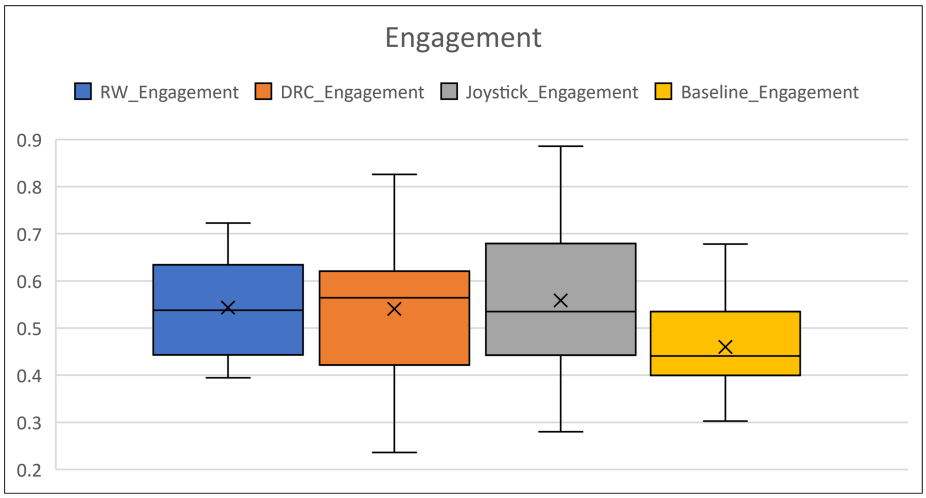

Figure 4.20: Box plot of mean-normalized values of Engagement Index for the three controller conditions and the baseline with 95% CI.

Friedman test was conducted to determine whether the mean-normalized score of the Engagement index differs between the three controller conditions compared to the baseline. The results show a significant difference, $\chi^2(3) = 9.160$, p = .027. Therefore, there is a significant difference in the mean-normalized Engagement index score among the three controllers and the baseline. Three Wilcoxon Signed Rank Tests were performed to determine if there was a statistically significant difference in engagement index values between the three controllers and the baseline. A total of 30 participants were used in the analysis. Bonferonni's corrected p-value for the three tests is 0.0167.

The first test revealed a statistically significant difference in engagement index values between the racing wheel condition and the baseline (z = -2.787, p = 0.005).

The second test revealed a statistically significant difference in engagement index values between the drone radio controller condition and the baseline (z = -2.849, p = 0.004).

The third test revealed a statistically significant difference in engagement index values between the joystick condition and the baseline ($z = -2.602$, $p = 0.009$).

These results indicate that all three controller conditions significantly affected the users' engagement compared to the baseline of no activity.

4.3.2 Racing wheel conditions of with and without wind.

	RW_Delta_AF	Wind_Delta_AF	Baseline_Delta_AF
Sample Size	22	22	22
Mean	**0.4884**	**0.4925**	**0.5105**
Standard Deviation	0.0727	0.0648	0.0591
Alpha	0.0500	0.0500	0.0500
Confidence Interval	0.0322	0.0287	0.0262

Table 4.17: General statistical data for the EEG data of Delta band for the racing wheel controller in with and without wind conditions compared to the baseline

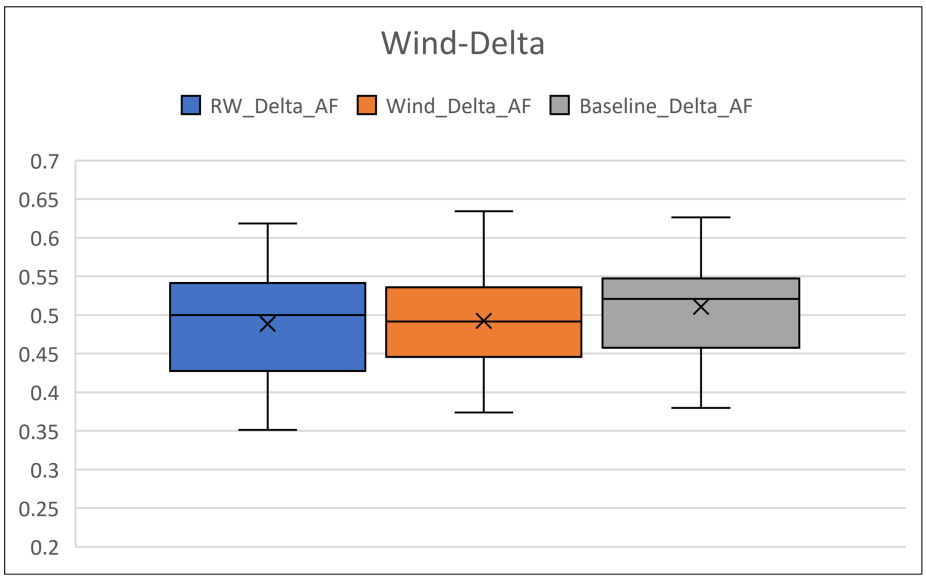

Figure 4.21: Box plot of mean-normalized values of Delta band for the racing wheel controller in with and without wind conditions compared to the baseline with 95% CI.

Friedman test was conducted to determine whether the Delta band's mean-normalized score differs between the racing wheel's without-wind and with-wind conditions. The results show non-significant difference, $\chi^2(2) = .091$, $p = .956$. Therefore, there is no significant difference in the mean-normalized score of the Delta band between the without-wind and with-wind conditions of the racing wheel.

	RW_Theta_AF	Wind_Theta_AF	Baseline_Theta_AF
Sample Size	22	22	22
Mean	**0.4126**	**0.4363**	**0.4194**
Standard Deviation	0.0662	0.0749	0.0742
Alpha	0.0500	0.0500	0.0500
Confidence Interval	0.0294	0.0332	0.0329

Table 4.18: General statistical data for the EEG data of Theta band for the racing wheel controller in with and without wind conditions compared to the baseline

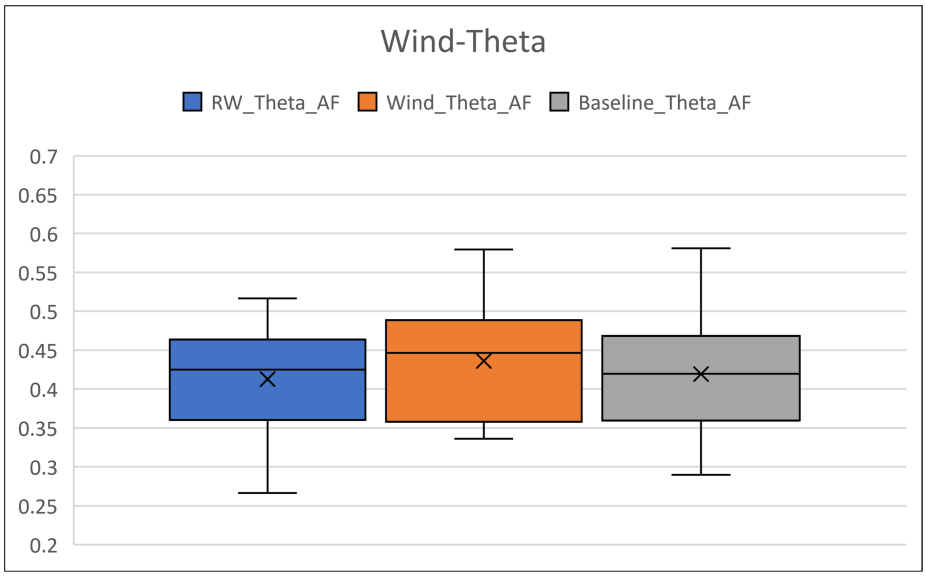

Figure 4.22: Box plot of mean-normalized values of Theta band for the racing wheel controller with and without wind conditions compared to the baseline with 95% CI.

Friedman test was conducted to determine whether the mean-normalized score of the Theta band differs between the without-wind and with-wind conditions of the racing wheel. The results show non-significant difference, $\chi^2(2) = .273$, p = .873. Therefore, there is no significant difference in the mean-normalized score of the Theta band between the without-wind and with-wind conditions of the racing wheel.

	RW_Alpha_AF	Wind_Alpha_AF	Baseline_Alpha_AF
Sample Size	22	22	22
Mean	**0.4589**	**0.4823**	**0.4748**
Standard Deviation	0.0979	0.0695	0.0830
Alpha	0.0500	0.0500	0.0500
Confidence Interval	0.0434	0.0308	0.0368

Table 4.19: General statistical data for the EEG data of Alpha band for the racing wheel controller in with and without wind conditions compared to the baseline

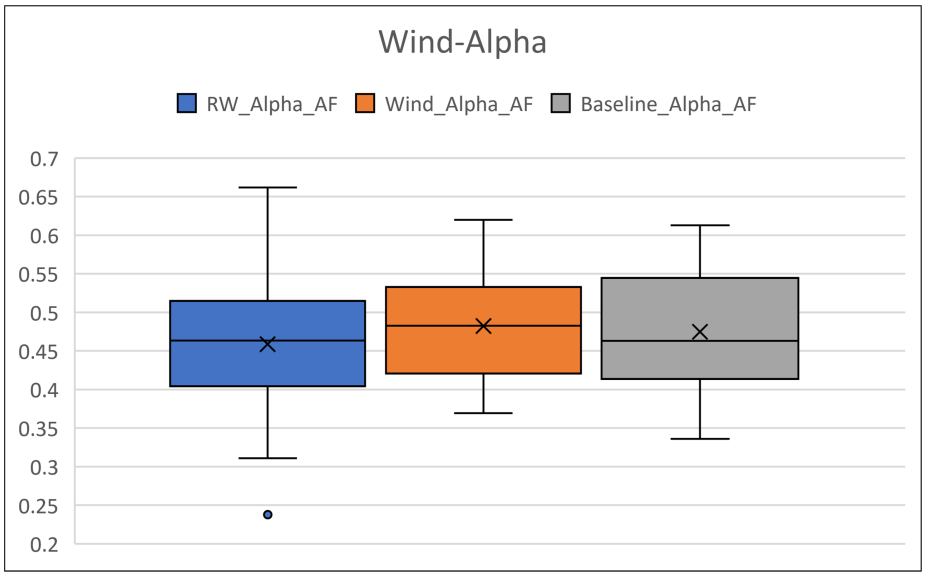

Figure 4.23: Box plot of mean-normalized values of Alpha band for the racing wheel controller in with and without wind conditions compared to the baseline with 95% CI.

Friedman test was conducted to determine whether the mean-normalized score of the Alpha band differs between the without-wind and with-wind conditions of the racing wheel. The results show non-significant difference, $\chi^2(2) = 2.273$, $p = .321$. Therefore, there is no significant difference in the mean-normalized score of the Alpha band between the without-wind and with-wind conditions of the racing wheel.

	RW_Beta_AF	Wind_Beta_AF	Baseline_Beta_AF
Sample Size	22	22	22
Mean	**0.4689**	**0.4368**	**0.3945**
Standard Deviation	0.0970	0.1399	0.0858
Alpha	0.0500	0.0500	0.0500
Confidence Interval	0.0430	0.0620	0.0380

Table 4.20: General statistical data for the EEG data of Beta band for the racing wheel controller in with and without wind conditions compared to the baseline

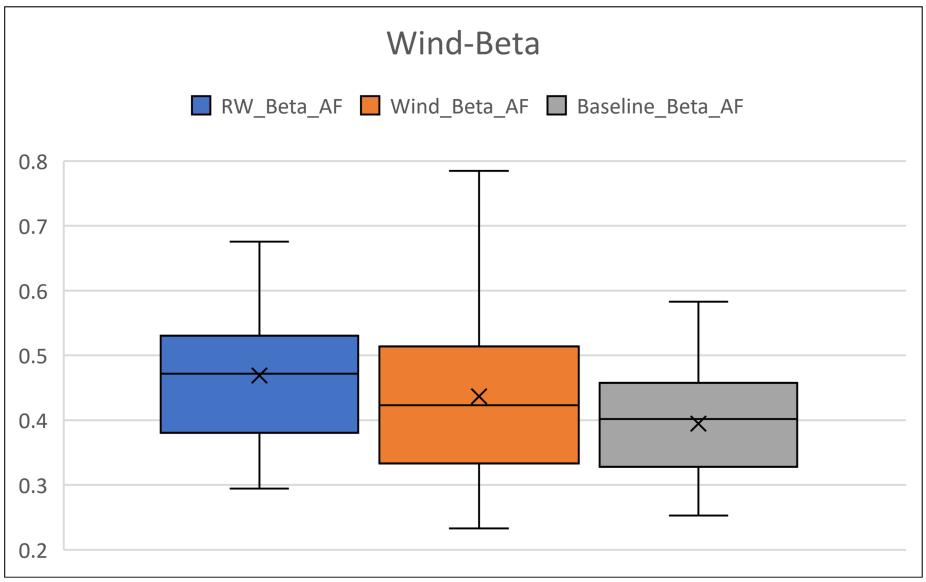

Figure 4.24: Box plot of mean-normalized values of Beta band for the racing wheel controller in with and without wind conditions compared to the baseline with 95% CI.

Friedman test was conducted to determine whether the mean-normalized score of the Beta band differs between the without-wind and with-wind conditions of the racing wheel. The results show non-significant difference, $\chi^2(2) = 5.727$, p = .057. Therefore, there is no significant difference in the mean-normalized score of the Delta band between the without-wind and with-wind conditions of the racing wheel.

	RW_Gamma_AF	Wind_Gamma_AF	Baseline_Gamma_AF
Sample Size	22	22	22
Mean	**0.4755**	**0.4406**	**0.3872**
Standard Deviation	0.1127	0.1425	0.0867
Alpha	0.0500	0.0500	0.0500
Confidence Interval	0.0500	0.0632	0.0384

Table 4.21: General statistical data for the EEG data of Gamma band for the racing wheel controller in with and without wind conditions compared to the baseline

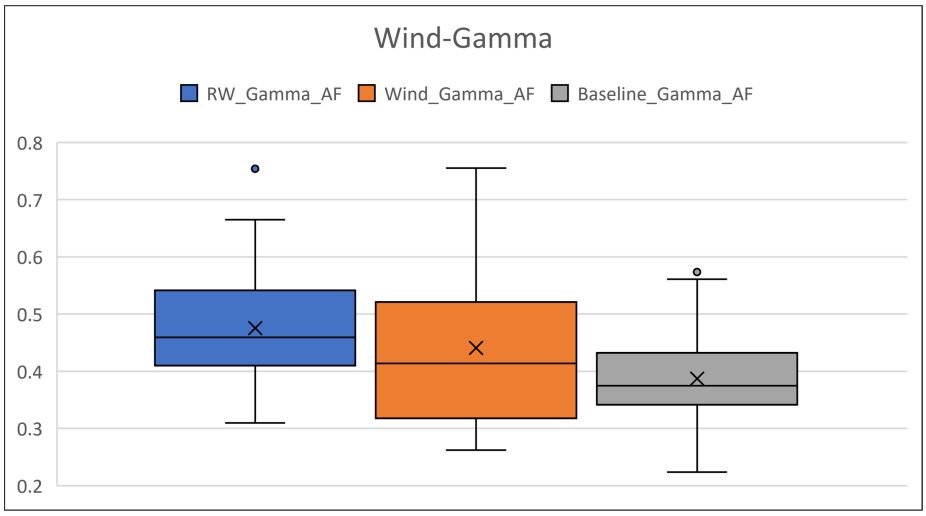

Figure 4.25: Box plot of mean-normalized values of Gamma band for the racing wheel controller in with and without wind conditions compared to the baseline with 95% CI.

Friedman test was conducted to determine whether the mean-normalized score of the Gamma band differs between the without-wind and with-wind conditions of the racing wheel. The results show non-significant difference, $\chi^2(2) = 5.182$, p = .075. Therefore, there is no significant difference in the mean-normalized score of the Gamma band between the without-wind and with-wind conditions of the racing wheel.

	RW_Engagement	Wind_Engagement	Baseline_Engagement
Sample Size	22	22	22
Mean	**0.5407**	**0.4735**	**0.4425**
Standard Deviation	0.0979	0.1385	0.0779
Alpha	0.0500	0.0500	0.0500
Confidence Interval	0.0434	0.0614	0.0345

Table 4.22: General statistical data for the EEG data of Engagement Index for the racing wheel controller in with and without wind conditions compared to the baseline

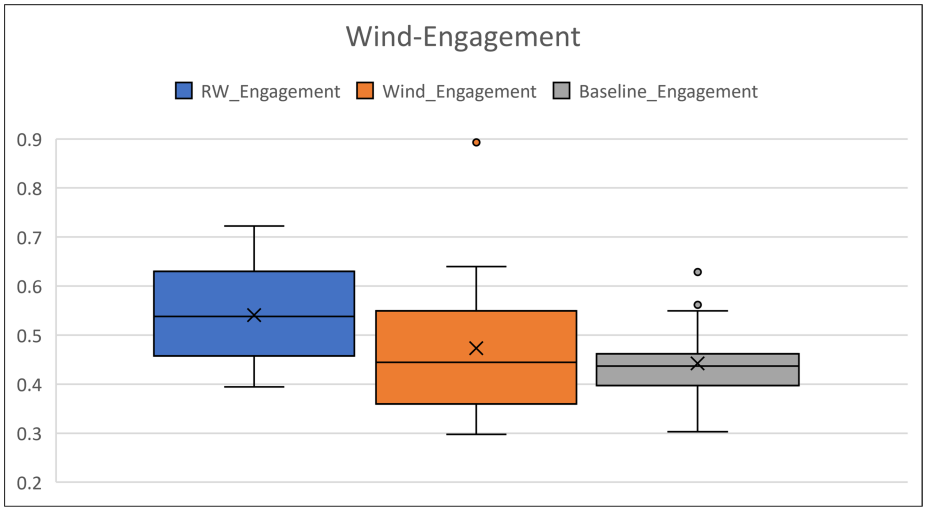

Figure 4.26: Box plot of mean-normalized values of Engagement Index for the racing wheel controller in with and without wind conditions compared to the baseline with 95% CI.

Friedman test was conducted to determine whether the mean-normalized score of the Engagement index differs between the without-wind and with-wind conditions of the racing wheel. The results show a significant difference, $\chi^2(2) = 8.818$, p = .012. Therefore, there is a significant difference in the mean-normalized score of the Delta band between the without-wind and with-wind conditions of the racing wheel.

Two Wilcoxon Signed Rank Tests were performed to determine if there was a statistically significant difference in engagement index values of the racing wheel between no-wind and wind conditions. A total of 22 participants were used in the analysis.

One test revealed that there was a statistically significant difference in engagement index values between the racing wheel no-wind condition and the baseline (z = -2.787, p = 0.005).

The other test revealed that there was no statistically significant difference in engagement index values between the racing wheel wind condition and the baseline ($z = -.179$, $p = .858$).

These results indicate that the racing wheel controller had a significant effect on the users' engagement compared to the baseline of no activity.

4.3.3 Experience vs Performance

From the pre-experiment survey, the users' previous experiences with the hardware to be tested were collected. The users rated their previous experiences on a Likert scale of five, where "0" represented "no experience" and "4" represented "used daily." Normalizing those values with equal weights and projecting them to a scale of 100, the experience score was determined.

$$ExperienceScore = (Sumofallvalues/MaximumValue) * 100 \tag{4.1}$$

Xbox	Racing Wheel	Drone RC	Joystick	VR	Experience Score
4	2	1	1	2	50
4	4	3	2	2	75
4	2	2	2	1	55
3	1	1	0	0	25

Table 4.23: A sample summary of how the pre-experiment survey data looked like for the experience portion of the survey.

The performance matric was determined based on three factors: (1) Root Mean Square Error (RMSE), (2) NASA-TLX Score, and (3) Time for Completion. As accuracy was the main factor to consider, the RMSE score was given a weight of 0.5, followed by the NASA-TLX score, a weight of 0.3, and the time for completion, a weight of 0.2. RMSE Score = ((13-RMSE)/13)*0.50, where 13 is the radius of each checkpoint. NASA-TLX Score = (100-NASA-TLX)*0.30 Time Score = (MaxTime-TimeTaken)/MaxTime*0.20, where MaxTime is the maximum time taken among all participants, and TimeTaken represents the time taken by the current participant.

From the two graphs presented above, we can clearly see that the previous experiences of the users played a linear relationship to how well they performed in the actual study.

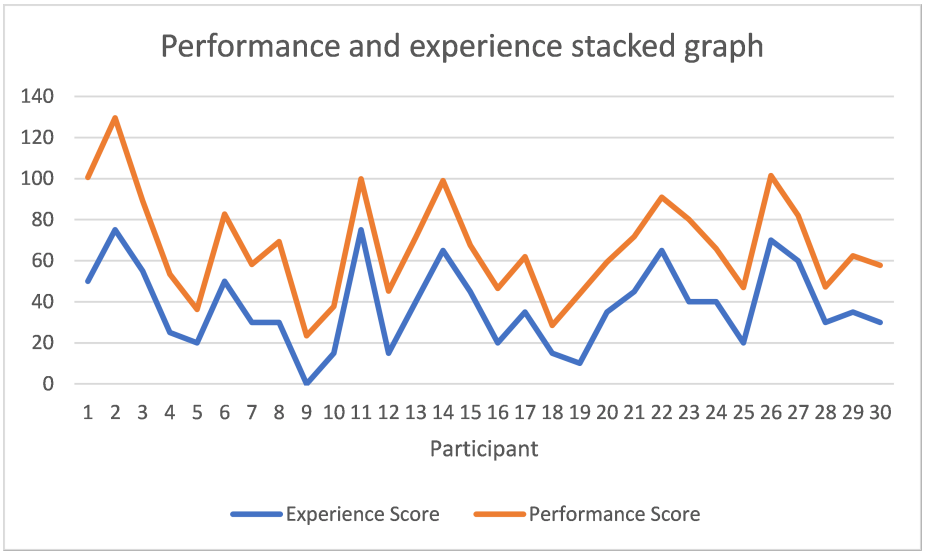

Figure 4.27: Performance and experience stacked graph.

Figure 4.28: Scatter plot of Experience vs Performance score

4.3.4 Cybersickness

	1 (Racing Wheel)	2 (Drone RC)	3 (Joystick)
Sample Size	22	22	22
Mean	**1.95**	**1.95**	**2.00**
Standard Deviation	1.36	1.25	1.45
Alpha	0.05	0.05	0.05
Confidence Interval	0.60	0.56	0.64

Table 4.24: Survey data of cybersickness for the three controller conditions and the with-wind condition.

A repeated measures ANOVA was performed to compare the effect of three controller conditions on the cybersickness experienced by the users. There was no significant difference in cybersickness across the three controller conditions ($F(2, 58) = .396$, $p = .675$, $\eta_p^2 = .233$).

	1 (Racing Wheel)	4 (RW-Wind)
Sample Size	22	22
Mean	**1.95**	**2.95**
Standard Deviation	1.36	2.28
Alpha	0.05	0.05
Confidence Interval	0.60	1.01

Table 4.25: Survey data of cybersickness for the with-wind and without-wind racing wheel controller conditions.

There was a significant difference in cybersickness score between the no-wind condition ($M = 1.95$, $SD = 1.36$) and the wind condition ($M = 2.95$, $SD = 2.28$); $t(21) = -2.569$, $p = .018$.

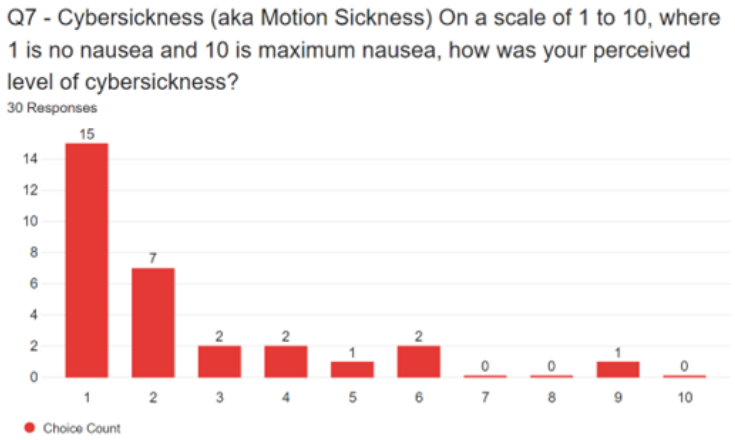

Figure 4.29: Cybersickness survey result for Racing Wheel.

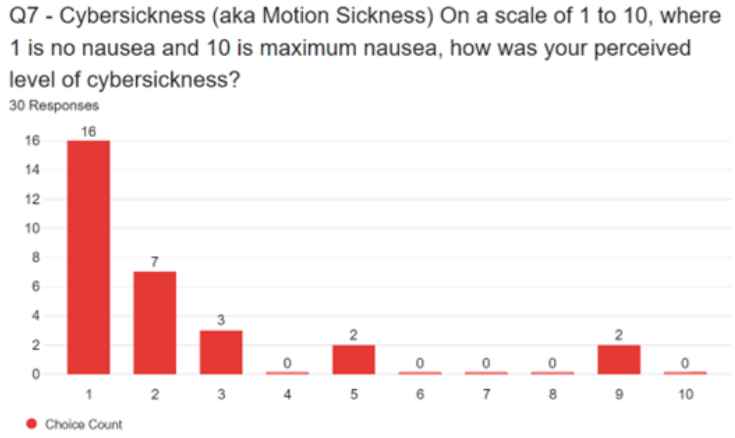

Figure 4.30: Cybersickness survey result for Drone Radio Controller.

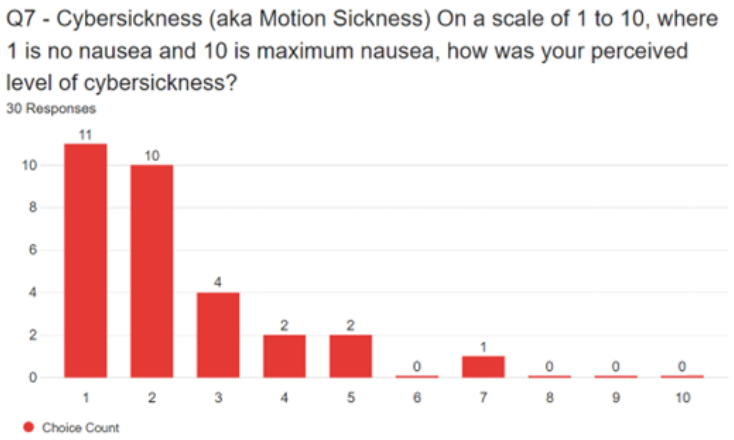

Figure 4.31: Cybersickness survey result for Joystick.

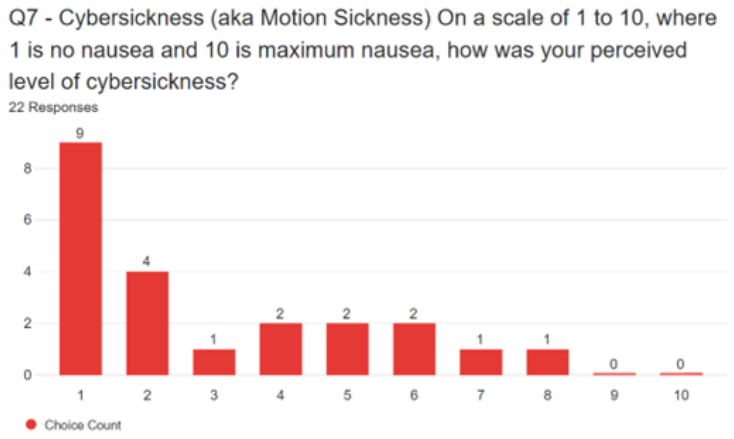

Figure 4.32: Cybersickness survey result for Racing Wheel with randomized wind condition.

4.4 Discussion

4.4.1 Findings

1. According to the survey, participants preferred the joystick controller over the other two controllers, but the result is not significant enough p is greater than 0.05.

2. According to the performance data gathered, there was no significant difference in performance for the three controller conditions. So, neither of the controllers can be said to perform better than the other in terms of user performance.

3. Racing wheel controllers with and without wind conditions significantly differed in NASA-TLX scores and RMSE values. Therefore, wind condition can make a significant impact on the performance and is required to be considered for VTOL simulations and training.

4. From EEG data, there was a significant difference in mean-normalized Beta band values for the three controllers compared to the baseline.

5. From EEG data, there was a significant difference in mean-normalized Engagement Index values for the three controllers compared to the baseline. Proceeding further to figure out which pair had significant differences, it was statistically deduced that all three controllers had significant differences compared to the baseline. All three p-values were significantly lower than the acceptable p-value of less than 0.0167 (Bonferroni corrected p-value). Out of the three, the drone radio controller had the lowest p-value of 0.004, followed by the racing wheel controller p-value of 0.005, ending with the joystick p-value of 0.009. Therefore, even though all three controllers had a significant engagement compared to the relaxed state of a user, the drone radio controller had the highest engagement, followed by the racing wheel controller, and the joystick had the least engagement compared to the three.

6. While looking at the box plots of EEG signals, it is notable that the beta values for the three controllers are significantly higher than the baseline, meaning the controllers required more attention to get the job done than the relaxed state.

7. Previous experience directly correlated to the performance of the participants. That means the more the population is used to the controller, the more viable the controller is for the VTOL aircraft. One might question whether VR training can directly transfer to real-life experience. Still, based on Drone training examples and pilot

training examples, we can say that having experience in any given controller can help in real-life vehicle maneuvers using the same controller. With that being said, too much exposure and muscle memory can be detrimental as well. For example, suppose someone is extremely used to flying an aircraft in a virtual game where the controller styles might be different than in real life. In that case, they might end up triggering certain actions unintentionally in real life due to muscle memory, which can cause serious damage.

8. From the EEG, one interesting finding was the racing controller with wind conditions. Even with all the randomized wind and the significantly high mental workload, the users' engagement with the task was close to the baseline, as if they were doing the work without actively focusing on the task. This might be because the users had already gone through six different runs (three for the tutorial, three for the main session), and by the time they reached this special condition, they got bored or got the hang of the task. Yes, their mental workload was high as they needed to adjust the sideways movement of the VTOL aircraft while moving forward, but it became more of a mere task that they wanted to get over with.

9. There was little to no cybersickness for the users for the three main controller conditions. Regarding the racing wheel controller condition with wind, there was mild cybersickness to some of the users, but not too harsh for them to be unable to complete the session. There were two participants as exceptions out of 30 participants, one of whom has general motion sickness issues in real life, which also translated to the VR world. Among all the conditions, the joystick had the least mean cybersickness reported, followed by the racing wheel, drone radio controller and the racing wheel with wind conditions.

4.4.2 Discussion

Looking at the results, we find that the joystick controller was the most preferred by the users, having the least amount of last rank preference at 23%, compared to 37% for racing wheel and 40% for drone racing wheel. The SPSS data did say that the preference was not significant enough. The joystick also had the lowest value of mean number of wrong entries compared to the rest; 0.2 for the joystick compared to 0.43 for the racing wheel and 0.57 for the drone radio controller. The EEG data showed that the joystick controller was very close to having a significant difference in Beta values (p=0.017, expected p¡0.0167 for it to be significant). The other two controllers did not even come close with p values of 0.022 and 0.037. As the beta value is associated with attention, it can be said that the joystick

almost required significant attention to operate compared to the rest. Furthermore, even though the Engagement index using EEG data showed all three controllers had significant engagement required compared to the baseline, the joystick required the least engagement with a p-value of 0.009, compared to 0.005 for the racing wheel and 0.004 for the drone radio controller. Ensuring a system is safe enough by keeping users engaged is essential, and requiring the least amount of engagement while it is significant enough is the ideal situation. The joystick is the winner among the three controls, based on the significant EEG data of the engagement index.

One issue with the joystick controller was mixing too many controls into one stick. Yes, it is intuitive. But, sometimes, very advanced manoeuvres are triggered by the users without them even realizing that they triggered it. For example, the user might want to go front-left. But when they are moving the stick, they mistakenly twist the stick to the left as well, making the car rotate left while moving forward and left. To solve this problem, the rotation portion of the control could be moved to a set of pedals for the legs, just like it is used by the helicopters in the current market. If sticking within joystick controls only is a must and no pedals can be used, the rotation can be placed on the thumb analogue stick, separating the rotational movement from the translational movement.

One might ask why the racing wheel did not pan out to be the best controller among the three. The answer to that question lies in three parts. First of all, the statistical data did not have significant differences that can support the claim. Secondly, the joystick is much more intuitive if we consider the motions and their controller equivalents; front by moving stick front, left moving stick left, and rotation clockwise by twisting the stick clockwise. The racing wheel cannot do that. Plus, having an extra dimension (up and down) can make the users think harder, as it goes against their everyday driving and the respective controls. Thirdly, many users said the racing wheel felt a bit move in terms of sensitivity and motion; the control in the simulation was said to be more sensitive than in real life, and the motion did not translate well compared to real-life driving. Previous experiences on such matters definitely affected their expected controls.

The performance of a user is directly related to their previous experiences, and past papers show that training in simulation can translate to better performance in real life. Therefore, the skills are likely to translate into the real world if the joystick or any other top-of-the-shelf controllers are used for VTOL aircraft. When designing the controls for VOTL aircraft, the existing market should be taken into consideration while making the least amount of changes to make it a viable and efficient option for users to catch onto it fast.

The wind condition for the racing wheel had a significant difference in performance

compared to the no-wind condition. It clearly shows that real-life environmental situations will affect driving performance, and the users should practice simulation with such conditions in mind.

4.4.3 Limitations

Like any other user study, there are some limitations to this study as well. First of all, as all three controllers are off-the-shelf controllers made for a wide variety of applications, their physical setup, controls and range of motions varied a lot. The specific controls had to be mapped to all three controllers to the best of our understanding, research and testing. For future studies, the existing controllers should be modified, or controllers should be designed specifically for VTOL aircraft.

Secondly, the testing system that was arranged for the study was in virtual reality. Even though we tried our best to keep it as real as possible, it is still a simulation. The users might not have acted exactly the same as they would have done if they were asked to trial a real VTOL aircraft.

Thirdly, as the time for the experiment was quite long, some of the users seemed to lose interest in some of the later conditions as they were going through the same track again and again. The randomized order of controllers took care of that issue, but it might have affected the performance of certain users and, indirectly, the metric of the surveys.

Chapter 5

Conclusion

5.1 Future work

Based on the results found and discussions made, there are three points to take away for future work:

1. External environmental conditions can directly affect the performance of a user. Further experiments can be carried out with wind, rain, snow, storm or other external conditions to help gather data with statistical significance among the three controllers.

2. People can practice in simulation first and then try the real-life VTOL aircraft. Suppose the data can be recorded for such scenarios. In that case, the policymakers might allow simulator training for VTOL aircraft to reduce the hours of actual flight time needed to get a license to drive such a vehicle.

3. The following modified controls could be considered for future user studies:

    - Joystick with pedals or thumb stick for rotational movement.
    - Drone radio controller with independent buttons for yaw.
    - Racing wheel with closer to real-life controls and sensitivity.

4. Another study could be conducted to consider whether the suggested controllers would be easier to accommodate pilots with real-life flying experiences. Also, the training effect on the VR simulation could be taken into account as well.

5.2 Conclusion

This study set out to determine the most effective controller for VTOL aircraft simulation, taking into consideration performance data, physiological data from EEG measurements, and user feedback. The findings weakly indicate that the joystick controller emerged as the preferred choice among users, with the lowest last-rank preferences, the fewest mean number of wrong entries, and notable engagement levels. While the EEG data showed that the joystick's difference in Beta values was almost significant, it strongly suggested that the joystick required significant attention, aligning with the engagement index results. The importance of user engagement in ensuring system safety cannot be understated, and the joystick's ability to strike a balance between engagement and effectiveness makes it the standout choice. That said, the repeated ANOVA result was not significant enough to strongly affirm the joystick as the best choice.

The study also identified a key issue with the joystick controller, namely, the potential for the unintentional triggering of advanced manoeuvres due to its integrated controls. To address this concern, separating rotational movement from translational movement, possibly using pedals or a thumb analogue stick, could improve user experience and prevent accidental actions.

As for why the racing wheel did not perform as well, several factors contributed to this outcome. Statistical data did not support its superiority, and the joystick's intuitiveness, mirroring real-world motions, played a significant role. Users' perceptions of sensitivity and motion also influenced their preference, as the simulation experience did not align with their real-life driving experiences. This underscores the importance of considering users' prior experiences when designing VTOL aircraft controls.

Furthermore, the study highlighted the impact of environmental conditions, particularly wind, on controller performance. Real-world scenarios and conditions should be incorporated into training and simulation exercises to prepare users effectively for varying situations. Further user studies should be carried out for the three controllers along with wind conditions to help get a more accurate picture of which controller would be accepted in the real world.

While the study provides valuable insights, it is essential to acknowledge its limitations, such as variations in off-the-shelf controllers, the simulated nature of the experiment, and potential user fatigue over extended testing periods. Future research in this area should explore the development of specialized VTOL aircraft controllers and strive for a more realistic simulation environment.

In summary, this study contributed valuable knowledge about manual controller pref-

erences and their impact on VTOL aircraft driving. The joystick controller stands out as the preferred choice, but the findings also emphasize the need for thoughtful controller design, user training, and consideration of real-world conditions in VTOL aircraft operations. These insights can guide the development of more effective and user-friendly manual control systems for this emerging technology.

References

[1] Product gallery - extreme 3d pro.

[2] Product gallery - g27 racing wheel.

[3] Tx12 mark ii radio controller.

[4] Anon.: Personal perspective on helicopter history, 1999.

[5] Volocopter: Aviation history – first piloted air taxi flight of vc200 – future of mobility, 2016. YouTube video.

[6] Volocopter 2x debuts at aero, 2019.

[7] Blackfly wants to be the flying car you can finally buy next year, and for cheap, 2021.

[8] Jetson one - take off in tuscany, 2021. YouTube video.

[9] Volocopter conducts south korea's first crewed public air taxi test flight, 2021.

[10] Jetson one product page, 2023.

[11] Opener product page, 2023.

[12] Volocopter product page, 2023.

[13] Sheikh Shahriar Ahmed, Grigorios Fountas, Ugur Eker, Stephen E Still, and Panagiotis Ch Anastasopoulos. An exploratory empirical analysis of willingness to hire and pay for flying taxis and shared flying car services. *Journal of Air Transport Management*, 90:101963, 2021.

[14] Edwin W Aiken and Kenneth H Landis. An assessment of various side-stick controller/stability and control augmentation systems for night nap-of-the-earth flight using piloted simulation," presented to_ the american helicopter society, palo alto, calif., apr. 1982. 11.

[15] Mahmoud I Al-Kadi, Mamun Bin Ibne Reaz, and Mohd Alauddin Mohd Ali. Evolution of electroencephalogram signal analysis techniques during anesthesia. *Sensors*, 13(5):6605–6635, 2013.

[16] D Aláez, X Olaz, M Prieto, J Villadangos, and JJ Astrain. Vtol uav digital twin for take-off, hovering and landing in different wind conditions. *Simulation Modelling Practice and Theory*, 123:102703, 2023.

[17] Ross D Arnold, Hiroyuki Yamaguchi, and Toshiyuki Tanaka. Search and rescue with autonomous flying robots through behavior-based cooperative intelligence. *Journal of International Humanitarian Action*, 3(1):1–18, 2018.

[18] Hasan Ayaz and Frédéric Dehais. *Neuroergonomics: The brain at work and in everyday life*. Academic Press, 2018.

[19] Nathan R Bailey, Mark W Scerbo, Frederick G Freeman, Peter J Mikulka, and Lorissa A Scott. Comparison of a brain-based adaptive system and a manual adaptable system for invoking automation. *Human factors*, 48(4):693–709, 2006.

[20] Peter A Bandettini. What's new in neuroimaging methods? *Annals of the New York Academy of Sciences*, 1156(1):260–293, 2009.

[21] Victor M Becerra. Autonomous control of unmanned aerial vehicles, 2019.

[22] Chris Berka, Daniel J Levendowski, Michelle N Lumicao, Alan Yau, Gene Davis, Vladimir T Zivkovic, Richard E Olmstead, Patrice D Tremoulet, and Patrick L Craven. Eeg correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation, space, and environmental medicine*, 78(5):B231–B244, 2007.

[23] Stephen Bridgewater. Electrify your commute, 2023.

[24] Leonard Bridgman. *Jane's All the World's Aircraft 1955-56*. McGraw, 1956.

[25] Yun-Hsuan Chen, Maaike Op De Beeck, Luc Vanderheyden, Evelien Carrette, Vojkan Mihajlović, Kris Vanstreels, Bernard Grundlehner, Stefanie Gadeyne, Paul Boon, and Chris Van Hoof. Soft, comfortable polymer dry electrodes for high quality ecg and eeg recording. *Sensors*, 14(12):23758–23780, 2014.

[26] Esmond Nigel Corlett, John R Wilson, and NIGEL CORLETT. *Evaluation of human work*. CRC Press, 1995.

[27] Gordon Darroch. Netherlands 'will pay the price' for blocking turkish visit – erdoğan. *The Guardian*.

[28] Arnaud Delorme and Scott Makeig. Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of neuroscience methods*, 134(1):9–21, 2004.

[29] Roger D Dias, Minhtran C Ngo-Howard, Marko T Boskovski, Marco A Zenati, and Steven J Yule. Systematic review of measurement tools to assess surgeons' intraoperative cognitive workload. *Journal of British Surgery*, 105(5):491–501, 2018.

[30] Daniel Dollinger, Philipp Reiss, Jorg Angelov, David Löbl, and Florian Holzapfel. Control inceptor design for onboard piloted transition vtol aircraft considering simplified vehicle operation. In *AIAA Scitech 2021 Forum*, page 1896, 2021.

[31] JM Drees. Prepare for the 21st century—the 1987 alexander a. nikolsky lecture. *Journal of the American Helicopter Society*, 32(3):3–14, 1987.

[32] Xu Duan, Songyun Xie, Xinzhou Xie, Ya Meng, and Zhao Xu. Quadcopter flight control using a non-invasive multi-modal brain computer interface. *Frontiers in neurorobotics*, 13:23, 2019.

[33] Guillaume JJ Ducard and Mike Allenspach. Review of designs and flight control techniques of hybrid and convertible vtol uavs. *Aerospace Science and Technology*, 118:107035, 2021.

[34] Art Equilibrium. Hq cyberpunk car, 2020.

[35] Frank Flemisch, Julian Schindler, Johann Kelsch, Anna Schieben, Christian Löper, Matthias Heesen, Daniel Damböck, Martin Kienle, Klaus Bengler, Jörg Dittrich, et al. Kooperative führung hochautomatisierter boden-und luftfahrzeuge am beispiel h-mode luft/boden. In *Kooperative Arbeitsprozesse*, 2009.

[36] Lisa R Fournier, Glenn F Wilson, and Carolyne R Swain. Electrophysiological, behavioral, and subjective indexes of workload when performing multiple tasks: manipulations of task difficulty and training. *International Journal of Psychophysiology*, 31(2):129–145, 1999.

[37] Frederick G Freeman, Peter J Mikulka, Mark W Scerbo, Lawrence J Prinzel, and Keith Clouatre. Evaluation of a psychophysiologically controlled adaptive automation system, using performance on a tracking task. *Applied Psychophysiology and Biofeedback*, 25:103–115, 2000.

[38] Frederick G Freeman, Peter J Mikulka, Mark W Scerbo, and Lorissa Scott. An evaluation of an adaptive automation system using a cognitive vigilance task. *Biological psychology*, 67(3):283–297, 2004.

[39] Jérémy Frey, Christian Mühl, Fabien Lotte, and Martin Hachet. Review of the use of electroencephalography as an evaluation method for human-computer interaction. *arXiv preprint arXiv:1311.2222*, 2013.

[40] Laurie A Garrow, Patricia Mokhtarian, Brian German, and Sreekar-Shashank Boddupalli. Commuting in the age of the jetsons: a market segmentation analysis of autonomous ground vehicles and air taxis in five large us cities. In *AIAA Aviation 2020 Forum*, page 3258, 2020.

[41] Leandre Grecia. Could this single-seat aircraft be the future of personal mobility?, 2021.

[42] Rebecca A Grier. How high is high? a meta-analysis of nasa-tlx global workload scores. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 59, pages 1727–1731. SAGE Publications Sage CA: Los Angeles, CA, 2015.

[43] Heidi Groshelle. Opener early access program moves blackfly closer to general release, 2023.

[44] Anna Gruetzmann, Stefan Hansen, and Jörg Müller. Novel dry electrodes for ecg monitoring. *Physiological measurement*, 28(11):1375, 2007.

[45] Bianca I Gursky and D Muller. Novel steering concepts for personal aerial vehicles. *DLR Institute of Flight Systems*, 2017(04/28), 2013.

[46] PA Hancock. Neuroergonomics: Where the cortex hits the concrete. *Frontiers in Human Neuroscience*, 13:115, 2019.

[47] Sandra G Hart. Nasa-task load index (nasa-tlx); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 50, pages 904–908. Sage publications Sage CA: Los Angeles, CA, 2006.

[48] Hal W Hendrick. The technology of ergonomics. *Theoretical Issues in Ergonomics Science*, 1(1):22–33, 2000.

[49] Keith C Hendy, Kevin M Hamilton, and Lois N Landry. Measuring subjective workload: when is one scale better than many? *Human Factors*, 35(4):579–601, 1993.

[50] Acuity Technologies Inc. At-10 vtol uas. Accessed: 2011.

[51] Robert Jenssen, Davide Roverso, et al. Automatic autonomous vision-based power line inspection: A review of current status and the potential role of deep learning. *International Journal of Electrical Power & Energy Systems*, 99:107–120, 2018.

[52] Altyngul T Kamzanova, Almira M Kustubayeva, and Gerald Matthews. Use of eeg workload indices for diagnostic monitoring of vigilance decrement. *Human factors*, 56(6):1136–1149, 2014.

[53] Waldemar Karwowski. Ergonomics and human factors: the paradigms for science, engineering, design, technology and management of human-compatible systems. *Ergonomics*, 48(5):436–463, 2005.

[54] Waldemar Karwowski. *International Encyclopedia of Ergonomics and Human Factors-3 Volume Set*. CRC Press, 2006.

[55] Waldemar Karwowski, Wlodzimierz Siemionow, and Krystyna Gielo-Perczak. Physical neuroergonomics: The human brain in control of physical work activities. *Theoretical Issues in Ergonomics Science*, 4(1-2):175–199, 2003.

[56] Anis Koubâa, Azza Allouch, Maram Alajlan, Yasir Javed, Abdelfettah Belghith, and Mohamed Khalgui. Micro air vehicle link (mavlink) in a nutshell: A survey. *IEEE Access*, 7:87658–87680, 2019.

[57] David Lee. Blackfly is latest attempt at flying car, 2018.

[58] Chin-Teng Lin, Chun-Hsiang Chuang, Chih-Sheng Huang, Shu-Fang Tsai, Shao-Wei Lu, Yen-Hsuan Chen, and Li-Wei Ko. Wireless and wearable eeg system for evaluating driver vigilance. *IEEE Transactions on biomedical circuits and systems*, 8(2):165–176, 2014.

[59] Eric Mack. Blackfly wants to be the flying car you can finally buy next year, and for cheap, 2018.

[60] William S Marras and Waldemar Karwowski. *Fundamentals and assessment tools for occupational ergonomics*. Crc Press, 2006.

[61] MasterPixel3D. Fantastic city generator, 2022.

[62] Andreas Meinel, Juan Sebastian Castaño-Candamil, Sven Dähne, Janine Reis, and Michael Tangermann. Eeg band power predicts single-trial reaction time in a hand motor task. In *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 182–185. IEEE, 2015.

[63] Peter J Mikulka, Mark W Scerbo, and Frederick G Freeman. Effects of a biocybernetic system on vigilance performance. *Human Factors*, 44(4):654–664, 2002.

[64] Noman Naseer, Hasan Ayaz, and Frederic Dehais. Portable and wearable brain technologies for neuroenhancement and neurorehabilitation, 2018.

[65] Attila Nemes. Tools for efficient soft computing modelling and feasible optimal control of complex dynamic systems. 2018.

[66] Attila Nemes and Gyula Mester. Energy efficient feasible autonomous multi-rotor unmanned aerial vehicles trajectories. In *Proceedings of the 4th International Scientific Conference on Advances in Mechanical Engineering, ISCAME*, pages 369–376, 2016.

[67] Thomas E Nygren. Psychometric properties of subjective workload measurement techniques: Implications for their use in the assessment of perceived mental workload. *Human factors*, 33(1):17–33, 1991.

[68] Oshkosh. Opener early access program moves blackfly closer to general release, 2023.

[69] Raja Parasuraman. Neuroergonomics: research and practice. *Theoretical issues in ergonomics science*, 4(1-2):5–20, 2003.

[70] Xiao-Wu Peng, Zhen-Cheng Xu, and Xiao-Chun Peng. Comparison of medical student's mental workload between vdt and paper-based reading. *Zhonghua lao Dong wei Sheng zhi ye Bing za zhi= Zhonghua Laodong Weisheng Zhiyebing Zazhi= Chinese Journal of Industrial Hygiene and Occupational Diseases*, 26(12):738–740, 2008.

[71] Florin Popescu, Siamac Fazli, Yakob Badower, Benjamin Blankertz, and Klaus-R Müller. Single trial classification of motor imagination using 6 dry eeg electrodes. *PloS one*, 2(7):e637, 2007.

[72] Charles Poussot-Vassal, Raphaëlle N Roy, Angela Bovo, Thibault Gateau, Frédéric Dehais, and Caroline Ponzoni Carvalho Chanel. A loewner-based approach for the approximation of engagement-related neurophysiological features. 2017.

[73] Lawrence J Prinzel, Frederick G Freeman, Mark W Scerbo, Peter J Mikulka, and Alan T Pope. A closed-loop system for examining psychophysiological measures for adaptive task allocation. *The International journal of aviation psychology*, 10(4):393–410, 2000.

[74] Lawrence J Prinzel III, Frederick G Freeman, Mark W Scerbo, Peter J Mikulka, and Alan T Pope. Effects of a psychophysiological system for adaptive automation on performance, workload, and the event-related potential p300 component. *Human factors*, 45(4):601–614, 2003.

[75] Fengxiang Qiao, Rong Zhang, and Lei Yu. using nasa-task load index to assess drivers' workload on freeway guide sign structures. In *ICCTP 2011: Towards Sustainable Transportation Systems*, pages 4342–4353. 2011.

[76] Mahjabeen Rahman, Waldemar Karwowski, Magdalena Fafrowicz, and Peter A Hancock. Neuroergonomics applications of electroencephalography in physical activities: a systematic review. *Frontiers in Human Neuroscience*, 13:182, 2019.

[77] Philipp Reiss, Daniel Dollinger, Christopher Schropp, David Löbl, and Florian Holzapfel. Multi crew coordination for remote piloted transition vtol. In *AIAA Scitech 2021 Forum*, page 1056, 2021.

[78] Aleksandar Rodić, Gyula Mester, and Ivan Stojković. Qualitative evaluation of flight controller performances for autonomous quadrotors. In *Intelligent Systems: Models and Applications: Revised and Selected Papers from the 9th IEEE International Symposium on Intelligent Systems and Informatics SISY 2011*, pages 115–134. Springer, 2013.

[79] Gavriel Salvendy. *Handbook of human factors and ergonomics*. John Wiley & Sons, 2012.

[80] Pranay Sinha, Piotr Esden-Tempski, Christopher A Forrette, Jeffrey K Gibboney, and Gregory M Horn. Versatile, modular, extensible vtol aerial platform with autonomous flight mode transitions. In *2012 IEEE aerospace conference*, pages 1–17. IEEE, 2012.

[81] Neville Anthony Stanton, Alan Hedge, Karel Brookhuis, Eduardo Salas, and Hal W Hendrick. *Handbook of human factors and ergonomics methods*. CRC press, 2004.

[82] RH Stone and G Clarke. The t-wing: a vtol uav for defense and civilian applications. *University of Sydney*, 2001.

[83] Kenneth I Swartz. Opener blackfly debuts at oshkosh. *Verti-flite*, 2018.

[84] Willi Tacke and Boric Marino. Flying pages europe sarl. *World Directory of Light Aviation 2015-16*, (1368–485X):206, 2015.

[85] Avinash Tandle, Nandini Jog, P D'cunha, and M Chheta. Classification of artefacts in eeg signal recordings and overview of removing techniques. *International Journal of Computer Applications*, 975:8887, 2015.

[86] Michal Teplan et al. Fundamentals of eeg measurement. *Measurement science review*, 2(2):1–11, 2002.

[87] Jan BF van Erp and Anne-Marie Brouwer. Neuroscience in ergonomics and human factors research and practice. *Tijdschrift voor Ergonomie jaargang*, 38(4), 2013.

[88] Graham Warwick. Volocopter goes bigger, faster, farther with voloconnect evtol, 2021.

[89] Chi Zhang, Hong Wang, and Rongrong Fu. Automated detection of driver fatigue based on entropy and complexity measures. *IEEE Transactions on Intelligent Transportation Systems*, 15(1):168–177, 2013.

[90] Jianhua Zhang, Zhong Yin, and Rubin Wang. Design of an adaptive human-machine system based on dynamical pattern recognition of cognitive task-load. *Frontiers in neuroscience*, 11:129, 2017.

[91] Yaoming Zhou, Haoran Zhao, and Yaolong Liu. An evaluative review of the vtol technologies for unmanned and manned aerial vehicles. *Computer Communications*, 149:356–369, 2020.

[92] Chen Ziyan and Liu Shiguo. China's self-driving car legislation study. *Computer Law & Security Review*, 41:105555, 2021.

[93] Glenn Zorpette. Finally, an evtol you can buy soonish: Opener's blackfly is the first of a radical new class of automated ultralight fliers. *IEEE Spectrum*, 60(1):32–57, 2023.

[94] Glenn Zorpette. I fly opener's blackfly evtol, 2023.

APPENDICES

Appendix A

SPSS Results

A.1 One-way ANOVA results for the three controller conditions

A.1.1 NASA-TLX Score

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
ControllerType	Sphericity Assumed	669.925	2	334.962	2.298	.110	.073	4.595	.448
	Greenhouse-Geisser	669.925	1.962	341.457	2.298	.111	.073	4.508	.444
	Huynh-Feldt	669.925	2.000	334.962	2.298	.110	.073	4.595	.448
	Lower-bound	669.925	1.000	669.925	2.298	.140	.073	2.298	.311
Error(ControllerType)	Sphericity Assumed	8455.590	58	145.786					
	Greenhouse-Geisser	8455.590	56.897	148.613					
	Huynh-Feldt	8455.590	58.000	145.786					
	Lower-bound	8455.590	29.000	291.572					

Table A.1: General Linear Model - Tests of Within-Subjects Effects

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
ControllerType	.981	.548	2	.760	.981	1.000	.500

Table A.2: General Linear Model - Mauchly's Test of Sphericity

A.1.2 Root Mean Square Error (RMSE)

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
ControllerType	Sphericity Assumed	.493	2	.247	1.217	.303	.040	2.435	.255
	Greenhouse-Geisser	.493	1.902	.259	1.217	.302	.040	2.316	.249
	Huynh-Feldt	.493	2.000	.247	1.217	.303	.040	2.435	.255
	Lower-bound	.493	1.000	.493	1.217	.279	.040	1.217	.187
Error(ControllerType)	Sphericity Assumed	11.750	58	.203					
	Greenhouse-Geisser	11.750	55.165	.213					
	Huynh-Feldt	11.750	58.000	.203					
	Lower-bound	11.750	29.000	.405					

Table A.3: General Linear Model - Tests of Within-Subjects Effects

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
ControllerType	.949	1.477	2	.478	.951	1.000	.500

Table A.4: General Linear Model - Mauchly's Test of Sphericity

A.1.3 Time Taken for Completion

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
ControllerType	Sphericity Assumed	25254.104	2	12627.052	2.078	.134	.067	4.156	.411
	Greenhouse-Geisser	25254.104	1.825	13841.089	2.078	.139	.067	3.791	.390
	Huynh-Feldt	25254.104	1.941	13013.402	2.078	.136	.067	4.033	.404
	Lower-bound	25254.104	1.000	25254.104	2.078	.160	.067	2.078	.286
Error(ControllerType)	Sphericity Assumed	352438.922	58	6076.533					
	Greenhouse-Geisser	352438.922	52.913	6660.766					
	Huynh-Feldt	352438.922	56.278	6262.457					
	Lower-bound	352438.922	29.000	12153.066					

Table A.5: General Linear Model - Tests of Within-Subjects Effects

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
ControllerType	.904	2.830	2	.243	.912	.970	.500

Table A.6: General Linear Model - Mauchly's Test of Sphericity

A.1.4 Time Spent Outside Track

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
ControllerType	Sphericity Assumed	192.031	2	96.015	.352	.704	.012	.705	.104
	Greenhouse-Geisser	192.031	1.903	100.890	.352	.694	.012	.671	.103
	Huynh-Feldt	192.031	2.000	96.015	.352	.704	.012	.705	.104
	Lower-bound	192.031	1.000	192.031	.352	.557	.012	.352	.089
Error(ControllerType)	Sphericity Assumed	15798.586	58	272.389					
	Greenhouse-Geisser	15798.586	55.198	286.219					
	Huynh-Feldt	15798.586	58.000	272.389					
	Lower-bound	15798.586	29.000	544.779					

Table A.7: General Linear Model - Tests of Within-Subjects Effects

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
ControllerType	.949	1.459	2	.482	.952	1.000	.500

Table A.8: General Linear Model - Mauchly's Test of Sphericity

A.1.5 Number of Wrong Entries

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
ControllerType	Sphericity Assumed	2.067	2	1.033	1.959	.150	.063	3.917	.390
	Greenhouse-Geisser	2.067	1.739	1.188	1.959	.157	.063	3.407	.361
	Huynh-Feldt	2.067	1.841	1.123	1.959	.154	.063	3.606	.372
	Lower-bound	2.067	1.000	2.067	1.959	.172	.063	1.959	.273
Error(ControllerType)	Sphericity Assumed	30.600	58	.528					
	Greenhouse-Geisser	30.600	50.443	.607					
	Huynh-Feldt	30.600	53.385	.573					
	Lower-bound	30.600	29.000	1.055					

Table A.9: General Linear Model - Tests of Within-Subjects Effects

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
ControllerType	.850	4.544	2	.103	.870	.920	.500

Table A.10: General Linear Model - Mauchly's Test of Sphericity

A.2 Paired T-Tests for with-wind and without-wind conditions of the racing wheel controller

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	NASA_TLX_1_RW - NASA_TLX_4_RW_Wind	-26.010	14.230	3.034	-32.319	-19.701	-8.573	21	.000
Pair 2	TimeTaken_1_RW - TimeTaken_4_RW_Wind	-6.38900	123.99119	26.43501	-61.36361	48.58561	-.242	21	.811
Pair 3	RMSE_1_RW - RMSE_4_RW_Wind	-.76320	.79398	.16928	-1.11523	-.41117	-4.509	21	.000
Pair 4	TimeOutsideTrack_1_RW - TimeOutsideTrack_4_RW_Wind	-8.13309	24.03603	5.12450	-18.79007	2.52389	-1.587	21	.127

Table A.11: T-Test - Paired Samples Test

		N	Correlation	Sig.
Pair 1	NASA_TLX_1_RW & NASA_TLX_4_RW_Wind	22	.504	.017
Pair 2	TimeTaken_1_RW & TimeTaken_4_RW_Wind	22	.521	.013
Pair 3	RMSE_1_RW & RMSE_4_RW_Wind	22	.460	.031
Pair 4	TimeOutsideTrack_1_RW & TimeOutsideTrack_4_RW_Wind	22	-.131	.561

Table A.12: T-Test - Paired Samples Correlations

A.3 EEG: Non-Parametric test results for the three controllers and the baseline

A.3.1 Delta

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
RW_Delta_AF	30	.48093467	.066902941	.351252	.618282
DRC_Delta_AF	30	.45536854	.082814159	.302710	.654519
Joystick_Delta_AF	30	.46164961	.060946528	.351177	.577112
Baseline_Delta_AF	30	.49017230	.068890858	.333548	.626292

Friedman Test

Ranks

	Mean Rank
RW_Delta_AF	2.80
DRC_Delta_AF	2.17
Joystick_Delta_AF	2.40
Baseline_Delta_AF	2.63

Test Statistics[a]

N	30
Chi-Square	4.120
df	3
Asymp. Sig.	.249

a. Friedman Test

Figure A.1: Friedman Test for K-Related Samples

A.3.2 Theta

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
RW_Theta_AF	30	.41758837	.070208725	.266214	.588253
DRC_Theta_AF	30	.42738904	.070938972	.283181	.616813
Joystick_Theta_AF	30	.42461031	.077623678	.278691	.589198
Baseline_Theta_AF	30	.41818297	.079363713	.289887	.586828

Friedman Test

Ranks

	Mean Rank
RW_Theta_AF	2.40
DRC_Theta_AF	2.50
Joystick_Theta_AF	2.57
Baseline_Theta_AF	2.53

Test Statistics[a]

N	30
Chi-Square	.280
df	3
Asymp. Sig.	.964

a. Friedman Test

Figure A.2: Friedman Test for K-Related Samples

A.3.3 Alpha

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
RW_Alpha_AF	30	.46598254	.097300830	.237869	.671701
DRC_Alpha_AF	30	.45505970	.064866505	.242814	.549594
Joystick_Alpha_AF	30	.46770128	.055828107	.316876	.562142
Baseline_Alpha_AF	30	.47207915	.080820808	.335776	.612720

Friedman Test

Ranks

	Mean Rank
RW_Alpha_AF	2.43
DRC_Alpha_AF	2.50
Joystick_Alpha_AF	2.47
Baseline_Alpha_AF	2.60

Test Statistics[a]

N	30
Chi-Square	.280
df	3
Asymp. Sig.	.964

a. Friedman Test

Figure A.3: Friedman Test for K-Related Samples

A.3.4 Beta

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
RW_Beta_AF	30	.48107594	.121478242	.294357	.721938
DRC_Beta_AF	30	.475762894	.1375258659	.2021164	.7717192
Joystick_Beta_AF	30	.49082752	.125115224	.235442	.703099
Baseline_Beta_AF	30	.40695009	.089501575	.252877	.582832

Friedman Test

Ranks

	Mean Rank
RW_Beta_AF	2.73
DRC_Beta_AF	2.47
Joystick_Beta_AF	2.87
Baseline_Beta_AF	1.93

Test Statistics[a]

N	30
Chi-Square	9.200
df	3
Asymp. Sig.	.027

a. Friedman Test

Figure A.4: Friedman Test for K-Related Samples

Test Statistics[a]

	Baseline_Beta_AF - RW_Beta_AF	Baseline_Beta_AF - DRC_Beta_AF	Baseline_Beta_AF - Joystick_Beta_AF
Z	-2.088[b]	-2.293[b]	-2.396[b]
Asymp. Sig. (2-tailed)	.037	.022	.017

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

Figure A.5: Wilcoxon Test

A.3.5 Gamma

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
RW_Gamma_AF	30	.48063193	.123748850	.309462	.753865
DRC_Gamma_AF	30	.45302085	.144370905	.218116	.774472
Joystick_Gamma_AF	30	.46915349	.133928452	.198531	.729893
Baseline_Gamma_AF	30	.39451949	.088055884	.223615	.573537

Friedman Test

Ranks

	Mean Rank
RW_Gamma_AF	2.83
DRC_Gamma_AF	2.60
Joystick_Gamma_AF	2.53
Baseline_Gamma_AF	2.03

Test Statistics[a]

N	30
Chi-Square	6.120
df	3
Asymp. Sig.	.106

a. Friedman Test

Figure A.6: Friedman Test for K-Related Samples

A.3.6 Engagement Index

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
RW_Engagement	30	.54308788	.103507611	.394570	.722446
DRC_Engagement	30	.54049323	.150245438	.235868	.825986
Joystick_Engagement	30	.55885246	.161521259	.279853	.885683
Baseline_Enagement	30	.46007850	.093292137	.302815	.677881

Friedman Test

Ranks

	Mean Rank
RW_Engagement	2.80
DRC_Engagement	2.57
Joystick_Engagement	2.73
Baseline_Enagement	1.90

Test Statistics[a]

N	30
Chi-Square	9.160
df	3
Asymp. Sig.	.027

a. Friedman Test

Figure A.7: Friedman Test for K-Related Samples

Test Statistics[a]

	Baseline_Engagement - RW_Engagement	Baseline_Engagement - DRC_Engagement	Baseline_Engagement - Joystick_Engagement
Z	-2.787[b]	-2.849[b]	-2.602[b]
Asymp. Sig. (2-tailed)	.005	.004	.009

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

Figure A.8: Wilcoxon Test

A.4 EEG: Non-Parametric test results for the with-wind and without-wind racing wheel controller conditions

A.4.1 Delta

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
RW_Delta_AF	22	.48838180	.072715257	.351252	.618282
Wind_Delta_AF	22	.49249773	.064771651	.373786	.634119
Baseline_Delta_AF	22	.51047379	.059121603	.379718	.626292

Friedman Test

Ranks

	Mean Rank
RW_Delta_AF	2.00
Wind_Delta_AF	1.95
Baseline_Delta_AF	2.05

Test Statistics[a]

N	22
Chi-Square	.091
df	2
Asymp. Sig.	.956

a. Friedman Test

Figure A.9: Friedman Test for K-Related Samples

A.4.2 Theta

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
RW_Theta_AF	22	.41262352	.066237338	.266214	.516662
Wind_Theta_AF	22	.43630936	.074855911	.336198	.579346
Baseline_Theta_AF	22	.41938214	.074240429	.289887	.581104

Friedman Test

Ranks

	Mean Rank
RW_Theta_AF	1.95
Wind_Theta_AF	2.09
Baseline_Theta_AF	1.95

Test Statistics[a]

N	22
Chi-Square	.273
df	2
Asymp. Sig.	.873

a. Friedman Test

Figure A.10: Friedman Test for K-Related Samples

A.4.3 Alpha

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
RW_Alpha_AF	22	.45885822	.097867546	.237869	.661918
Wind_Alpha_AF	22	.48232976	.069453332	.369245	.619654
Baseline_Alpha_AF	22	.47482989	.083023595	.335776	.612720

Friedman Test

Ranks

	Mean Rank
RW_Alpha_AF	1.77
Wind_Alpha_AF	2.23
Baseline_Alpha_AF	2.00

Test Statistics[a]

N	22
Chi-Square	2.273
df	2
Asymp. Sig.	.321

a. Friedman Test

Figure A.11: Friedman Test for K-Related Samples

A.4.4 Beta

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
RW_Beta_AF	22	.46888931	.096993774	.294357	.675260
Wind_Beta_AF	22	.43678872	.139900421	.233138	.784518
Baseline_Beta_AF	22	.39450267	.085786318	.252877	.582832

Friedman Test

Ranks

	Mean Rank
RW_Beta_AF	2.41
Wind_Beta_AF	1.86
Baseline_Beta_AF	1.73

Test Statistics[a]

N	22
Chi-Square	5.727
df	2
Asymp. Sig.	.057

a. Friedman Test

Figure A.12: Friedman Test for K-Related Samples

Test Statistics[a]

	Baseline_Beta_AF - RW_Beta_AF	Baseline_Beta_AF - Wind_Beta_AF
Z	-2.088[b]	-.601[b]
Asymp. Sig. (2-tailed)	.037	.548

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

Figure A.13: Wilcoxon Test

A.4.5 Gamma

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
RW_Gamma_AF	22	.47547724	.112697057	.309462	.753865
Wind_Gamma_AF	22	.44060791	.142512857	.262220	.755105
Baseline_Gamma_AF	22	.38720844	.086675143	.223615	.573409

Friedman Test

Ranks

	Mean Rank
RW_Gamma_AF	2.36
Wind_Gamma_AF	1.95
Baseline_Gamma_AF	1.68

Test Statistics[a]

N	22
Chi-Square	5.182
df	2
Asymp. Sig.	.075

a. Friedman Test

Figure A.14: Friedman Test for K-Related Samples

Test Statistics[a]

	Baseline_Gamma_AF - RW_Gamma_AF	Baseline_Gamma_AF - Wind_Gamma_AF
Z	-2.520[b]	-1.023[b]
Asymp. Sig. (2-tailed)	.012	.306

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

Figure A.15: Wilcoxon Test

A.4.6 Engagement Index

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
RW_Engagement	22	.54074564	.097923237	.394570	.722446
Wind_Engagement	22	.47352182	.138493205	.297580	.893080
Baseline_Enagement	22	.44246098	.077876788	.302815	.628411

Friedman Test

Ranks

	Mean Rank
RW_Engagement	2.50
Wind_Engagement	1.64
Baseline_Enagement	1.86

Test Statistics[a]

N	22
Chi-Square	8.818
df	2
Asymp. Sig.	.012

a. Friedman Test

Figure A.16: Friedman Test for K-Related Samples

Test Statistics[a]

	Baseline_Engagement - RW_Engagement	Baseline_Engagement - Wind_Engagement
Z	-2.787[b]	-.179[b]
Asymp. Sig. (2-tailed)	.005	.858

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

Figure A.17: Wilcoxon Test

A.5 Cybersickness

A.5.1 Three controller conditions

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Controller	Sphericity Assumed	.467	2	.233	.396	.675
	Greenhouse-Geisser	.467	1.482	.315	.396	.614
	Huynh-Feldt	.467	1.543	.302	.396	.622
	Lower-bound	.467	1.000	.467	.396	.534
Error(Controller)	Sphericity Assumed	34.200	58	.590		
	Greenhouse-Geisser	34.200	42.980	.796		
	Huynh-Feldt	34.200	44.749	.764		
	Lower-bound	34.200	29.000	1.179		

Table A.13: General Linear Model - Tests of Within-Subjects Effects

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Controller	.651	12.039	2	.002	.741	.772	.500

Table A.14: General Linear Model - Mauchly's Test of Sphericity

A.5.2 No-wind and with-wind racing wheel conditions

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	CS_1_RW - CS_4_RW_Wind	-1.00000	1.82574	.38925	-1.80949	-.19051	-2.569	21	.018

Table A.15: T-Test - Paired Samples Test

		N	Correlation	Sig.
Pair 1	CS_1_RW & CS_4_RW_Wind	22	.598	.003

Table A.16: T-Test - Paired Samples Correlations

Appendix B

Poster

Figure B.1: Poster designed for user study invitation