What is the Usability Perception of a VR Flight Training System?

By

Luis Alejandro Morales Trujillo

Supervisor: Dr. Brett Wilkinson

Thesis
Submitted to the College of Science and Engineering in partial fulfilment of the requirements for the degree of Master of Science (Computer Science) at Flinders University – Adelaide, Australia.

Master of Science (Computer Science)

College of Science and Engineering
October 2021

TABLE OF CONTENTS

ABSTRACT

Human-Computer Interactions (HCIs) have been evolving from the early days of using the command line or low-level programming language to communicate with the device to the widespread usage of smartphones in today's world. Moreover, relatively novel technologies such as Virtual Reality hardware and software are emerging as the new trend of HCI. This research describes the process and techniques in which the project Virtual Reality Flight Training Experience (VRFTE) system was evaluated from a usability perspective. The findings in the literature contributed to defining a strategy of combining some of the most widely-used usability methodologies in desktop, web, mobile and serious game applications that are often used for usability testing within the VR industry. Due to the lack of a well-defined methodology aimed for VR products, the application of traditional heuristics evaluation, informal usability surveys and observation techniques are combined or adjusted to the requirements and nature of the system.

The results obtained from the usability testings and the analysis revealed that their previous experiences significantly influence the user perception of usability in VR and flight training scenarios. Thus, subjective results are required to be compared with the obtained data from the system logs, as well as the processed analysis of the feedback and reviews from the users. Overall, the perception of the VRFTE system is the positive spectrum, and the obtained feedback implies that some improvements and fixes of the interactions could increase the efficiency, effectiveness and satisfaction of the system’s usage.

DECLARATION

I certify that this thesis does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text.

Signed..

Date...

ACKNOWLEDGEMENTS

Through the challenges that the last year has imparted, I have experienced a constant lack of sense of time, in which the days have come and gone without my notice. This created a sense of urgency and pressure that I had not felt in a long time. However, now that I see the light at the end of the tunnel, a nostalgic feeling runs across my mind, realising the closure of another chapter in my life. One thing that I am confident about is that this last year has given me the tools and knowledge to make a difference in my profession and future endeavours; I hope to become a wiser person who employs the hardships as the fuel to go beyond my limits.

My profound gratitude to my supervisor, Dr Brett Wilkinson, who supported and motivated me throughout this project regardless of his busy schedule and occupations. His contagious positivity and passion for the industry could not have been more inspirational to keep on track with the proposed plan. I would also like to thank all the students from Flinders University and the other institutions that participated in this research.

My infinite thanks to my parents, Alejandro and Marcela, who, despite the distance, were constantly encouraging me to continue against the odds. And lastly, I would like to thank my partner Maria and my friends, who understood me in the stressful moment and encouraged me in the moment of despair.

Luis Morales

October 2021

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE: INTRODUCTION

Digital driven technologies seem to be at the core of all human interactions in today's world. The boom of smartphones with over six billion active devices, allegedly more than half the population on the Earth have used a smartphone (O'Dea, 2021), is driving this technology focused communication. With that premise, it is reasonable to consider the future and the type of novel human-computer interactions (HCI) to come. As discussed during an interview with Mark Zuckerberg by Stein (2021), big tech companies such as Facebook intend to transition and evolve their products; with the challenges presented during the COVID Pandemic in 2020 and 2021, Virtual Reality (VR) and Augmented Reality (AR) are being developed and implemented to widen the adoption of these technologies, pursuing the engagement of the public in a virtual world. Although VR growth has been predominantly in the entertainment industry, particularly in the development of videogames, VR systems are also gaining popularity and reception in other fields such as health, construction, psychology, medical mobility treatments and aviation. Thus, it has permitted new problem-solving approaches and alternatives for training operations that were usually dealt with using traditional learning pathways such as reading manuals or actual on-field training. For example, some research argues the effectiveness of using VR serious game that provides task-driven scenarios for aircraft evacuation training of crew and passengers (Feng, González, Amor, Lovreglio, & Cabrera-Guerrero, 2018). Furthermore, including the concepts of serious games (games that are not designed solely for entertainment purposes) in VR experiences increases the possible application of these systems.

VR research opportunities are inspiring projects that, by innovating the communication supported by computer-user interactions, improve knowledge retention and skill acquisition against traditional methods of education (Chittaro & Buttussi, 2015). In contrast, some might consider the complexity involved in designing and developing a VR system a burden that could affect the fundamental simplicity implied in the early serious games (Manuel, José, Manuel, Iván, & Baltasar, 2019). In spite of the risk, VR experiences could be seen in the success of training the operation of plant and heavy machinery such as firefighters' trucks, military equipment, aircraft maintenance and ambulances (de Armas, Tori, & Netto, 2020). Therefore, considering the possibilities of the future, in which a VR training system could enhance users' safety operations and procedures in their work environment, is promising.

A review of the current literature was conducted to understand VR systems, their applications through recent years, products developed for training programs within various industries and the most widely used usability evaluations techniques for software, serious games and VR, to define a suitable usability testing approach for the Virtual Reality Flight Training Experience (VRFTE) system designed and developed for this project. Coverage of the motivation, the design, development cycle and prototyping of the VRFTE project intends to provide the background for the analysis of the findings in the usability evaluations carried out on the system with the participation of volunteer users. VRFTE is a system developed in Unreal Engine 4.26 game engine published by Epic Games; the fundamental interactions are programmed in the C++ programming language, using the inclusion of Blueprints for animated components. A custom-made 3D model of the Cessna 172SP aircraft was designed and sculpted with Autodesk's modelling and rendering software 3DS Max. Additionally, texture creation and video editing were completed with Adobe Creative Solutions (Photoshop 2021, Premier Pro 2022). From the mockup design to prototype release, the overall development cycle took approximately four months.

Consequently, with the increasing interest in VR, the number of Head-Mounted Displays (HMDs) built that are used for virtual reality immersion has exploded in popularity and as shown in Figure 1, is expected to exceed 16 million units by the end of 2021. The chosen HMD for implementing the VRFTE system is the Oculus Quest 2. Primary reasons include the hardware's affordability, compatibility with UE4 and portability.

Removed due to copyright restriction

Figure 1 - VR HMD unit sales worldwide from 2019-2024 (Statista,2021)

To understand the usability evaluation process within this research, according to ISO 9241 reviewed by Bevan, Carter, and Harker (2015), usability is determined by the degree to which an HCI product can be used with effectiveness, efficiency and satisfaction. This research aims to answer the proposed question of "What is the Usability Perception of a VR Flight Training System?", by combining usability evaluation methodologies such as heuristics evaluation, qualitative review, cognitive walkthrough and controlled experiments. From applying heuristics evaluation surveys to the collection of data within each recorded session per user and the use of observation and interview techniques, the individualised analysis of each user perception is graphically described in this paper.

This research presents the usability evaluation of a VR Flight Training System with the background of the Virtual Reality state of technology. Considering that standards of usability testing of VR serious games are not yet established (Yanez-Gomez, Cascado-Caballero, & Sevillano, 2017), a hybrid approach has been adopted, blending traditional evaluation methodologies and techniques in Human-Computer Interaction (HCI), combined with novel techniques that are focused on answering how to appraise the player's presence linking playability, reception, and satisfaction in the process of task-driven scenarios.

This study is driven by the research questions (RQ) below, which are the foundation of the usability evaluation of the VRFTE application.

- RQ1: How does the VRFTE System in the designed tasks motivate the user to practice them multiple times in the session?
- RQ2: What factors provide the best playability experience for the user?
- RQ3: How upgradable is the system according to the feedback of the user?
- RQ4: What are the benefits to the user by giving them feedback on their performance?
- RQ5: What is the reception of the participant and the likelihood of using the VRFTE system again?

The remainder of this paper is divided into seven chapters. Chapter 2 provides a review of the literature that discusses the fundamentals for analysis and evaluation. Chapter 3 describes the software development cycle adopted for the built VR system. Chapter 4 compiles the proposed methodology used for tackling the usability evaluation process of the VR product. Chapter 5 presents the overall qualitative and quantitative results obtained throughout the two usability testing techniques. Chapter 6 discusses the patterns and analyses the results within the participants' sample. Chapter 7 outlines the author's conclusion and future research possibilities in usability of VR applications. Lastly, at the end of the document research supported material is attached in the appendix.

CHAPTER TWO: LITERATURE REVIEW

This chapter details some of the research conducted in the field of usability evaluation for VR products with an emphasis on serious games applications and is structured into four sections. The first section reviews the evolution of technology usability, followed by the fundamental principles of usability. The next section discusses virtual reality in the serious games market and finally associates these technologies with studies in applying a mixed methodology of usability evaluation for human-computer interaction (HCI) with VR head-mounted displays (HMD). Although the arguments of this paper aim to evaluate the usability of a VR training system, the initial query of the literature generated a quantity of papers that were beyond the scope of this thesis. Hence, a clear delineation of the search strategy, databases used, and selection of papers, was defined as follows:

- The search terms for the literature were defined within the concepts of the research title and expanded to capture a broader range of research related to usability, serious games and virtual reality. For the search strings S1 to S4 (see below), the literature's hierarchy and constraints were considered, such as the publications date not before 2016 and type of paper as peer-review journal articles and conferences papers. Nonetheless, in some of the literature, older publications were constantly referenced by other publications, which made those seminal papers of interest for this review.
  S1: “Usability” **AND** “Evaluation” **AND** ”Virtual Reality” **OR** “Serious Games”
  S2: “Usability Methodologies” **AND** “Serious Games” **OR** “Usability Testing”
  S3: “Usability Procedures” **OR** “Usability Principles” **OR** “Usability Practices”
  S4: “Usability” **AND** “Software Design Cycle” **OR** “Game Development Cycle”
- While the Flinders Library FindIt search engine provided clear filters and papers linked to the various databases it referenced, the additional use of Google Scholar and databases such as the ACM Digital Library and the IEEE Xplore were essential for correlating studies within the search terms.

2.1. Background

Throughout the years of humankind discoveries, inventions and technological devices, their success or failure has been heavily influenced by how practical and easy to use they were. For instance, the “wheel” invented by the Mesopotamians approximately 3500 B.C has evolved from its original design of an axis with a cylinder wood shape (used as a tool for transportation of goods) to incorporating modern materials in the frame, built for more complex equipment or systems (e.g. Vehicles) (Faiella, 2006). Nonetheless, the current wheel, in its essence, is the same successful idea that revolutionised transportation, and has become integrated into everyday activities. With the same principle, computational devices are becoming part of our daily lives to the degree where it might be something that every person requires to interact with at least once in their lifetime. Despite the conspicuous usage of computers, it is relatively frequent to find problems in their usability on both hardware and software functionalities, resulting in millions of products that finish in failure.

With regards to computers, the importance of user interfaces (UI) and user experience (UX) has not been a novelty. As described by Nielsen (1994), the user opinion is connected to the purpose, functionality and development of a software application that directly affects its accomplishments in the competitive market of information technologies. From an initial premise of developing for experts, where usability was not an important role, to today's trend of placing the user feedback as a key factor in finding a balance of usable built solutions. The reality is that there is no reason to create desktop, web, mobile or VR applications if they will never be used or accepted by the end-user, this lack of use by the users will result in less profits or eminent losses.

Therefore, in this research, the causes, studies, and concepts are associated with usability from its beginnings to the growing fields and industries implementing its different methodologies. Some research, in particular, explain the Heuristic Evaluation (HE) method, referencing Nielsen's first conceptualisation of usability and comprising components required in any system to be considered usable (Abulfaraj & Steele, 2020), such as learnability, efficiency, memorability, errors, and satisfaction. Abulfaraj and Steele also conducted interviews to evaluate the evaluators' understanding of heuristics fundamentals.

An area of focus for the literature was to analyse the novel usage of VR in the aviation industry. From the author's personal experience with updated documentation, procedures, and training protocols; the research was conducted based on non-pilot, enthusiast and some experienced users who voluntarily participated in the testing session of the VRFTE. Therefore, this thesis' core motivation originates from the difficulties that pilots faced throughout their training process, from theoretical learning based on aircraft manufacturers manuals to muscle memory acquisition of normal, abnormal, and emergency standard procedures. The aircraft chosen as the environment platform is the Cessna T-41D (Military version) or C172D "Skyhawk", which is the most produced light aeroplane in history and the most common used for basic training programs in aviation schools around the globe (Buscombe, 2016).

One of the main terminologies commonly known by pilots in their flight training is "flows". A flow is a series of steps that the pilot performs, strictly following a checklist of procedures detailed in the aircraft operating manual. In most cases, these flows are essential actions that the awareness, muscle memory, proficiency, and cross-check with the official documentation determine an aircraft's proper operation. In an emergency situation, these procedures are of critical importance to prevent an accident (Company, 2008). According to the research by Kharoufah, Murray, Baxter, and Wild (2018), approximately 75% of all aircraft accidents and incidents are due to human factors, some involved the lack of training or awareness of the crew, resulting in fatal consequences such as the Air France 447 of the 1st of June 2009, in which the design of the flight controls lack of haptic feedback from both sides of the cockpit, and the reduced knowledge of the inexperienced second officer in an abnormal operation contributed to the chain of errors.

While most of the flight experience is acquired by actual missions in the designated aircraft, military institutions and airline companies invest heavily in full flight simulators (FFS) (Figure 2) training programs. Administration (2021) explains that an approved FFS by international and local regulating organisations such as the United States Federal Aviation Administration (FAA) and the Australian Civil Aviation Safety Authority (CASA), is considered full size replica of a specific make, model, series and type of flight deck, which has the capacity to represent ground and flight operations with outside deck visual and motion feedback of at least 3 axis of movement. This system requires to have the technological equipment and software built into the simulators, categorised into four levels of complexity, mechanical axis, and simulation capabilities. As shown in Table 1 in Australia, approved and certified FFS are level designated (Authority, 2016). In spite of the evident advantages of these training systems, the cost is too high for regular training and independent pilots, resulting in an increasing trend of consumer simulator games, such as Microsoft Flight Simulator 2020 or X-Plane 11, and the growing development of VR applications with HMDs as an alternative for flight training, helping to are narrow the gap against FFS.

Table 1 - FFS Classification

FFS Type	Description	Motion Axis	Add-ons
A	Early visual systems, old version, barely use in the market and usally for older aircraft models.	3	Night visuals
B	Scarcely use for older aircraft models, providing only up to 80 percent of initial training and 100 percent if its recurrent training.	3	Night visuals, ground handling simulation
C	Highly use for private jets and operators, all instruments are required in fidelity, landing capabilities and traffic patterns.	6	night and dusk visuals, dynamic control loading, higher fidelity
D	Most use for commercial aircrafts and airline training programs, simulates everything that could be presented in the operation of an aeroplane, it also includes data and performance tolerance and analysis features.	6	night and dusk visuals, dynamic control loading, highest fidelity

Removed due to copyright restriction

Figure 2 - Cessna Citation FFS Level D (International, 2020)

2.2. Usability Fundamentals

Traditional methodologies and techniques in usability evaluation for HCI have been used to develop desktop, web, mobile applications and videogames for decades. Moreover, the history of such applications has shown us that specific patterns of human behaviours are repeated; these patterns are simplified as principles aiming to be applied to answer what makes computational user interfaces easy to use, learn and experience. The most fundamental of these principles are the ones briefly mentioned in the previous section; they constitute the basis of the Usability Heuristics method. Another fundamental usability principle is the application of Nielsen's revised ten usability heuristics. It should be noted that Nielsen’s usability heuristics are not perfect or absolute in guaranteeing a positive perception from a user perspective. However, some researchers and evaluators argue that these ten principles are the commandments to be followed in usable HCI (Marcus & Rosenzweig, 2020). Despite the similar premise with the ten commandments of Christianity, usability heuristics are considered by some to be a broad design guideline instead of mandates. Researchers aiming to evaluate IT applications combining these principles with other methodologies (Abulfaraj & Steele, 2020) tend for innovation and newer technologies that are increasingly becoming popular and adopted by the public.

To support the research presented within the following chapters, it is important to understand the principles of Nielsen's HE that are included in numerous studies (Sauro & Lewis, 2016), as follow:

I. Visibility of System Status – Open, continuous, and timed communication of the system with the user, giving the needed information that helps the user to understand and operate the application.
II. Match Between the System and the Real World – The application and interface should be presented logically and in a language easy for the user to comprehend and recognise. In some cases, a skeuomorphic design approach, utilising metaphors and affordance, could be considered to facilitate familiarity with a physical experience to a digital interaction (Spiliotopoulos, Rigou, & Sirmakessis, 2018).
III. User Control and Freedom – Provide the user with the ability to undo and redo, understanding that sometimes human actions are unpredictable or, a user may change their change of mind. This concept should be clearly visible to the user and not a hidden feature that would not be used.
IV. Consistency and Standards – Predictable human interactions to improve user's learnability, considering internal consistency that implies a recognisable feature of an established application; an example would be the gestures in iOS devices. Consideration also needs to be given to external consistency/standards, which are interactions that are constantly used in similar applications, for instance, the from and destination location fields of most navigational apps.
V. Error Prevention – Straightforward design for error prevention, limiting the interactions of the user to the key features necessary at that point of the application, in case they might click or interact in an unintended way.
VI. Recognition vs Recall in UIs – By using cues in the system, it is possible to make memory retrieval easier for the user. Hence recognition will be a better practice to facilitate placing the user in a recognisable context.
VII. Flexibility and Efficiency of Use – A highly usable application should be flexible enough to be friendly to first users and efficient for advanced or experienced users by providing options for how an interaction could be made. An example of this could be the word processing program that is being used for this thesis, where it is possible to copy a piece of text by interacting with the mouse right-click in the required text, by a shortcut such as CTRL+C or by the menu bar.
VIII. Aesthetic and Minimalist Design – Providing the right amount of visibility of the relevant and functional information to the user, aiming for signal-to-noise ratio, determining the amount of information given to the user simply and concisely. Therefore, a system should not have elements in its interfaces to beautify the UI if these do not support the content or primary focus of the system.
IX. Help Users Recognize, Diagnose and Recover from Errors – Perfect systems are not true, and users will encounter errors or make unintentional actions. Therefore, by providing the correct

information to the user, their experience could be enhanced, giving them confidence in the application. Informing them about the error, explaining what went wrong and giving them a solution or pathway to solving it, these supportive interactions would make the system more usable for the end-user (NNgroup, 2019).

X. Help and Documentation – Supporting the previous principle, this one aims to prevent errors by supporting the user with the information required when they might get stuck with something. Thus, it is crucial to evaluate the system in three scenarios; the first involves making it easy to search for help either by documentation or by the application contact support, the next stage requires that the documentation is focused on the user's tasks, and finally guaranteeing that there are easy to follow and concrete steps to be carried out by the user within the documentation.

As long as the design and development of a system are defined following the HE principles and Nielsen's Heuristics, the ability to increase the usability perception of the users is somewhat straightforward. Nevertheless, a key factor to consider is the inclusion of usability testing by volunteer participants through the different cycles or versions of the system. On the one hand, having a transparent development process, timeline and milestones, allow for systematically program usability in advance to the release or launch of the system. Still, an understanding of the development and the objective of the testing are as crucial as the evaluation itself. Therefore, defining limitations according to the budget, such as the sample of participants, the number of iterations, and the testing review, is key to preventing the lack of usability in the final product and its success (Dunn & Hayes, 2020). Also, selecting between a qualitative or quantitative evaluation permits identifying the type of observations, surveys and quantity of participants needed for the test. As explained on NNgroup (2018), in a qualitative assessment, the argument of "Less is more" applies in a pattern that corresponds to better observation of the user's interactions, targeting flaws in the system and extracting visual cues from the users; on the contrary with a quantitative approach, a bigger sample size of participants tends to be more favourable for the collection of data by surveys, written feedback and log records of the system, in order to be analysed with metrics in post-evaluation.

It is also worth noting that the evaluators' experience in usability fundamental principles and in conducting usability testing, greatly influence the methodology that should be used. For instance, if the researchers are starting in usability evaluation, but the design and development team have experience and are dynamic, the quantitative methods could provide the information required for a better system and prevent failures in the analysis. Otherwise, missing points could be present if qualitative methods are applied by an inexperienced evaluator, resulting in a lack of a usable system. However, an experienced evaluator could use the fast-pacing qualitative assessment to identify the flaws by observation and communication techniques to extract enough information from a smaller sample of participants (Geisen, 2017). In reality, there is no definitive methodology to tackle usability evaluation, but actually a variety of tools that researchers could use and combine to get the targeted results.

2.3. Virtual Reality and Serious Games

As the evaluated application of this research combines a serious game approach into a VR HMD system, it is needed to contemplate both concepts and the reasoning of what design decisions are made when designing for a usable training system.

Regardless of age, a big misconception is an assumption that VR is a novel and more expensive way to play video games. That statement could not be farther from the essence of Virtual Reality. Studies on virtual reality can be traced to the beginning of computers, where developers speculated on the applicability of their UIs with VR and even considered a wholly immersive environment (Bunnenberg, 2018). Consequently, with the growth of computer applications in the 1990s, the evolution of portable devices in the 2000s and the increased usage, affordance and accessibility to the internet generated innovative ways to see the VR applicability within an application. From collaborative platforms in an industrial setup to people's virtual interactions with other users and the boom of videogames (Eller, Bittner, Dombois, & Rüppel, 2018) have transformed that first recognition of VR to what it is today and what it will become in the future. In today's perception, VR is an application/software built from the combination of software and hardware. The user is immersed in a 3D environment that allows them to interact realistically with the objects and tasks created within the system. Petar and Sanja Maravić (2020) argue that implementing game design and development methodologies such as Agile or Extreme Programming produces the best results for effective application based on technology limitations. Although the limitations are holding some of the ideas for VR technologies, the constant advances in haptic feedback, eye tracking functionality in HMDs, and newer hardware that would allow higher levels of immersion, are reducing those limitations. Also, there is to be a consideration between game and non-game VR applications, which could determine why the author's opted for a serious game approach to the VRFTE system and the effects on the methodology and results of this study.

Firstly, if considerations within the design are made for a VR application to be as realistic as possible without any gamification, in that case, the 3D immersive environment would tend to be a straight replica of what the user could visually perceive in the real world. Palmas, Labode, Plecher, and Klinker (2019) explained that this strategy had been regularly used in industry-focused applications to support employees in their training of specific tasks, resulting in a mitigation of the higher cost of practising them in real-world scenarios Figure 3. Nonetheless, while effective in reproducing the environments, the lack of cues, instructions, challenges, and feedback impacts the training efficiency, usability, and results for the developed VR apps. In the end, even with a realistic VR environment, if the user does not have similar or better tools for interaction with the world elements in it then the notion of immersion could be lost. Thus, from the user perception, when there are no apparent benefits in the usage of VR from the traditional ways of doing things, why would they bother to stick or engage in it.

Removed due to copyright restriction

Figure 3 - Comparison between real-world setup and virtual training environment (Palmas, Labode, Plecher, & Klinker, 2019)

Secondly, when the orientation of the VR application design is on providing tasks, instructions, and challenges for an specific activity that does not involve entertainment; the trend is still to analyse it from a videogame perspective, following a game development approach, that depending on the end goal of the application, it could be considered a gamified product or a serious game. Consequently, for the previous example, if the realistic environment's requirements are made, including user cues to follow specific interactions, it will be in the spectrum of serious games. Fundamentally, VR systems have been built regardless of their level of immersion for the purpose of training workers/users in several domains such as health, defence, education, and transportation areas, many of which are considered a serious game (SG) (Azadegan & Riedel, 2012). Just like commercial videogames, SGs contemplate gameplay setups, user interactions, challenges and objectives. Despite the similarities with a gamified application, SGs are built from the ground-up, designed with an educational framework and pedagogical characteristics, which are concepts outside this thesis' scope (Menin, Torchelsen, & Nedel, 2018).

Lastly, the significance of immersion and the sense of presence within a VR application are directly connected to the data perceived by our senses. As Slater and Sanchez-Vives (2016) explained, the user perception requires the sensory inputs received in the brain and involves a more complex process of compiling expectations, prior experiences and beliefs, which creates a mental chain reaction in the sense of presence of the VR application. This level of immersion could vary from user to user, influencing the development of the application according to the targeted audience. Some studies analysed the use of HMDs, controls, audio cues and haptic feedback in systems that incorporated a more robust integration of hardware, and the findings are promising in terms of the improved effectiveness of the VR systems for the goal of user performance within the assigned tasks (Wu, Yu, & Gu, 2020). Moreover, as mentioned before, technology advances in hardware, processing power, software and affordance of HMD devices will put into a whole new era, the conceptualisation of user immersion and presence (Kilteni, Groten, & Slater, 2012). Research in this field will involve a growing user expectation, their perception and the consequences that it might create from an ethical and health perspective to realise the benefits that advances could bring.

2.4. Usability Methods and Techniques for Virtual Reality

One of the most common approaches to usability for VR systems is to observe it from the lessons learnt in game development. Some methodologies focus on visualizing the idea of the game and releasing teaser information to the public, to perceive the audience reception. The other will use evaluation methods such as recording and observing the testing phases, where participants interact with an early stage of the application; this allows for a redrafting the original design or a more unified vision of the end-product. In some cases, a larger sample size of users is tested on a prototype version of the game, commonly known as alpha and beta testing, providing a quantitative and survey type feedback that is intended for post-review for correction and version improvements (Hookham, Nesbitt, & Kay-Lambkin, 2016). Yanez-Gomez et al. (2017) explained that developing a new evaluation technique tailored for videogames is an approach that some development teams might see fit for their development capabilities.

In comparison, off-the-shelf techniques consider the differences between features in commercial software, this includes public available reviews, that are scored and detailed by selected individuals who could be considered similar to a test participant in a traditional HE methodology. However, this methodology does not consider the development process, and is mainly implemented in the production phase of the final product. But perhaps this approach applies in a post-release scenario and for future patches or updates. From another standpoint, adopting traditional evaluation techniques involving observation, documentation, and survey techniques makes it easier to shape results-focused despite the differences that could be seen on videogame hardware or a VR system for this research (Hookham et al., 2016). Subsequently, when developing, implementing, and testing a prototype of the product, there are two approaches, an evaluation by experts in the field, also known as Inspection Method, or evaluations involving users/participants, classified into Inquiry Methods and Testing Methods (Yanez-Gomez et al., 2017).

Various guidelines in heuristics and usability have been developed to help design game user experiences, from contemplating the HE principles to applying psychology techniques to extract and portray the player acceptance of the game interfaces (Hookham et al., 2016). In addition, the primary purpose of using these methodologies for the research presented in this thesis, is to emphasise usability itself, making the VRFTE application easy to interact with, learn from and engage with. Given the importance discovered in the literature, the evaluation process will focus on particular criteria such as: Learnability, Reliability, Efficiency, Utility and Memorability. However, as the research literature has suggested, the requirements are not limited to only these outlined factors. They could also be affected more subjectively, specifically user satisfaction or more objective variables such as the overall playability in-game. Nevertheless, the use of usability evaluation techniques in VRFTE's development phases could reduce the issues related to poor practices of usability implementations (Fernandez, Insfran, Abrahão, Carsí, & Montero, 2012).

Regarding a virtual reality system such as VRFTE, when considering the effects in the design and development of immersive components against the level of realism of the interactions with an aircraft. It was crucial to define a balance between the final product unavoidable limitations and the actual capabilities and goal of the system. The balance aims to elicit an emotional reaction from the user, to increase the immersive possibilities of the system and the presence recognition of the user within the VR environment. As described by Pallavicini, Pepe, and Minissi (2019), within their studies and evaluation of the usability between a VR immersive game and a non-immersive videogame, they have shown that the evaluation criteria of usability favoured playing in VR, linked to the user increased results in the level of satisfaction. While the same game was played with traditional controllers and an external screen, that experience score was slightly lower. However, their study argues that the current state of the technology would not inflict a significant difference of how a user would prefer VR from non-immersive experiences, which could mean that the future holds the last word of how that perception would be changed.

The majority of usability methodologies and evaluation techniques involve collecting and analysing information from each user or participant during a testing session. Hence, the related works in VR systems evaluation involve a combination of different methods. The literature review commonly found that the system usability scale (SUS) was included to acquire quantitative results for further analysis. The SUS is a simple questionnaire that catches the user's answer on a scale of positive to negative sense (e.g. Strongly Agree to Disagree Strongly), which should use a value score of 1-5 (Swan & Gabbard, 2014) as shown in Figure 4. A significant component in the evaluation was the number of tests and interactions divided by the different versions of the VR system (Miglani, Kidambi, & Mareguddi, 2020). Alternatively, a qualitative approach usually evaluates smaller projects or highly efficient design, development and testing teams. In addition, it involves communication with the user on a one-on-one basis, incorporating observation techniques on the user responses, actions and expectations; it also considers the body language, tone and formulating open questions, that allows guiding the participant to provide further information (Sagar & Saha, 2017).

	The System Usability Scale Standard Version		Strongly disagree				Strongly agree
			1	2	3	4	5
1	I think that I would like to use this system frequently.		O	O	O	O	O
2	I found the system unnecessarily complex.		O	O	O	O	O
3	I thought the system was easy to use.		O	O	O	O	O
4	I think that I would need the support of a technical person to be able to use this system.		O	O	O	O	O
5	I found the various functions in the system were well integrated.		O	O	O	O	O
6	I thought there was too much inconsistency in this system.		O	O	O	O	O
7	I would imagine that most people would learn to use this system very quickly.		O	O	O	O	O
8	I found the system very awkward to use.		O	O	O	O	O
9	I felt very confident using the system.		O	O	O	O	O
10	I needed to learn a lot of things before I could get going with this system.		O	O	O	O	O

Figure 4 - Standardised SUS (Sauro & Lewis, 2016)

As described in previous sections, usability evaluation and testing require a broad approach to what has been used in recent years in VR systems and what innovation could be brought to the field (Karre, Mathur, & Reddy, 2020). Although, the requirements for good practices of usability inclusion in the design and development process seemed to be more feasible following a hybrid methodology between heuristics evaluation and data logs collection. Therefore, the use of scored quantitative surveys and qualitative observations (with an interview format) will be the chosen tools that this research will employ throughout the methodology and the analysis chapters.

CHAPTER THREE: VR APPLICATION DEVELOPMENT

This study focused its methodology, results and analysis on the Virtual Reality Flight Training Experience (VRFTE) designed and developed as part of this project. VRFTE has been developed in conjunction with another researcher, while methodological evaluation approaches and outcomes have differed. The system is inspired by the prior experience of the author, as well as the growing implementations and interest in VR technologies in the field of personal training. For instance, using HMDs and VR environments with a serious game methodology has allowed studies in accident prevention and fire evacuation drills for tech manufacturers, depots and firefighters drills training (Ha, Lee, Lee, Cha, & Kim, 2016). Concerning the focus on the aviation industry by the VRFTE, there has always been profound attention to alternatives for training in maintenance, flight and ground related procedures due to the high cost of real-world aviation operations. Full flight simulators, pool ditching training, evacuations drills and maintenance mentoring are programs already used in the industry. However, with the increase of more affordable VR HMDs and commercial simulators, the horizon of VR training systems is expanding exponentially, making it a suitable field of choice for evaluating its usability. The following subsections detail the resources utilised, the design and development process, and the learning experiences in building the VRFTE system.

3.1. Hardware, Design and Development Tools

After establishing a clear idea of what was expected from the VRFTE application, the project team evaluated the options of game engines between Unity and Unreal Engine. This was based on the learning curve, programming language, VR capabilities, 3D modelling and the support of the available hardware provided by Flinders University and the developers' additional devices. The hardware is described below:

- HMDs: Oculus Quest and Quest 2 (LCD, res 1832x1920 per eye, 120 Hz, Snapdragon XR2) with dual controllers and hand tracking capabilities on Quest 2.
- Laptop Platform for Design and Development: Asus ROG Zephyrus GX502GW (Intel i7 9750H, Nvidia GeForce RTX 2070, Display Pantone 15.2'' 240Hz, 2Tb SSD storage, 32GB RAM).
- External Monitor: Xiaomi Mi Curved 34'' Freesync and 144Hz, HDR-capable for 3D Modeling.
- High-Speed Cables: USB 3.1 - 3m length.

Figure 5 - Figure 3 - UE 4.26 User Interface

Unreal Engine was chosen as the core development and testing platform with the outlined benefits and the open-source nature, easy access to information, and authors' prior experience with the software (Figure 5) (Engine, 2020). However, additional design, modelling and rendering software were utilised in the development process:

- Game Engine: Unreal Engine 4.26 + Programing Language C++.
- Version Control: GitHub Desktop 2.9.3.
- Modelling and Rendering Software: 3DS Max 2021 by Autodesk, TerreSculptor 2.0.
- Texturing and Design Software: Adobe Creative Photoshop 2021.
- Video Rendering Software: Adobe Creative Premier Pro 2021.

3.2. Design and Development Process

To generate the research question and create the base prototype to be analysed, the first phase of the VRFTE project included a general idea of the author's expectations. This initial pre-design and mockup were vastly broad and ambitious for the limited timeframe available on the research. After initial planning and consideration of minimum criteria, a more structured and achievable development was created, in which, instead of incorporating a full animated, aircraft model, the scope of the project focused on the primary objects and interactions of the cockpit panel that participants would be tested on (Figure 6). Accordingly, the Cessna 172SP or T41 Mescalero military designation was made in relation to its popularity among flight training schools due to its success and reliability (Buscombe, 2016). With that in mind, a defined list of specific tasks were selected that are required for the essential operation of an aircraft while on the ground.

Figure 6 - Cessna 172SP / T41D Initial Mockup (Research, 2017)

The total selected procedures were limited to five flows. Each Flow is designed in a tutorial mode concept that the user interacts with in a step-by-step process considering the allocation of actions within the cockpit, as Figure 5 and described in Table 2.

Table 2: Cessna 172SP/T41D Selected Flows

Tasks - Flows	Checklist - Steps
I. Initial Cockpit Check	(Q) Parking Brake – ON (F) Master Switch – ON (C) Battery Voltage – Check (I) Landing Lights – ON (J) Taxi Lights – ON (K) Navigation Lights – ON (H) Beacon – ON (L) Strobes – ON (P) Flaps – Extended
II. Before Starting Engines	(S) Fuel Selector – BOTH (Q) Parking Brake – ON (H, I, J, K, L) Lights – OFF (F) Master Switch – ON (A) Fuel Quantity – CHECK (N) Power – IDLE (O) Mixture Rich (Full Forward) (AB) Avionics Bus 1 & Bus 2 –CHECK OFF

III. Engines Start	(N) Throttle – OPEN ½" (F) Master Switch – ON (G) Fuel Pump – ON (O) Mixture – LEAN ½" (H) Beacon – ON (E) Ignition Switch – START (O) Mixture – RICH (Advance when the engine starts) (E) Ignition Switch – Check BOTH (B) Oil Pressure – CHECK (AB) Avionics Bus 1 & Bus 2 –ON (D) Transponder – ON
IV. Emergency Engine Shutdown on Ground	(O) Mixture – CUT-OFF (All out) (R) Fuel Shutoff Knob – PULL OUT (E) Ignition Switch – OFF (F) Master Switch – OFF
V. Engine Failure during Take-Off Roll	(N) Throttle – IDLE (Q) Parking Brake – Slowly to ON (P) Wing Flaps – Retract (O) Mixture – CUT-OFF (E) Ignition Switch – OFF (F) Master Switch - OFF

Phase two involved the initial sketches and outlining the learning curve that was required for the selected tools. Documentation, tutorials and training packages were used for obtaining the needed knowledge to start the game development cycle (Sparks, 2017). This scheduling had been planned in the project management software Jira and designed to follow the Agile SCRUM methodology. The development team consisted of two computer science students. One of them mainly focused on the programming elements, the functionality and logic of the system, while the author's role was designing, modelling and system evaluator. According to Higuchi and Nakano (2017), the chosen development methodology of SCRUM allowed for a reduced game development team to review and correct their deliverables in smaller portions; these portions are defined as Sprints, where a series of tasks were to be completed, tested and errors fixed in order to continue the development. Alternatively, Aleem, Capretz, and Ahmed (2016) argued that a game development software engineering process contributes to increasing productivity of the involved team members, where the whole picture of the project could be evaluated in real-time with the production and development teams. Fundamentally, it explains that instead of a traditional life cycle of pre-production, development, production and post-production, the nature of having both phases in parallel provides cyclical feedback to smaller deliverables, improving the game development process. For VRFTE applying these fundamentals, a timeline (Appendix A) was followed, with some amendments due to issues presented along the exporting and prototyping process.

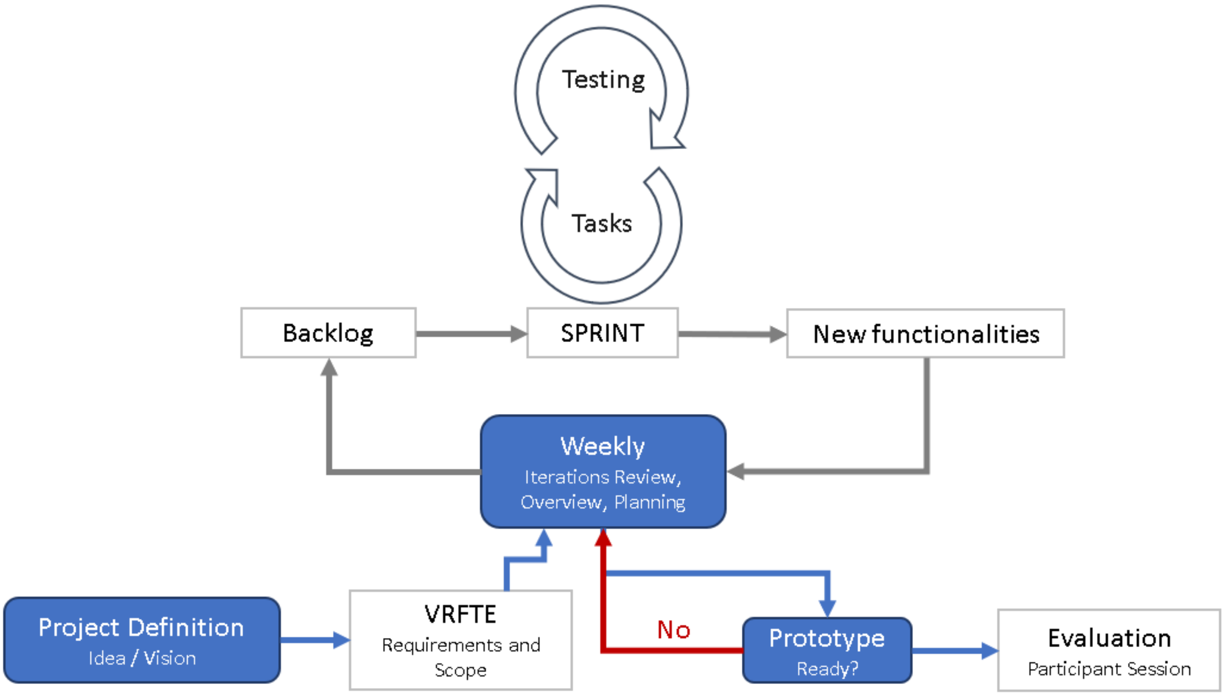


Figure 7 - VRFTE SCRUM Stages, based on (Azanha, Argoud, Camargo Junior, & Antoniolli, 2017)

The third phase followed the iterative behaviour of Figure 7, including integration with a repository version control created in the GitHub platform and committing all changes for each sprint using the desktop application (Tom, Chris, Hyett, & Scott, 2008). The team evaluated both modelling design and programming on the weekly meeting, with an additional online or face-to-face session with the supervisor. From a modelling perspective, an aircraft 3D model made from scratch was the most suitable option to maintain consistency, reliability and flexibility in exporting supported formats across the rendering software and the game engine. Thus, by using the official documentation and resources from Epic Games publisher of UE4.26, modelling the Cessna 172SP was made with a similar step by step process, from using the essential tools of object creation and polygon shaping such as extrude, bevel, and inset. The sculpting iterations were based on the official blueprints of the aircraft obtained by the aircraft specification manual (Figure 8) and preserving the real-world scale dimensions, as well as the standards from 3DS Max axis configurations.

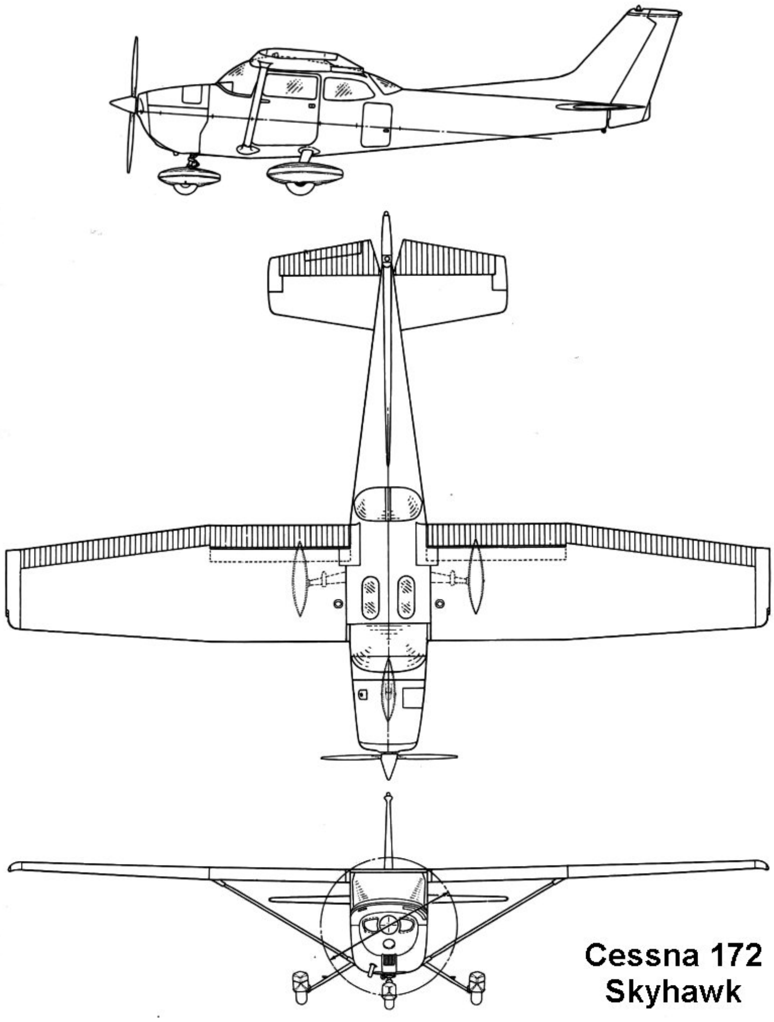


Figure 8 - Cessna 172SP Exterior Blueprint

Having finished the low poly exterior model of the aeroplane, the sculpting of the internal cockpit, buttons and handles were based on the researched media from the Xplane 11 simulator engine and the actual photos obtained by the author. Using modifiers built in the rendering software, such as TurboSmooth, a layer-based layout of components for object grouping was implemented. Phase four detailed the process of parameters, texturing and rigging of mechanical objects that were to be animated and included in the functionalities of the testing segments within the final prototype. These segments were individually exported files that would be automatically linked in UE4 by the scripted add-on commands. The first functional version was the cockpit panel with basic textures and a beta test for user interactions of a programmed tutorial of the chosen Flows. This beta contributed to an initial qualitative evaluation, with the results fed into future iterations to refine and improve the prototype.

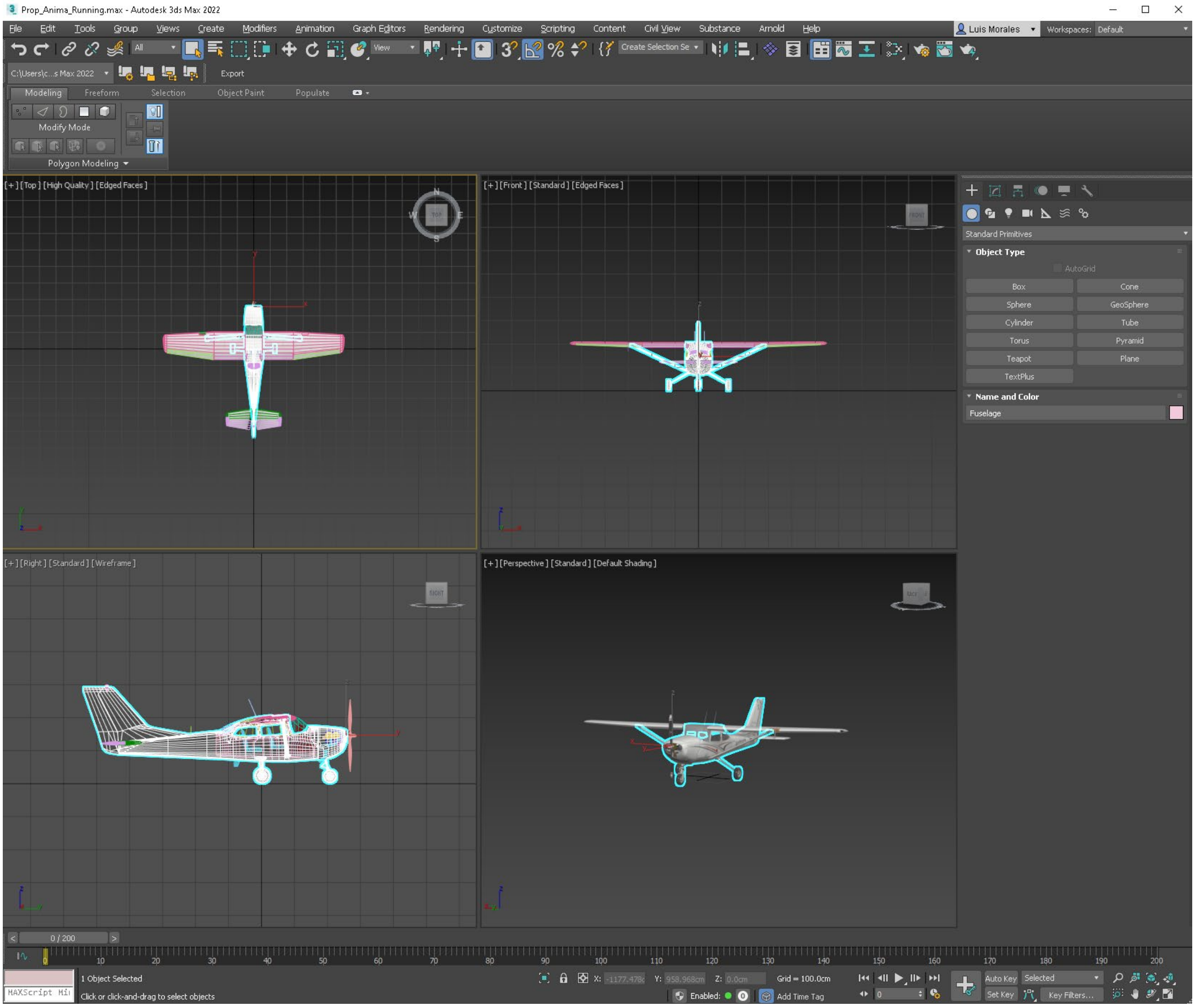


Figure 9 - Custom 3D Model for VRFTE

Phase five included integrating all 3D assets created in the modelling software with UE4 VR built environment and functionalities (Figure 9). Before exporting the objects, skeletal meshes, animations and settings, the use of instances for each asset was designed for optimisation purposes. Despite the optimising approach, the background code needed some restructuring to accept the inheritance behaviour of the scene objects. The use of pointers that reference the instances in UE4 was the strategy for seamless assimilation in the final application prototype. In addition, the inclusion of animations and sound effects for user feedback was essential to increase the sense of presence and immersion. These audio and visual feedback was included in the tutorial of all flows and the evaluation scenario of Flow No. 5. One of the visual cues was to highlight in green fluorescent colour, the instrument for the step the participant was performing. This increased the system's gamification and provided the correct sequence of steps for the tutorial (Figure 10). However, for the evaluation aspect, which was intended to be conducted without visual cues, this green highlight was removed. Only audio feedback and the complete Flow 5 checklist in the middle of the cockpit panel were provided to the participant (Figure 11).

Figure 10 - Flow 3 Tutorial Steps (VRFTE)

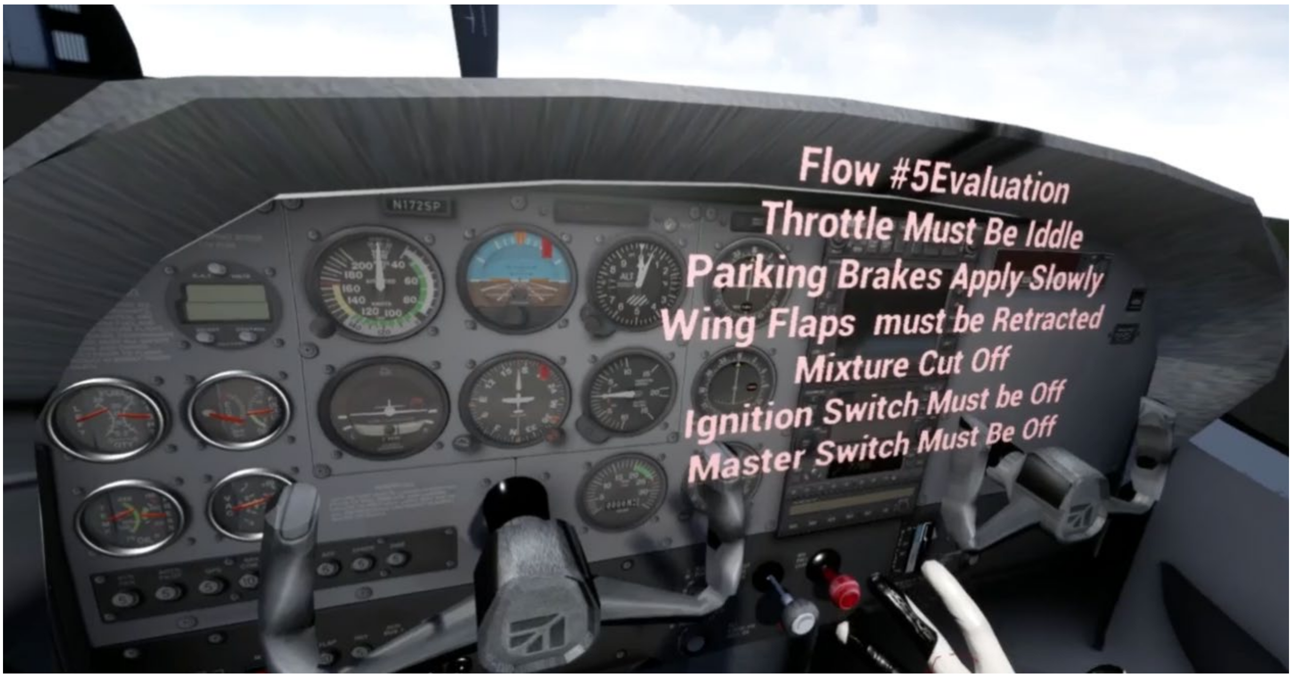


Figure 11 - Flow 5 Evaluation (VRFTE)

Lastly, to improve the world outside the cockpit of the plane open-source data was obtained through a geographic information system (GIS) from the USGS global database (USGS, 2021) and then processed by the software TerreSculptor to create bitmaps with terrain elevation shading of the Goolwa Aerodrome located south of Adelaide in Middleton, South Australia. These bitmaps were compatible with a new feature of terrain modelling added to Unreal Engine in version 4.26. By importing such a file, the game engine could recreate elevation into a plane object (that is a geometric shape not an aircraft) that is reshaped according to the depth of field within the scale of the bitmap image. For detailing and optimisation, the terrain sculpting mode in UE4 was used to decrease the rendering requirements, and polygon count within VRFTE.

3.3. Challenges and Learnt Lessons

When the tools were selected, an initial proposal was to purchase an existing market-built 3D models to accelerate the production. This resulted in acquiring a Cessna 172D 3D model from an independent creator on the website TurboSquid. The artist that created this 3D model claimed its compatibility with the rendering software and game engine. However, this decision became one of the most significant issues presented throughout the modelling and development process. When the modifications, rigging and texturing were completed, the model's inconsistent polygon structure, incompatible scale, incorrect real-world coordinates, and exporting issues to UE4 made it an unusable asset for the final integration. As a result, the modelling of a custom-made Cessna 172SP/T41D was the solution adopted, as described in the previous section.

The differences between the game engine and the modelling software created an iterative process of testing the 3D axis world orientation (Figure 12), where in UE4 the "x" axis faces forward, whilst in 3Ds Max is the opposite. To solve this was necessary to incorporate scripting of automated commands in 3DS Max to correctly define the settings of the transformed MAX format into FBX UE4 compatible format. In addition, a setup of duplicates between committed versions created an issue in the definition of blueprints and pointers of reference due to the inconsistency of object naming, grouping and classification. Therefore, a clean-up of all assets and a compilation of the final version in the repository allowed the successful deployment of the final 3D model with functionalities and serious game components to be ready for testing.

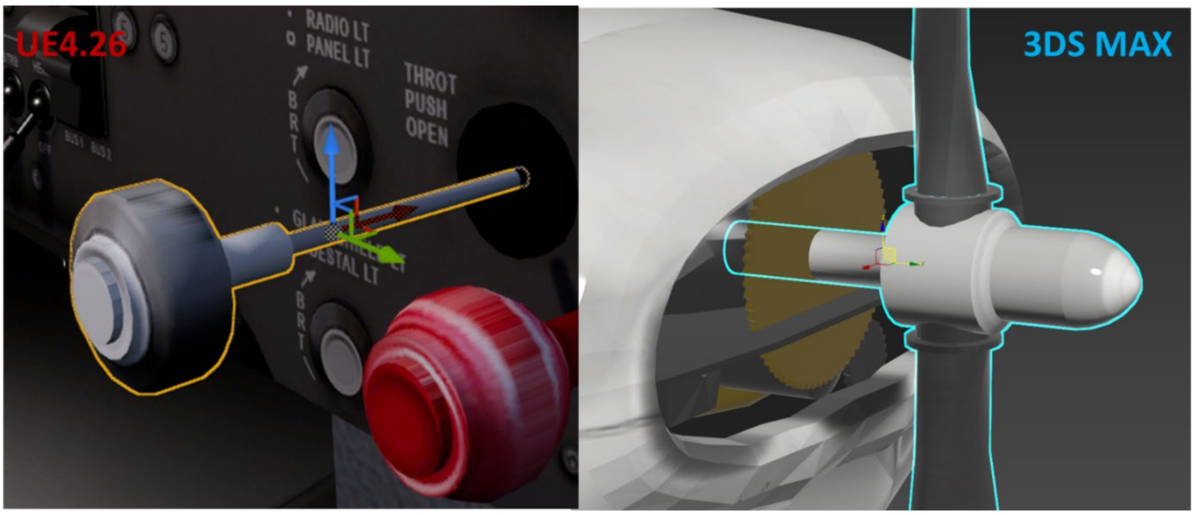


Figure 12 - World Axis orientation UE4 vs 3Ds Max

One of the challenges for the project was the complexity of the models and the processing cost associated with this complexity. The intended deployment platform utilised Android APK executables which were system limited in size. Therefore, there was a need to optimise and reduce the size of the models ready for exporting into the Oculus Quest 2. Therefore, when the terrain model was implemented as an externally visible element from the cockpit, the overall polygon count and rendering load increased considerably, resulting in an early failed attempt of reproducing the Goolwa Aerodrome. To identify the correct proportion of the terrain size and map optimization, a total of three maps were produced with different sizes and polygon details. The original scene was the largest on a real-world scale of 1000 x 1000 metres, followed by a scale reduced by half resulting in 500 x 500 metres and the last scene of 250 x 250 metres. Objectively, the latter option was picked as the solution as it provided enough visual elements and surrounding awareness without compromising the actual design of the system, thus becoming part of the final prototype, as shown in Figure 13.

Figure 13 - Terrain Scene External View (250m x 250m)

An issue that arose during the development of VRFTE had an impact on the modelling activities completed by the author. Hand tracking was implemented which resulted in the deletion of prebuilt collisions and code for the Oculus Quest controllers. This caused an issue with the rigged and animated models created in 3DS Max and then imported into UE4, in which had to be resolved by editing the position and direction parameters of components that had already been implemented. Due to the parallel modelling of the in-game assets, some of these improved changes in the scene became a case of concern that required solutions in both the modelling application and the programming logic of the developer. For this reason, the final version of the VRFTE system needed a complete review to dispose of redundant and unnecessary components, resources and code that were impacting the overall performance.

CHAPTER FOUR: METHODOLOGY

A common trait found in the literature review is that either software engineering, game development, or VR development processes consider usability a critical attribute of any system. Thus, as a relatively recent field in the consumer computing market, VR systems are seeing increasing adoption of usability methodologies and testing techniques used within the development and post-production cycles, despite this it is yet to have its own standards. As a result, the usability evaluation of VR products is approached by applying either traditional, new or adapted methodologies. This chapter discusses the research questions as well as usability evaluations protocols implemented for the VRFTE.

4.1. Research Questions

As discussed in Chapter 1, the usability evaluation of the VRFTE system aims to answer the following questions:

- RQ1: How does the VRFTE System in the designed tasks motivate the user to practice them multiple times in the session?
- RQ2: What factors provide the best playability experience for the user?
- RQ3: How upgradable is the system according to the feedback of the user?
- RQ4: What are the benefits to the user by giving them feedback on their performance?
- RQ5: What is the reception of the participant and the likelihood of using the VRFTE system again?

4.2. Usability Implemented

Based on the literature review, it is excellent practice to follow and implement the ten fundamental heuristics principles in an HCI application. For VRFTE, from the conceptualisation of the project to the development process, an iterative assessment of the viable capabilities and the scope of the research reduced the priority of some of the principles. They also constituted more realistic milestones for the programmed testing sessions. Otherwise, the complexity of the project would have gone beyond the available time and research limitations.

Through analysing the possibilities of the VRFTE, five of the heuristics fundamentals were the centre of the design and the development to provide a suitable user experience. In addition, four were partially implemented from the remaining principles, whereas principle VII was discarded in the prototype version; this approach is detailed in Table 3.

Table 3: VRFTE Applicable Heuristics of Chapter 2.2.

Principle No.	Description	Implementation	Reason
I	Visibility of System Status	Yes	Inclusion of Highlighted visual instructions for interaction
II	Match Between the System and the Real World	Yes	English language and designed VR environment as familiar setup of an actual aircraft
III	User Control and Freedom	Partial	In a degree with individual buttons that allow the user to change, repeat or finalise the tutorial and evaluation
IV	Consistency and Standards	Yes	VRFTE maintains 3D objects with features and dimensions used in a real-world setup such as handles, key switches, buttons and levers.
V	Error Prevention	Partial	Functionality within the tutorial to prevent the user from continuing into the next step of the Flow if the action performed is in the incorrect position
VI	Recognition vs Recall in UIs	Yes	Audio cues of the interacted objects that are originally from real-world sounds, as well as the implementation of the same concepts in a primary aircraft that simulate some of the cockpit elements that are similar in a car
VII	Flexibility and Efficiency of Use	No	VRFTE has linear interactions and is out of its scope to provide a different option for hand tracking or the use of other peripherals to operate the elements
VIII	Aesthetic and Minimalist Design	Yes	Textures, 3D elements and written text are provided with functionality in mind; there are not unrequired elements within the required interactions
IX	Help Users Recognise, Diagnose and Recover Errors	Partial	When an action or state of an object is incorrect, there is a message that explains the user being unable to continue onto the next step
X	Help and Documentation	Partial	Although no written information of the system functionalities is provided, a video introduction and tutorial are given to the participants and is projected as an alternative on TV Screens around the evaluation session

The heuristics fundamentals became a tool for the system's development and testing process. Their argument was revisited in each sprint to better adapt them to the two usability evaluation sessions that were run and will be explained in the following section. Likewise, a conscious development team targeting satisfaction by putting themselves into the thoughts, aesthetic appeal and possible preferences of the user (feelings), focusing on effectiveness by task completion and adequacy, and placing efficiency on the performance concerning error reduction and maximising the UI aids for the evaluation session (Yanez-Gomez et al., 2017). Admittedly, these heuristics are not the only major usability concepts that are influencing the VRFTE system. Terms such as playability and haptics feedback are becoming part of the structural heuristics tackled in serious games. As a gamified system, VRFTE considered, to some degree, the playability linked to user satisfaction.

4.3. Usability Testing

The methods in usability testing employed across the two sessions of the VRFTE usability evaluation were classified based on the VR evaluation index framework in Table 4. This framework illustrates how the system was evaluated in a qualitative review approach in a pilot testing session (initial investigatory testing session, not aviation professional), conducted with a sample of three participants. Despite the low number of users involved in this initial test, the evaluation was intentionally designed to acquire rapid results in the mid-development process of the application, helping to adjust, correct and improve the usability for the final prototype and subsequent evaluation.

The actual prototype evaluation session involved a mix of controlled experiment and cognitive walkthrough approaches in a quantitative review of fifteen participants. This approach generated the log records within the application in text format, which were manually analysed by exporting them into MS Excel for data organization. This evaluation was only executed following the ethics approval of the Human Research Ethics Committee (Project ID 4487) (Appendix C). Additionally, all participants completed a demographic survey incorporated into the findings (Appendix B) and read and agreed to the consent form (Appendix C) provided at the introduction of the test session.

Table 4: VR Evaluation Index Framework, based on (Wang, Li, & Zhu, 2019)

System Purpose (P)	Criteria Index (C)	Feature Index (F)
VRFTE Usability	System UI (u1)	Textures Quality (u1.1)
		Text messages (u1.2)
		Interface Style (u1.3)
	Mechanics (u2)	Highlighted Instruction (u2.1)
		Hand Tracking (u2.2)
		Undo or Redo (u2.3)
	Playability	Object Interactions (u3.1)
		Tutorial Video (u3.2)
		Audio and visual feedback (u3.3)

While testing of the application occurred during development, this was focused on finding bugs and issues with the application. Formal user evaluations had to be conducted to determine the usability of the application. These user evaluations are detailed as follows:

Test 1 – Qualitative Review: This usability evaluation methodology was performed by testing three participants who experienced a beta version of the system. At the same time, the development of the prototype was a little over half of the expected progress. Although the planning was to conduct this testing with expert users, due to the lack of volunteers and available users with experience in flight training, the sample were interested students from the University of South Australia undertaking aviation education, who had heard about the project from an external source. The setup for this test was in a controlled room, and the VRFTE early version was connected by cable to the main computer running the UE4 game engine. The aircraft 3D model only included the cockpit panel without final texturization and with limited interactable elements: the throttle handle, mixture handle, parking brake and yokes. The test itself involved an early stage of one of the tutorial tasks, specifically the Flow 1 “Initial Cockpit Check”; however, the in-game text checklist within the UI was static and restricted to the right side of the cockpit. The objective of this test was to understand the ease-of-use perception by observation analysis of the user actions while in VR and a post-interview based on an empirical method of using an informal usability survey (Appendix D).

Test 2 – Mix of Controlled Experiment and Cognitive Walkthrough Review: On this approach, the goal of mixing both methodologies was to contemplate applicable metrics from the feature index on Table 4. In addition, by incorporating a modified informal usability survey based on the original SUS, explained by Lewis (2018) the standard range between highly or strongly with agree or disagree could be adapted to the metrics required in the respective research. Firstly, to understand the hybrid evaluation, it is essential to discuss their differences. Karre et al. (2020) describe a controlled experiment as a widely used method that utilises presence surveys and informal usability surveys to obtain the users' reactions post usability test. Whereas cognitive walkthrough requires the involvement of evaluators to record the participants' usability reaction by observation during the testing session.

With this distinction in mind, recruitment for the evaluation commenced in August 2021. A total twenty initially interested participants were invited for testing session.

Even though the actual number of participants was reduced to fifteen, the sample size permitted the quantitative testing approach to continue. The individual evaluation was performed in respective dedicated times of approximately thirty minutes per participant, with allocations of the session distributed in a timeframe of two weeks. Prior to the testing day, users received via email an introductory video “VRFTE Tutorial” (Appendix E), in which the author orientates the user on the basic procedures, terminology and general idea of the research. This introductory video was also displayed on the TV screens in the room where the tests were conducted. This meant that where a participant had not watched the video, they could opt to before the start of their session (Figure 14).

Figure 14 - VRFTE Participant watching VRFTE Tutorial

After a brief explanation of the Oculus Quest 2 HMD, safety considerations and reinforcing their ability to stop the session at any moment, the participants were presented with the in-game instructional tutorial of the flows in Table 1. While in the session, the facilitator engaged in subjective communication using one of the three most used techniques in the cognitive walkthrough evaluation, such as Echo, where the user generates a question and the evaluator answers by rephrasing the same question, motivating the user to give more information. The second technique, Boomerang, returns the user's query with another question to obtain their own opinion on the question. And the Columbo technique, where the facilitator formulates a non-leading question to the user in order to get more information for a particular interaction that the user might have missed (Martins, Kirner, & Kirner, 2015). For this research, in addition to the observed interactions, the boomerang technique was constantly used to obtain information from the user. At the end of the VRFTE in-game tutorial, the user was directed to initiate the evaluation scenario, which required the participant to follow a flow checklist without the highlighted visual cues. It was during this phase where the evaluator analysed the gestures, body language and possible questions asked by the user.

Lastly, when the user testing session was completed, the participant was provided with the adapted SUS questionnaire (Appendix F). The participant was offered privacy and time to answer at their convenience. This technique intends to acquire an objective review and feedback that could portray the user's perception of the usability of the VRTFE. This survey data is then tabulated in a CSV file for results analysis, as shown in chapters five and six.

CHAPTER FIVE: RESULTS

In this chapter, the findings collected by observation, system logs, and user feedback are presented and organised according to the methodologies and techniques that were followed in this research. In particular, both qualitative and quantitative results supported the initial premise of including usability evaluation as the pointer to design an easy-to-use system so that the users could achieve specified tasks with the parameters of efficiency, effectiveness and satisfaction. From the initial qualitative review used within the development cycle to the combined methods of having a controlled experiment and executing a cognitive walkthrough; helped to obtain valuable data of the user reactions in the testing scenarios, where the participant went through the VRFTE interactive tutorial and a evaluation module. The results are highly valuable for the system evaluation of usability perception and how they were obtained, and the data collected are discussed in the following sections.

5.1. Qualitative Review

This test was performed in the development process by observing the user while interacting in a VR 3D environment that included the basic cockpit panel and some of the primary interactable elements of the final design. The test settings were based on Flow 1, programmed to only react to the actions of the user and without audio or system message feedback. Three participants were involved in the test, and the results of the interviews were summarized in Table 5 and Table 6.

Table 5 - Interview Questions based on Appendix D

Interview Questions	
1	First Impressions?
2	Difficulties using the HMD?
3	Thoughts on the User Interface and layout?
4	Accurate experience for training?
5	What would you change or improve?
6	Would you use this type of VR app for learning a skill or training in other fields? If so, give an example.

Table 6 - Qualitative Review Notes

<table>
<tr><th rowspan="2">Participant</th><th rowspan="2">Test Observations</th><th colspan="3">Interview</th></tr>
<tr><th>Questions</th><th>Body Language</th><th>Feedback</th></tr>
<tr><td rowspan="3">VRBT1</td><td rowspan="3">- Manifested surprise
- Express positiveness on the object's reaction to the hand tracking
- Asked about where to press for most of the instructions</td><td>1</td><td>Straight and without hesitation</td><td>Surprised by the immersion, even though it was not realistic to be sitting in front of a panel surrounded by blue skies and the monochromatic tone of the floor.</td></tr>
<tr><td>2</td><td>It took him some time to answer and hesitation</td><td>She expressed a little discomfort on the blurriness of the image and how the head band straps were difficult to tighten.</td></tr>
<tr><td>3</td><td>Confused face</td><td>She recalled it being blurry and without enough detail, but it could be directly related to an aircraft cockpit.</td></tr>
</table>

	- Impatience in some of the object animation delays	4	Head tilted to the side, thinking expression	Believes that it could be, but still has a long way, there is no tutorial or description of how to proceed correctly.
		5	Calm expression, show confidence while talking	Feedback, resolution of the environment and responsiveness of the object's interaction
		6	Positive facial expression, smiling and joking tone	Yes, driving another type of vehicle, expressed the interest in boats
VRBT2	- Constantly touching the HMD - Constant movement in the chair, trying to realign himself with the VR environment - Questioning about where to press multiple times -Manifested some degree of motion sickness	1	Doubting facial expression	The user felt that he could not adjust the HMD properly and affect his test performance.
		2	Confidence with a serious face	Band straps tightness and lenses ergonomics
		3	Neutral expression	A little blurry but overall expressed that it was suitable for his experience.
		4	Confused expression and some delay to answer	Considers that it could be beneficial for a training program.
		5	Serious expression and quick response	The HMD straps, and the inclusion of some mechanism of providing step by step instructions.
		6	Delay to answer and relax facial expression	Yes, in factories when employees need to operate specific panels that activate factory systems.
VRBT3	- He took some time to adjust the straps of the HMD and adjusting his seated position - Straight posture and confidence, few questions but direct to missing information that he was expecting - Tried to interact with the environment outside the cockpit panel and provided elements.	1	Neutral expression	Noticed that the hand tracking was accurate and with low latency
		2	Quick response and without hesitation	Expressed some discomfort due to the amount of pressure of the straps, but that was necessary to improve the visual resolution.
		3	Thinking expression and short delay to answer	Though it was too bright, in general was a clear layout, although the panel information was not possible to read the label of the objects.
		4	Hesitation and serious facial expression	He manifested that it has possibilities and could be refined as a training option.
		5	Cheerful expression and straight to the point	The lighting, the instructions setup, the lack of information provided to the user, and more details in the cockpit elements and panel.
		6	Started with a serious expression, followed by a joking tone.	Yes, he believes that this type of app could be used in all kinds of training, even in sports.

5.2. Controlled Experiment and Cognitive Walkthrough

The approach to combine both methodologies permitted the recompilation of empirical information from the facilitator's observations, the in-game logs of each participant's interactions and the recorded feedback from the modified SUS questionnaire answered at the end of their tests. The fifteen participants had individual test sessions, in which they completed the tutorial activities based on all the flows from Table 1 and the evaluation task. The usability testing was conducted in a controlled environment, where the facilitator provided the introduction to the assessment, including support and considerations. Firstly, in terms of the overall demographics of the fifteen participants, a graphical analysis of previous experience in VR is demonstrated in Figure 15, in addition to a pie chart in Figure 16 that describes their previous Flight experience. Secondly, all annotated observations were taken in an MS Excel spreadsheet; to illustrate the results per user, the participant with ID VREC7 is taken as an example, and the comments are made in Table 7.

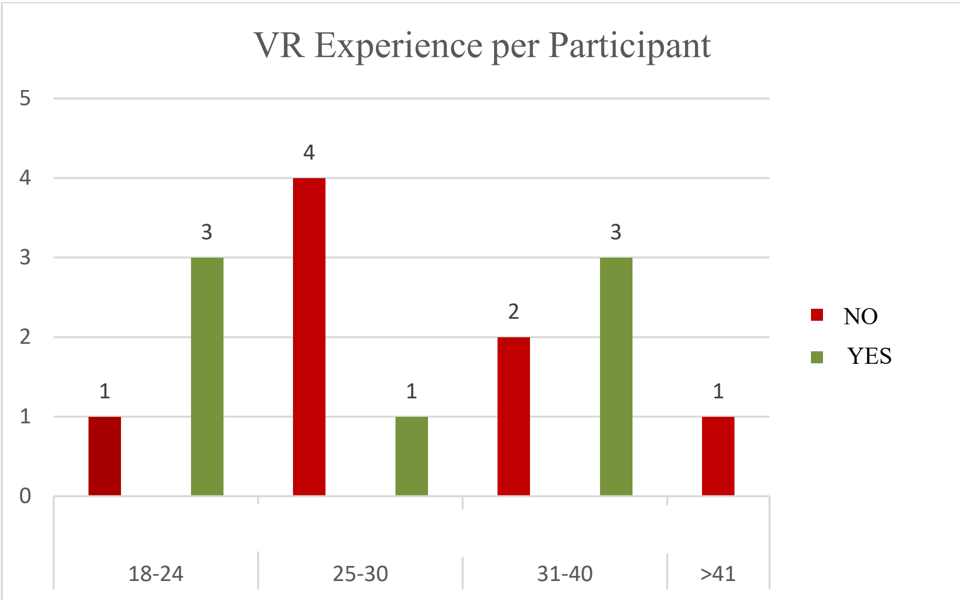


Figure 15 - Participants by VR Experience

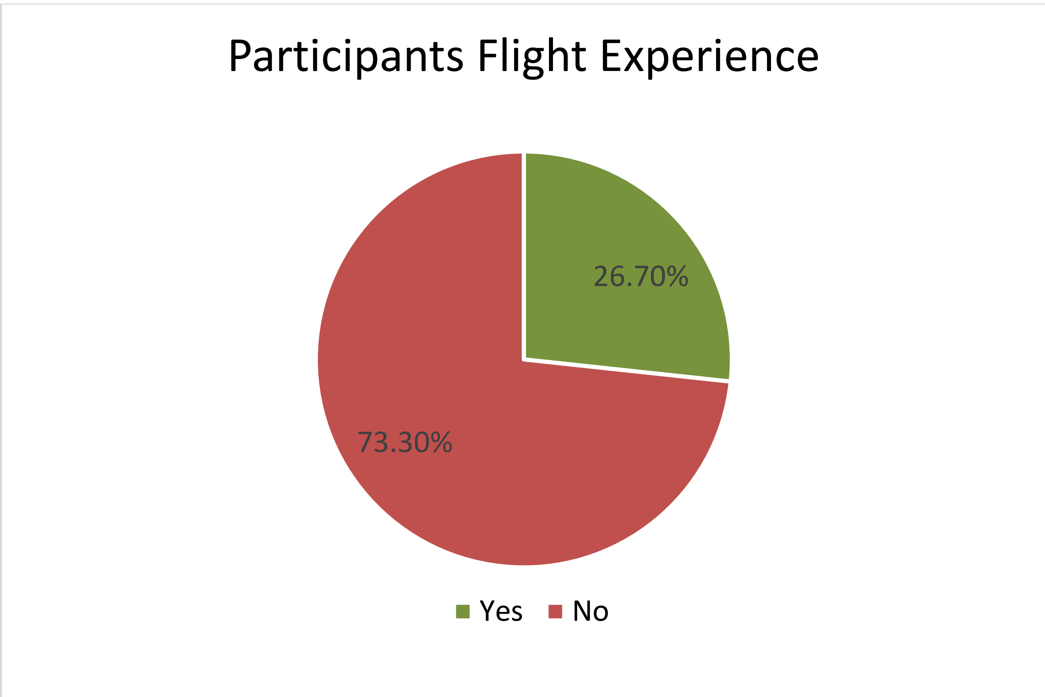


Figure 16 - Participants by Flight Experience

Table 7: Annotated Observations - Cognitive Review

VRFTE Cognitive Review	
Participant	**Observations Flow 5 Test**
VREC7	Confidence and good posture
	Cheerful attitude and willingness to properly listen to the facilitator's instructions
	Explained each performed step, and with a tone expecting confirmation
	No rush, however, the grabbing gesture for the parking brake was significantly pronounced

Consequently, recorded data of the user interactions were logged and stored in a table that detailed the number of iterations made per step in the complete evaluation. As an example, Table 8 describes the results from the user VREC7 according to Appendix I.

Table 8: Flow 5 Evaluation Actions Log

FLOW	Flow5
TRAINING OR EVALUATION	Training Activated? 0
Row Labels	**Num Iterations**
VREC7	**14**
Iddle ThrottlePosition	1
KeySwitch Cooling Off Engine	1
MasterSwitch Off	4
MasterSwitch On	3
Mixture Iddle	1
Navigation lights On	1
The Parking Brake Is On	1
ThrottleFull	1
WingsFlaps Up	1
Grand Total	**14**

On the other hand, the standardised SUS survey results are the core usability questions exposed to the participants immediately after the testing session was performed. Based on Figure 4, the VRFTE system implemented these set of questions and obtained an average result of 85.83% usable perception from all the users shown in Table 9.

Table 9 - VRFTE Data SUS Score

Participant ID	Q1: I Think I would like to use this VR application frequently	Q2: I found the VRFTE app unnecessary complex	Q3: I though the VRFTE app was easy to use	Q4: I will need the support of a technical person to be able to use the VRFTE app	Q5: I found the various functions in the VRFTE app were well integrated	Q6: I thought there was too much inconsistency in this VR app	Q7: I would imagine that most would learn to use this VR app very quickly	Q8: I found the VR app very awkward to use	Q9: I felt very confident using the VRFTE app	Q10: I needed to learn a lot of things before I could get going with this VR app	SUS Raw	SUS Score
VRYC1	5	2	5	2	4	2	5	2	5	1	35	87.5
VRSY2	2	1	3	4	2	2	4	4	1	5	16	40
VRTU3	5	1	4	1	5	1	4	2	5	3	35	87.5
VRTU4	5	2	5	2	5	1	4	2	5	1	36	90
VRAK5	5	1	4	1	5	1	5	1	5	2	38	95
VRAI6	5	3	5	1	5	1	5	1	5	1	38	95
VREC7	5	1	5	2	5	1	5	1	5	1	39	97.5
VRAW8	5	1	5	1	5	1	4	1	5	1	39	97.5
VRCD9	5	1	5	1	5	1	5	1	5	1	40	100
VRTD10	4	1	4	4	5	2	4	2	4	1	31	77.5
VRPM11	5	1	5	5	4	2	5	1	4	1	33	82.5
VRTA12	4	2	4	1	5	1	4	2	4	1	34	85
PTAK1	5	1	4	4	5	1	5	2	4	2	33	82.5
PTAK2	5	1	4	4	5	1	4	1	5	2	34	85
PTAC7	5	1	4	1	4	2	5	1	3	2	34	85
								Average			34.33	85.83%

Finally, a graphical analysis of the tabulated data from the compilation of all participants surveys (Appendix F) and processed answers (Appendix H). The diamond diagram in Figure 17 shows the covered area of positive and negative perception according to the classification of participants by their experience levels in virtual reality environments or flight experience as an aviator.

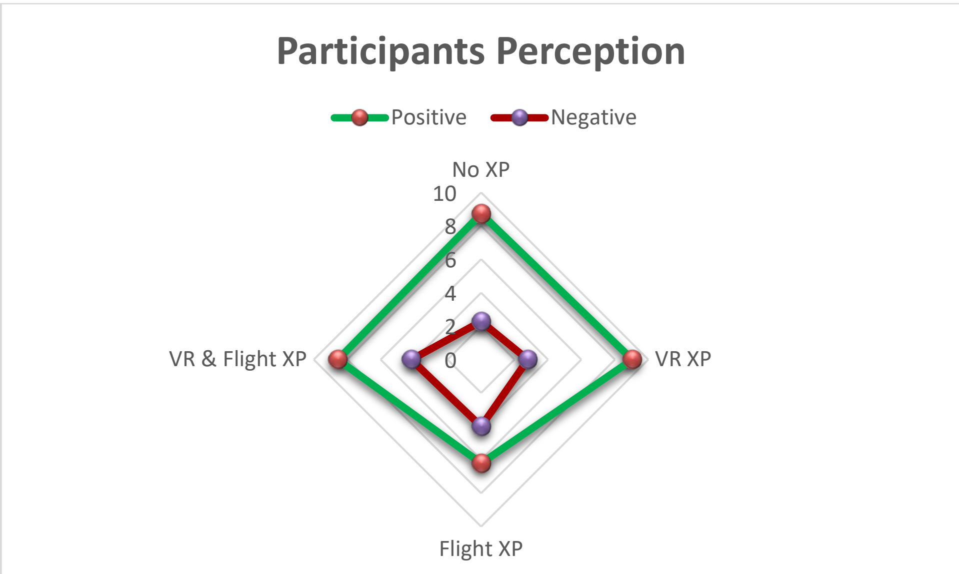


Figure 17 - Usability Perception (Diamond Diagram)

CHAPTER SIX: DISCUSSION AND ANALYSIS

Based on the results of the two iterations of usability evaluation made during the development and the completion of the VRFTE prototype, this study can answer the research questions presented in Section 4.1. It is crucial to note that the industry where this research was done has some regulations and guidelines for software that involves training and safety procedures. Therefore, the arguments that are presented in this paper are oriented academically and are outside the scope of being adopted as a primary tool for skill and knowledge training in the operation of the Cessna 172SP (Ronell, 2020). However, the results and conclusion from the research could provide some basis for further study and the development of a more robust training system.

RQ1: Does the VRFTE System designed tasks motivated the user to practice them multiple times throughout the session? – In both qualitative and quantitative results, the user's interaction in an immersive environment allowed for an increased interest in performing the required flows with the instructions provided and the opportunity to visually interact with the designed elements of the VR environment.

According to the results in section 5.1, the application of heuristics evaluation techniques and a qualitative review approach provided an initial understanding of the problems in the usability design, specifically in the 3D model and the programmed interactions. As Asghar, Cang, and Yu (2018) explained, including a more personal communication with the user in the early stages of the development provides precise information of the flaws, features to be improved, and new ideas to the development team that might have been omitted in the first design. One of the repeated observations of all three participants was the time taken and learning process of adjusting the Oculus Quest 2 HMD; this indirectly is reflected in the comments mentioned by all of them in the interview segment of the test, as shown in Table 6.

While the nature of subjectivity involved in the comments from each user could widely vary, the results from the interview showed similar opinions on specific issues or features. On the positive side, participants agreed on the possibilities and advantages that the VRFTE system could bring to training personnel studying to become pilots or interested in flying. On the other hand, similar feedback involving the lack of more straightforward instructions or the inclusion of step-by-step aids for the flow procedures made them feel lost or unsure where to press or which portion of the cockpit panel they should focus their attention. Subsequently, the absence of haptic or audio feedback along with the implemented visual feedback affected the sense of presence and immersion. Lastly, and perhaps the most critical opinion in all three users was the uncomfortable design of the straps in the HMD;

this issue generated some discomfort that could have influenced some bias in the observations the participants made about the experience.

These early findings allowed the design and development team to arrange some additional features and improvements within the development cycle. These included implementing guided instructions within the tutorial (highlighting in green the object to be activated or deactivated). Additionally, the inclusion of real-world cockpit sound and object interaction audio, provided a more immersive experience. In regard to the HMD straps, a feasible solution would have been to replace them with a custom made strap that includes memory foam and easier clutches for tightening and a release mechanism. Thus, the final test briefing prior to the in-game evaluation includes an understanding of the straps' functionality and how to adjust them, aiming for comfort and better visual resolution.

RQ2: What provided the best playability experience for the user? – As outlined above, in the development cycle, the portability of the Oculus Quest 2 HMDs, the system's visual and audio feedback and tutorial assistance interactions increased the overall sense of accessibility in the users. As a result, by defining a specific objective within the evaluation module, the user sense of satisfaction was obtained by overcoming the challenge of interacting with the system without the assistance of the tutorial cues. Consequently, from the survey results the VRFTE could be consider a highly usable system, with factors such as the degree of immersion and the easy to follow interactions.

RQ3: How upgradable is the system according to the feedback of the user? Within the comments taken by the facilitator, there is a high perception of the possibilities that the system could implement to become a more immersive version and an easier to use application. However, some limitations are evident as not all the suggestions from the users could be achieved or are part of the system's primary purpose.

With the modifications and lessons obtained in the first usability test, a controlled experiment approach and cognitive walkthrough evaluation methodologies were used to evaluate the final version of the VRFTE prototype. As per section 5.2. the findings compiled a classification of the participants by age and level of experience in VR or Flight training, as well as the user testing techniques of observation, communication and review of the user feedback.

Firstly, from a quantitative analysis of the results, it could be seen that the percentage of participants with and without experience in VR is approximately the same, with no experience being 8 participants (53.33%), and 7 participants (46.67%) having some VR experience as presented in Figure 15. In comparison, a lower 26.7% of users had flight experience. These details, when compared to the diagram in Figure 17, allowed for a visual impression of the negative or positive user perception within the selected sample. Although a slight variation, the diagram suggests that users with no previous experience in either VR or flight training had a better positive perception of the VRFTE system. This is explained by the distance between the purple and red dot in the diamond area of the No Experience category, as demonstrated by the larger gap. By contrast, participants who had Flight experience have a smaller distance between the red and purple dot in their diamond area, and therefore a lower level of satisfaction with VRFTE. This could imply that due to their previous experience in flying, they would be considered at the expert review level, thereby applying a more objective and critical approach to the test. The usability perception of the system could vary significantly, but not for the lack of usability but more for the constructive feedback that could come from a selection of typical users in the target audience, in this research, a flight training scenario.

Secondly, as the tutorial and test were designed in a task-based procedure, the observation of how the participant interacted with the system provided findings that touched even some of the personality traits of each participant. For instance, a deeper analysis of the obtained data and information from participant VREC7 could provide an idea of a character that is easy-going and spontaneous, who is willing to participate without hesitation; this is correlated to the demographics details of VREC7, which placed her in the category of No Experience. From the facilitator's comments in Table 7, it could be assumed that the user's personality infers analytical thinking, in which actions taken were made with verbal confirmation or reception of approval from the facilitator, providing a higher degree of security for the step to be taken. Appendix G demonstrates a user's interaction with Flow 3 within the tutorial portion of the test. Table 8 describes the number of actions per step that were taken to complete Flow 5, which motivated further analysis of the usability in terms of the interactable objects' sensitivity to the user actions. It is clear that an unintended selection of the navigation lights was made and that the system performed several involuntary actions on the master switch button. By combining the information in Table 7 with Table 8, it could be argued that there is an issue in the usability of these buttons, where the collision of the user's hands and the instruments could be imprecise, causing the user to activate and deactivate the element without being aware of it. This type of analysis allowed for a better understanding of the unavoidable bugs that could be presented in a VR system, helping to improve the production and post-production cycles.

Finally, the mixed methodologies used in this research contributed to evaluating the ease of use of the VRFTE system from an objective and subjective point of view. The overall findings have shown that the users' age did not significantly influence the perception of usability; instead, their prior experiences with similar environments or setups slightly increased their expectations, hence influencing their system usability perception.

RQ4: What are the benefits to the user by giving them feedback on their performance? – By comparing the first usability test in the earlier stages of development with the final prototype test, it could be demonstrated that improved feedback to the user allowed for better results in their interactions and in following the instructions required in the tasks assigned. Nonetheless, due to the limitations of usability testing iterations, it is not possible to determine the exact benefits apart from the possible improvement of accuracy in their interactions.

RQ5: What is the likelihood of the participant using the VRFTE system again? – To answer this, all the applied techniques and acquired results from chapters four and five contribute to the correlation of the quantitative scores with the analysed annotated information from the observed participant interactions. However, one of the most straightforward results is the obtained 85.83% of the SUS survey (Table 9); this average of the overall SUS scores from each participant yields a single number representation of how usable is the VRFTE system in relation to the user perception. On that account, that quick argument allocates the evaluation in a general overview of usability by including the results in Figure 17; it could be argued that the collected answers and calculated data from Appendix H supports the SUS score obtained. By looking at the outer nodes of positive perception, it is conspicuous to see that the diamond shape on green colour covers the red nodes of negative perception, implying that the reception surpassed the rejection of the VRFTE system.

Accordingly, with the cognitive data and the sores outlined above, it could predict the chances of the participant's using the VRFTE system in the future are considerable. In spite of the higher expectations from experienced participants that affected the feedback, the general acceptance of the system is proportional to the designed scope of the project. Consequently, it could be implied that the purpose of the usability evaluation was achieved, allowing to identify the flaws, possible improvements and the users' perception.

CHAPTER SEVEN: CONCLUSIONS

This thesis has aimed to analyse and evaluate the usability of the Virtual Reality Flight Training Experience System developed and motivated by the author's findings on the lack of VR interactive tools for aviators. As a result, a lower level of proficiency training in-flight crews is a consequence of the COVID-19 pandemic, which affected most industries between 2020 and 2021.

The literature review clarified a pathway proven in the implementation of a mixed approach in usability testing, furthermore, by applying some of the usability heuristics principles throughout the development cycle and running an early usability qualitative test with three participants, allowed to minimise some of the issues that could have influenced the prototype test involving a considerable number of users. The last test was conducted in a controlled experiment setup approach, employing a System Usability Scale (SUS) survey customised for this VR system. The captured feedback from the participants provided a clearer picture of the usability perception, along with the annotated observations made during the test session of each user. In addition, the log recorded feature built into the system permitted an additional dataset that complemented the findings on the other techniques.

To conclude, although there are no specific standards for usability evaluations in VR products, the most common methodologies could be adapted to the needs of a particular system. Hence, in the case of the VRFTE system, the implementation of the techniques of qualitative review for the early usability test and followed by the use of cognitive walkthrough and controlled experiment methodologies, allowed for results that were a combination of objective and subjective information of the user's interactions. On the one hand, it was evident that the age difference among the participants was not an influencing factor in the usability reception of the system. On the other hand, their previous experience in virtual reality or flight training environments impacted their perception of usability. In some cases, through observation and linkage to the collected data, the expectations from users with experience were higher than those without any experience. In the latter, perhaps the VR immersion's novelty influenced the positive feedback in a subjective model. However, the knowledgeable participants went beyond the limitations and scope of the system prototype, trying to objectively provide the improvements required for the system to become closer to reality.

7.1. Future Work

This section describes the areas of improvement of this project that could have been implemented with more time, more usability testing iterations and an increased budget.

On the VRFTE system, as Ronell (2020) explained, the allocation of defined and structured development processes would require the involvement of typical users, meaning the inclusion of expert reviews in different iterations of usability testing within the development and production cycles. Additionally, the complexity of aviation simulation requires the reliability of the final product and involvement of high-quality 3D models, in which official sources and real-world capture of textures and material could significantly enhance the sense of immersion in the simulated scenarios (Jafer & Durak, 2017).

Although the implementation of hand tracking helped improve the sense of presence in the system, an increase of the budget would have allowed the acquisition of haptic feedback gloves. Functionality within the design could have been developed to elevate the immersion and the usability interactions of the elements.

Finally, one of the emerging usability evaluation methodologies is automated Task Tree Generation. By acquiring, organising and processing the dataset logs of the user session, could allow for analysis with machine learning algorithms that generate prediction patterns (Harms, 2019). However, this approach is in its early stages, and for the issue of usability some of the details and findings from the subjective perceptions and human interactions observations required that a manual input from a facilitator is provided. Therefore, this novel approach requires more profound research and testing.

REFERENCES

Abulfaraj, A., & Steele, A. (2020). *Detailed Usability Heuristics: A Breakdown of Usability Heuristics to Enhance Comprehension for Novice Evaluators*, Cham.

Administration, F. A. (2021). *Flight Simulation Training Device Qualification Guidance*. United States: United States Department of Transportation Retrieved from https://www.faa.gov/about/initiatives/nsp/ac/

Aleem, S., Capretz, L. F., & Ahmed, F. (2016). Game development software engineering process life cycle: a systematic review. *Journal of software engineering research and development, 4*(1), 1-30. doi:10.1186/s40411-016-0032-7

Operational Standards and Requirments - Approved Flight Simulators (FSD-1), Part 60 C.F.R. (2016).

Azadegan, A., & Riedel, J. C. K. H. (2012). Serious Games Integration in Companies: A Research and Application Framework. In (pp. 485-487): IEEE.

Azanha, A., Argoud, A. R. T. T., Camargo Junior, J. B. d., & Antoniolli, P. D. (2017). Agile project management with Scrum. *International journal of managing projects in business, 10*(1), 121-142. doi:10.1108/IJMPB-06-2016-0054

Bevan, N., Carter, J., & Harker, S. (2015). *ISO 9241-11 revised: What have we learnt about usability since 1998?* Paper presented at the International Conference on Human-Computer Interaction.

Bunnenberg, C. (2018). Virtual Time Travels? Public History and Virtual Reality. *Public History Weekly, 2018*(3). doi:10.1515/phw-2018-10896

Buscombe, J. (2016). Cessna 172 The Most Popular Light Aircraft in History. Retrieved from https://airshare.com/blog/cessna-172/

Chittaro, L., & Buttussi, F. (2015). Assessing Knowledge Retention of an Immersive Serious Game vs. a Traditional Education Method in Aviation Safety. *IEEE Transactions on Visualization and Computer Graphics, 21*(4), 529-538. doi:10.1109/TVCG.2015.2391853

Company, C. A. (2008). *Cessna Model 172S Operational Manual*. Wichita, Kansas, USA: Cessna Aircraft Company.

de Armas, C., Tori, R., & Netto, A. V. (2020). Use of virtual reality simulators for training programs in the areas of security and defense: a systematic review. *Multimedia Tools and Applications, 79*(5-6), 3495-3515. doi:10.1007/s11042-019-08141-8

Dunn, P. G., & Hayes, A. (2020, July 19-24). *The Problems with Usability Testing*, Cham.

Eller, C., Bittner, T., Dombois, M., & Rüppel, U. (2018). *Collaborative Immersive Planning and Training Scenarios in VR*, Cham.

Engine, U. (2020). Make Something Unreal. Retrieved from https://www.unrealengine.com/en-US/

Faiella, G. (2006). *The technology of Mesopotamia*: The Rosen Publishing Group, Inc.

Feng, Z., González, V. A., Amor, R., Lovreglio, R., & Cabrera-Guerrero, G. (2018). Immersive virtual reality serious games for evacuation training and research: A systematic literature review. *Computers and education, 127*, 252-266. doi:10.1016/j.compedu.2018.09.002

Fernandez, A., Insfran, E., Abrahão, S., Carsí, J. Á., & Montero, E. (2012). Integrating Usability Evaluation into Model-Driven Video Game Development. In (Vol. 7623, pp. 307-314). Berlin, Heidelberg: Berlin, Heidelberg: Springer Berlin Heidelberg.

Geisen, E. (2017). *Usability testing for survey research* (1st edition ed.): Cambridge, Massachusetts : Morgan Kaufmann.

Ha, G., Lee, H., Lee, S., Cha, J., & Kim, S. (2016). *A VR serious game for fire evacuation drill with synchronized tele-collaboration among users*. Paper presented at the Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology, Munich, Germany. https://doi-org.ezproxy.flinders.edu.au/10.1145/2993369.2996306

Harms, P. (2019). Automated Usability Evaluation of Virtual Reality Applications. *ACM transactions on computer-human interaction, 26*(3), 1-36. doi:10.1145/3301423
Higuchi, M. M., & Nakano, D. N. (2017). Agile Design: A Combined Model Based on Design Thinking and Agile Methodologies for Digital Games Projects. *Revista de Gestão e Projetos, 8*(2), 109-126. doi:10.5585/gep.v8i2.528
Hookham, G., Nesbitt, K., & Kay-Lambkin, F. (2016). *Comparing usability and engagement between a serious game and a traditional online program*. Paper presented at the Proceedings of the Australasian Computer Science Week Multiconference, Canberra, Australia. https://doi-org.ezproxy.flinders.edu.au/10.1145/2843043.2843365
International, F. S. (2020). Cessna Citation FFS Simulator. Retrieved from https://resources.flightsafety.com/cessna-citation-iii-vi-vii/cessna-citation-iii-vi-vii-simulator-profiles/
Jafer, S., & Durak, U. (2017). *Tackling the complexity of simulation scenario development in aviation*. Paper presented at the Proceedings of the Symposium on Modeling and Simulation of Complexity in Intelligent, Adaptive and Autonomous Systems, Virginia Beach, Virginia.
Karre, S., Mathur, N., & Reddy, Y. (2020). Understanding usability evaluation setup for VR products in industry: a review study. *Applied computing review : a publication of the Special Interest Group on Applied Computing, 19*(4), 17-27. doi:10.1145/3381307.3381309
Kharoufah, H., Murray, J., Baxter, G., & Wild, G. (2018). A review of human factors causations in commercial air transport accidents and incidents: From to 2000–2016. *Progress in Aerospace Sciences, 99*, 1-13.
Kilteni, K., Groten, R., & Slater, M. (2012). The Sense of Embodiment in Virtual Reality. *Presence (Cambridge, Mass.), 21*(4), 373-387. doi:10.1162/PRES_a_00124
Lewis, J. R. (2018). The System Usability Scale: Past, Present, and Future. *International journal of human-computer interaction, 34*(7), 577-590. doi:10.1080/10447318.2018.1455307
Manuel, P. C. V., José, P. C. I., Manuel, F.-M., Iván, M.-O., & Baltasar, F.-M. (2019). Simplifying the Creation of Adventure Serious Games with Educational-Oriented Features. *Educational technology & society, 22*(3), 32-46.
Marcus, A., & Rosenzweig, E. (2020). *Design, User Experience, and Usability. Design for Contemporary Interactive Environments 9th International Conference, DUXU 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part II* (1st ed. 2020. ed.): Cham : Springer International Publishing : Imprint: Springer.
Martins, V. F., Kirner, T. G., & Kirner, C. (2015). *Subjective Usability Evaluation Criteria of Augmented Reality Applications*, Cham.
Menin, A., Torchelsen, R., & Nedel, L. (2018). An Analysis of VR Technology Used in Immersive Simulations with a Serious Game Perspective. *IEEE Computer Graphics and Applications, 38*(2), 57-73. doi:10.1109/MCG.2018.021951633
Miglani, A., Kidambi, S., & Mareguddi, P. (2020). Play to Improve: Gamifying Usability Evaluations in Virtual Reality. In (Vol. 12423, pp. 248-266). Cham: Cham: Springer International Publishing.
Nielsen, J. (1994). *Usability engineering*: Morgan Kaufmann.
NNgroup (Producer). (2018, September 28, 2021). Design Process. *Usability with 5 users*. Retrieved from https://www.youtube.com/watch?v=RhgUirqki50
NNgroup (Producer). (2019, September 28, 2021). Help Users Recognize, Diagnose and Recover from Errors. *Usability Heuristics*. Retrieved from https://www.youtube.com/watch?v=cCun-ReLTFI
O'Dea, S. (2021). Number of smartphone users worldwide from 2016 to 2026. Retrieved from https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/
Pallavicini, F., Pepe, A., & Minissi, M. E. (2019). Gaming in Virtual Reality: What Changes in Terms of

Usability, Emotional Response and Sense of Presence Compared to Non-Immersive Video Games? *Simulation & gaming, 50*(2), 136-159. doi:10.1177/1046878119831420

Palmas, F., Labode, D., Plecher, D. A., & Klinker, G. (2019, 4-6 Sept. 2019). *Comparison of a Gamified and Non-Gamified Virtual Reality Training Assembly Task.* Paper presented at the 2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games).

Petar, Č., & Sanja Maravić, Č. (2020). DEVELOPMENT CONCEPTS OF VIRTUAL REALITY SOFTWARE. *Acta Technica Corvininesis, 13*(3), 23-29.

Ronell, M. (2020). *Discussion of aviation software oversight improvement*. Paper presented at the Proceedings of the 2020 ACM SIGPLAN International Symposium on New Ideas, New Paradigms, and Reflections on Programming and Software, Virtual, USA. https://doi-org.ezproxy.flinders.edu.au/10.1145/3426428.3426926

Sagar, K., & Saha, A. (2017). Qualitative usability feature selection with ranking: a novel approach for ranking the identified usability problematic attributes for academic websites using data-mining techniques. *Human-centric computing and information sciences, 7*(1), 1-24. doi:10.1186/s13673-017-0111-8

Sauro, J., & Lewis, J. R. (2016). Chapter 8 - Standardized usability questionnaires. In J. Sauro & J. R. Lewis (Eds.), *Quantifying the User Experience (Second Edition)* (pp. 185-248). Boston: Morgan Kaufmann.

Sparks, M. (2017). Skills Needed to Create a Serious VR Game with a Game Engine. Retrieved from https://learningsolutionsmag.com/articles/2396/skills-needed-to-create-a-serious-vr-game-with-a-game-engine

Spiliotopoulos, K., Rigou, M., & Sirmakessis, S. (2018). A comparative study of skeuomorphic and flat design from a ux perspective. *Multimodal Technologies and Interaction, 2*(2), 31. doi:10.3390/mti2020031

Statista. (2021). Virtual Reality Headset unit sales worldwide from 2019 to 2024. Retrieved from https://www.statista.com/statistics/677096/vr-headsets-worldwide/

Stein, S. (2021). Mark Zuckerberf on Facebook's VR future: New sensors on Quest Pro, fitness and a metaverse for work. *CNET.* Retrieved from https://www.cnet.com/features/mark-zuckerberg-on-facebook-vr-future-new-sensors-on-quest-pro-fitness-and-a-metaverse-for-work/

Swan, J. E., & Gabbard, J. L. (2014, 29 March-2 April 2014). *Quantitative and qualitative methods for human-subject experiments in Virtual and Augmented Reality.* Paper presented at the 2014 IEEE Virtual Reality (VR).

Tom, P.-W., Chris, W., Hyett, P. J., & Scott, C. (2008, August 25, 2021). Where the world builds software. 2.9.3. Retrieved from https://github.com/

USGS. (2021). Oceania South Sector GIS. Retrieved from https://www.usgs.gov/products/maps/topo-maps

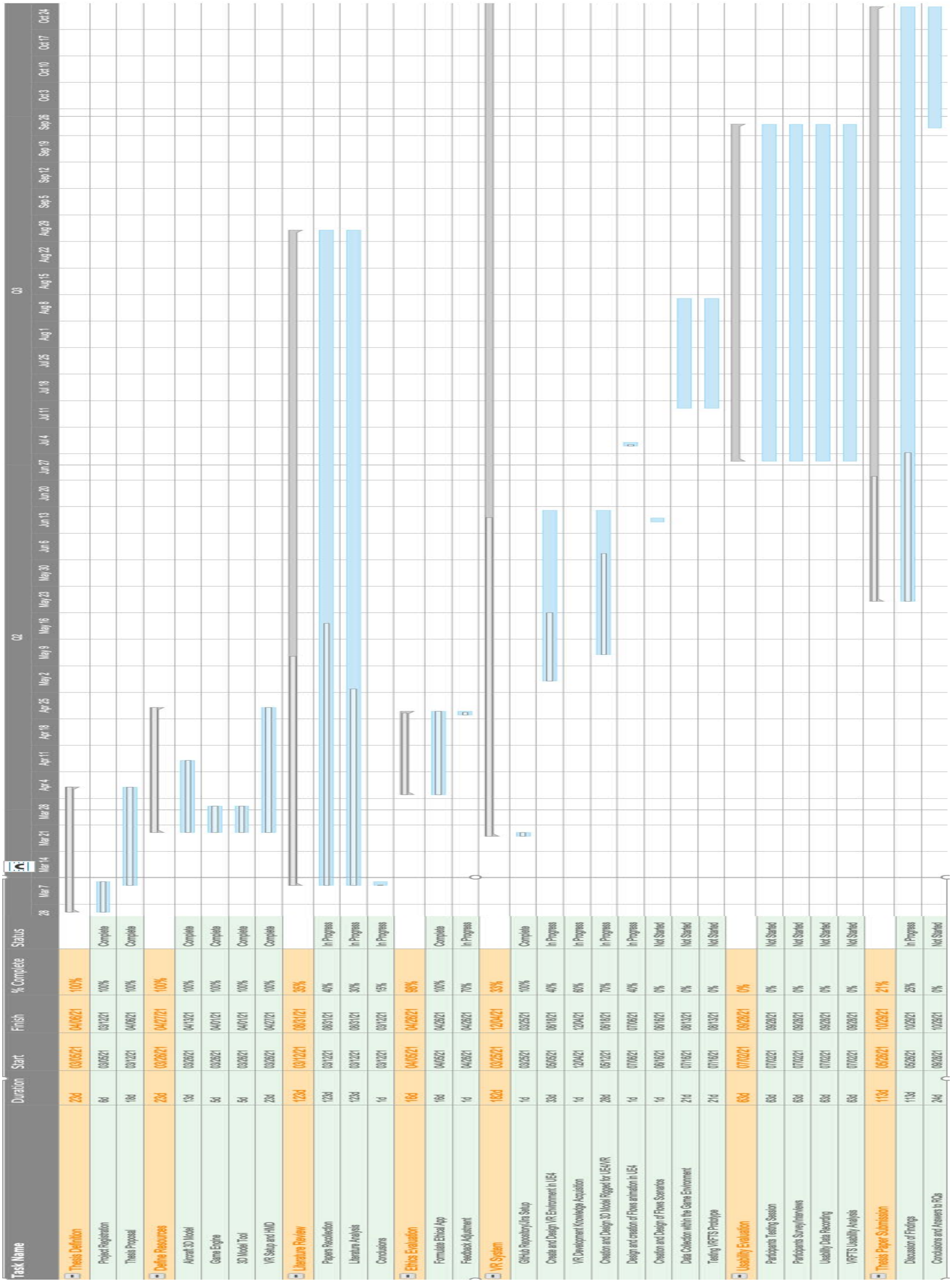
Wang, S., Li, B., & Zhu, Y. (2019). Comprehensive evaluation of usability at the mobile end interface. *IOP Conference Series: Materials Science and Engineering, 573*, 012037. doi:10.1088/1757-899x/573/1/012037

Wu, B., Yu, X., & Gu, X. (2020). Effectiveness of immersive virtual reality using head-mounted displays on learning performance: A meta-analysis. *British journal of educational technology, 51*(6), 1991-2005. doi:10.1111/bjet.13023

Yanez-Gomez, R., Cascado-Caballero, D., & Sevillano, J.-L. (2017). Academic methods for usability evaluation of serious games: a systematic review. *Multimedia Tools and Applications, 76*(4), 5755-5784.

APPENDIXES

Appendix A:



Task Name	Duration	Start	Finish	% Complete	Status
Thesis Definition	23d	03/05/21	04/06/21	100%	
Project Registration	6d	03/05/21	03/12/21	100%	Complete
Thesis Proposal	18d	03/12/21	04/06/21	100%	Complete
Define Resources	23d	03/26/21	04/27/21	100%	
Aircraft 3D Model	13d	03/26/21	04/13/21	100%	Complete
Game Engine	5d	03/26/21	04/01/21	100%	Complete
3D Model Tool	5d	03/26/21	04/01/21	100%	Complete
VR Setup and HMD	23d	03/26/21	04/27/21	100%	Complete
Literature Review	123d	03/12/21	08/31/21	35%	
Papers Recollection	123d	03/12/21	08/31/21	40%	In Progress
Literature Analysis	123d	03/12/21	08/31/21	30%	In Progress
Conclusions	1d	03/12/21	03/12/21	15%	In Progress
Ethics Evaluation	16d	04/05/21	04/26/21	98%	
Formulate Ethical App	16d	04/05/21	04/26/21	100%	Complete
Feedback Adjustment	1d	04/26/21	04/26/21	70%	In Progress
VR System	182d	03/25/21	12/04/21	33%	
GitHub Repository/Jira Setup	1d	03/25/21	03/25/21	100%	Complete
Create and Design VR Environment in UE4	33d	05/05/21	06/18/21	40%	In Progress
VR Development Knowledge Acquisition	1d	12/04/21	12/04/21	60%	In Progress
Creation and Design 3D Model Rigged for UE4/VR	28d	05/12/21	06/18/21	70%	In Progress
Design and creation of Flows animation in UE4	1d	07/06/21	07/06/21	40%	In Progress
Creation and Design of Flows Scenarios	1d	06/16/21	06/16/21	0%	Not Started
Data Collection within the Game Environment	21d	07/16/21	08/13/21	0%	Not Started
Testing VRFTS Prototype	21d	07/16/21	08/13/21	0%	Not Started
Usability Evaluation	63d	07/02/21	09/28/21	0%	
Participants Testing Session	63d	07/02/21	09/28/21	0%	Not Started
Participants Survey/Interviews	63d	07/02/21	09/28/21	0%	Not Started
Usability Data Recording	63d	07/02/21	09/28/21	0%	Not Started
VRFTS Usability Analysis	63d	07/02/21	09/28/21	0%	Not Started
Thesis Paper Submission	113d	05/26/21	10/29/21	21%	
Discussion of Findings	113d	05/26/21	10/29/21	25%	In Progress
Conclusions and Answers to RQs	24d	09/28/21	10/29/21	0%	Not Started

Appendix B:

Demographics Survey

The following survey aims to understand the participant's previous experience, in order to position the project Virtual Reality Flight Training Experience in the current state of VR. In addition, this project has been approved by Flinders University's Human Research Ethics Committee to comply with the privacy and safety of the participants.

* Required

1. Unique ID *

2. Age *

- ◯ 18-24
- ◯ 25-32
- ◯ 30-40
- ◯ >40

3. What is your experience playing/interacting with a Flight Simulator? *

	1	2	3	4	5	
Without Experience	◯	◯	◯	◯	◯	Experience

6/15/2021

4. Have you used or played any VR game/application? *

- ◯ Yes
- ◯ No

5. If you have experienced VR, How comfortable you feel with the use of this technology? *

	1	2	3	4	5	
Strongly uncomfortable	◯	◯	◯	◯	◯	Strongly comfortable

6. Educational Institution *

- ◯ Flinders University
- ◯ University of South Australia
- ◯ Adelaide University
- ◯ Torrens University
- ◯ FTA
- ◯ Hartwig Air
- ◯ Aero Star Aviation Academy
- ◯ Other

7. Degree *

Appendix C: Consent Form and Ethics Letter of Approval.

Flinders
UNIVERSITY
ADELAIDE • AUSTRALIA

PARTICIPANT INFORMATION SHEET AND CONSENT FORM

Title: Evaluating the Usability of a Virtual Reality Flight Training Experience Application

Chief Investigator
Dr Brett Wilkinson
College of Science and Engineering
Flinders University
Tel: +61 8 82013662
Email: brett.wilkinson@flinders.edu.au

Co-Investigator
Mr Luis Morales Trujillo
College of Science and Engineering
Flinders University
Email: mora0157@flinders.edu.au

My name is Luis Morales Trujillo, and I am a Flinders University Master's student. I am undertaking this research as part of my degree. For further information, you are more than welcome to contact my supervisor (Chief-Investigator) or myself (Co-Investigator). Contact details listed above.

Description of the study
This project will investigate the overall usability of a Virtual Reality serious game titled VRFTE different usability evaluation methods. These usability evaluation methods will also be measured to determine which one is the most applicable to Virtual Reality applications/serious games.

VRFTE is a Virtual Reality serious game that has been developed at Flinders University and has several training, knowledge and skill acquisition goals aimed towards staff and students at Flinders University. The VR experience is designed to train fundamental flows in an Aircraft, improving skills and giving the necessary knowledge to perform certain operational tasks, evaluating the success of this type of VR training environments. This project is supported by Flinders University, College of Science and Engineering.

Purpose of the study
This project aims to find out the overall usability of the VRFTE serious game to make it fit for release to staff and students at Flinders University. It also aims to determine the most applicable usability evaluation method for Virtual Reality serious games. As a response to the lack, of an effective usability evaluation method for mobile serious games, and mobile serious game designers currently utilise methods that may be obsolete or do not apply to mobile serious games.

inspiring achievement

This research aims to evaluate the usability perception of an aviator and non-aviator participants in a VR training environment by the execution of defined tasks.
For the resolution and analysis of this study, a VR serious game has been prototyped based on the Cessna T41-Mescalero (C172 Skyhawk) cockpit, which is one of the most commonly used aircraft for primary flight training. Consequently, a series of normal and emergency procedures are designed within the scope of tasks scenarios, with the purpose to engage participants to perform the required steps correctly.
A testing session per participant will be recorded, to obtain their performance data and feedback. The latter will be through the use of usability evaluation surveys, which would be aimed to record the individual motives and qualitative opinion on the system.

Benefits of the study
Your participation will help to determine the overall usability of VRFTE in training environments that are based on the aircraft scenario presented but not restricted to only this. It will also determine the most applicable usability evaluation method for Virtual Reality applications/serious games. This in turn will aid in setting a benchmark for what counts as a usable VR serious game aimed for knowledge acquisition, skill and training personnel.

Participant involvement and potential risks
If you agree to participate in the research study, you will be asked to:

- Complete a Participant Registration Survey.
- To use a Virtual Reality headset.
- To complete the tutorial and task within the VRFTE environment.
- Complete two questionnaires on a computer that ask about your experience with VRFTE, and your perception with usability questionnaires.

Participation is entirely voluntary.

There will be a negligible risk to you the participant, which could be represented as mild motion sickness that could generate nausea, nevertheless, this case should be a rare side effect due to the nature of the virtual reality experience; you will be told beforehand to cease the training module if there is the slightest discomfort.

The Oculus Quest 2 is the Head Mounted Display (HMD) hardware chosen for this project, there are certain considerations that are outlined by the Oculus Health and Safety Warnings section of the manual that you need to be aware of, such as an approximate probability of 1 in 4000 people might have seizures. This is a consequence of light flashes/patterns, and could be presented while watching TV, playing video games or in this case experiencing Virtual Reality. Therefore, if you have had a seizure, loss of awareness, or other symptom linked to an epileptic condition, you should consult with your general practitioner before using this technology.

Consequently, you will be asked to engage with the Virtual Reality Headset, the surveys on a computer webform, and if voluntary selected a final verbal interview of your experience.

The researchers do not expect the questions from the multiple usability questionnaires and Participant registration Survey to cause any harm or discomfort to you. However, if you experience feelings of distress as a result of participation in this study, please let the research team know immediately. You can also contact the following services for support:

- Lifeline – 13 11 14, www.lifeline.org.au
- Beyond Blue – 1300 22 4636, www.beyondblue.org.au

Withdrawal Rights
You may, without any penalty, decline to take part in this research study. If you decide to take part and later change your mind, you may, without any penalty, withdraw at any time without explaining. To withdraw, please contact the Chief Investigator or Co-Investigator, or you may just refuse to answer any questions / close the internet browser and leave the online survey / not participate in exercises at any time. Any data collected up to the point of your withdrawal will be securely destroyed.

My decision not to participate or to withdraw from this research study will not affect my relationship with Flinders University and its staff and students.

Confidentiality and Privacy
Only researchers listed on this form have access to the individual information provided by you. Privacy and confidentiality will be assured at all times. The research outcomes may be presented at conferences, written up for publication or used for other research purposes as described in this information form. However, the privacy and confidentiality of individuals will be protected at all times. You will not be named, and your individual information will not be identifiable in any research products without your explicit consent.

No data, including identifiable, non-identifiable and de-identified datasets, will be shared or used in future research projects without your explicit consent.

Data Storage
The information collected may be stored securely on a password-protected computer and/or Flinders University server throughout the study. Any identifiable data will be de-identified for data storage purposes unless indicated otherwise. All data will be securely transferred to and stored at Flinders University for at least five years after the publication of the results. Following the required data storage period, all data will be securely destroyed according to university protocols.

How will I receive feedback?
mailto:human.researchethics@flinders.edu.au
On project completion, a summary of the outcomes will be provided to all participants via email upon request by the participant.

Ethics Committee Approval
Once conditional approval has been provided, Prof Clare Pollock's office will be contacted to gain permission to recruit Flinders Students.

Queries and Concerns
Queries or concerns regarding the research can be directed to the research team. If you have any complaints or reservations about the ethical conduct of this study, you may contact the Flinders University's Research Ethics & Compliance Office team via telephone 08 8201 3116 or email human.researchethics@flinders.edu.au.

Thank you for taking the time to read this information sheet which is yours to keep. If you accept our invitation to be involved, please sign the enclosed Consent Form.

CONSENT FORM

Consent Statement

- ☐ I have read and understood the information about the research, and I understand I am being asked to provide informed consent to participate in this research study. I understand that I can contact the research team if I have further questions about this research study.
- ☐ I am not aware of any condition that would prevent my participation, and I agree to participate in this project.
- ☐ I understand that I am free to withdraw at any time during the study and that my withdrawal will not affect my relationship with Flinders University and its staff and students.
- ☐ I understand that I can contact Flinders University's Research Ethics & Compliance Office if I have any complaints or reservations about the ethical conduct of this study.
- ☐ I understand that my involvement is confidential, and that the information collected may be published. I understand that I will not be identified in any research products.

I further consent to:

- ☐ completing usability questionnaires.
- ☐ completing a Participant Background Survey.

Signed:

Name:

Date:

18 June 2021

HUMAN ETHICS LOW RISK PANEL

APPROVAL NOTICE

Dear Dr Brett Wilkinson,

The below proposed project has been **approved** on the basis of the information contained in the application and its attachments.

Project No:	4487
Project Title:	What is the usability perception of a VR Flight Training System?
Primary Researcher:	Dr Brett Wilkinson
Approval Date:	18/06/2021
Expiry Date:	30/10/2021

Please note: *Due to the current COVID-19 situation, researchers are strongly advised to develop a research design that aligns with the University's COVID-19 research protocol involving human studies. Where possible, avoid face-to-face testing and consider rescheduling face-to-face testing or undertaking alternative distance/online data or interview collection means. For further information, please go to https://staff.flinders.edu.au/coronavirus-information/research-updates.*

RESPONSIBILITIES OF RESEARCHERS AND SUPERVISORS

1. Participant Documentation

Please note that it is the responsibility of researchers and supervisors, in the case of student projects, to ensure that:

- all participant documents are checked for spelling, grammatical, numbering and formatting errors. The Committee does not accept any responsibility for the above mentioned errors.
- the Flinders University logo is included on all participant documentation (e.g., letters of Introduction, information Sheets, consent forms, debriefing information and questionnaires – with the exception of purchased research tools) and the current Flinders University letterhead is included in the header of all letters of introduction. The Flinders University international logo/letterhead should be used and documentation should contain international dialing codes for all telephone and fax numbers listed for all research to be conducted overseas.

2. Annual Progress / Final Reports

In order to comply with the monitoring requirements of the *National Statement on Ethical Conduct in Human Research 2007 (updated 2018)* an annual progress report must be submitted each year on the approval anniversary date for the duration of the ethics approval using the HREC Annual/Final Report Form available online via the ResearchNow Ethics & Biosafety system.

Please note that no data collection can be undertaken after the ethics approval expiry date listed at the top of this notice. If data is collected after expiry, it will not be covered in terms of ethics. It is the responsibility of the researcher to ensure that annual progress reports are submitted on time; and that no data is collected after ethics has expired.

If the project is completed *before* ethics approval has expired please ensure a final report is submitted immediately. If ethics approval for your project expires please either submit (1) a final report; or (2) an extension of time request (using the HREC Modification Form).

For student projects, the Low Risk Panel recommends that current ethics approval is maintained until a student's thesis has been submitted, assessed and finalised. This is to protect the student in the event that reviewers recommend that additional data be collected from participants.

3. Modifications to Project

Modifications to the project must not proceed until approval has been obtained from the Ethics Committee. Such proposed changes / modifications include:

- change of project title;
- change to research team (e.g., additions, removals, researchers and supervisors)
- changes to research objectives;
- changes to research protocol;
- changes to participant recruitment methods;
- changes / additions to source(s) of participants;
- changes of procedures used to seek informed consent;
- changes to reimbursements provided to participants;
- changes to information / documents to be given to potential participants;
- changes to research tools (e.g., survey, interview questions, focus group questions etc);
- extensions of time (i.e. to extend the period of ethics approval past current expiry date).

To notify the Committee of any proposed modifications to the project please submit a Modification Request Form available online via the ResearchNow Ethics & Biosafety system. Please note that extension of time requests should be submitted prior to the Ethics Approval Expiry Date listed on this notice.

4. Adverse Events and/or Complaints

Researchers should advise the Executive Officer of the Ethics Committee on 08 8201-3116 or human.researchethics@flinders.edu.au immediately if:

- any complaints regarding the research are received;
- a serious or unexpected adverse event occurs that effects participants;
- an unforeseen event occurs that may affect the ethical acceptability of the project.

Yours sincerely,

Ms Camilla Dorian

on behalf of

Human Ethics Low Risk Panel
Research Development and Support
human.researchethics@flinders.edu.au
P: (+61-8) 8201 2543

Flinders University
Sturt Road, Bedford Park, South Australia, 5042
GPO Box 2100, Adelaide, South Australia, 5001

http://www.flinders.edu.au/research/researcher-support/ebi/human-ethics/human-ethics_home.cfm

ResearchNow
Ethics & Biosafety

Flinders University Proactively supporting our Research

Appendix D:

VRFTE Usability-Interview Questionnaire
Participant ID:

Usability testing interview form.

1. Unique ID of the participant:

Enter your answer

2. Could you describe in few words your first impressions of the experience with this VR training application?

Enter your answer

3. Did you find any particular difficulties using this type of Hardware?

Enter your answer

4. On the other hand, briefly what are your throughs on the User Interface and the cockpit layout?

Enter your answer

5. Do you believe the experience was accurate enough to be used as a training tool in the aviation industry? and why?

Enter your answer

6. Is there anything you will change or improve in the VR experience?

Enter your answer

7. Would you be willing to use this type of VR application for learning a skill or training in other fields? Could you provide an example?

Enter your answer

Appendix E:

Figure 18 - VRFTE Tutorial

Appendix F:

VRFTE Usability Questionnaire

Participant ID:

Usability testing evaluation form of participants perception by the end test session.

1. It was easy to adapt myself to this VR environment

	1	2	3	4	5	6	7	8	9	10	
Strongly Disagree	○	○	○	○	○	○	○	○	○	○	Strongly Agree

2. The use of VR hardware was simple and straightforward

	1	2	3	4	5	6	7	8	9	10	
Strongly Disagree	○	○	○	○	○	○	○	○	○	○	Strongly Agree

3. It was simple to use this application

	1	2	3	4	5	6	7	8	9	10	
Strongly Disagree	○	○	○	○	○	○	○	○	○	○	Strongly Agree

4. I was able to understand the tutorial completely

	1	2	3	4	5	6	7	8	9	10	
Strongly Disagree	○	○	○	○	○	○	○	○	○	○	Strongly Agree

5. It was easy to learn how to use the VR Interface

	1	2	3	4	5	6	7	8	9	10	
Strongly Disagree	○	○	○	○	○	○	○	○	○	○	Strongly Agree

6. I felt Comfortable using this VR training experience

	1	2	3	4	5	6	7	8	9	10	
Strongly Disagree	○	○	○	○	○	○	○	○	○	○	Strongly Agree

7. The tutorial part gave me the knowledge that I needed the training session

	1	2	3	4	5	6	7	8	9	10	
Strongly Disagree	○	○	○	○	○	○	○	○	○	○	Strongly Agree

8. The difficulty of the flight training tasks was adequate for my learning experience

	1	2	3	4	5	6	7	8	9	10	
Strongly Disagree	○	○	○	○	○	○	○	○	○	○	Strongly Agree

9. I believe performing this training using this VR experience regularly will benefit my knowledge in flight training flows

	1	2	3	4	5	6	7	8	9	10	
Strongly Disagree	○	○	○	○	○	○	○	○	○	○	Strongly Agree

10. The information and instructions provided by the facilitator were clear

	1	2	3	4	5	6	7	8	9	10	
Strongly Disagree	○	○	○	○	○	○	○	○	○	○	Strongly Agree

11. The interface and immersive environment were pleasant

	1	2	3	4	5	6	7	8	9	10	
Strongly Disagree	○	○	○	○	○	○	○	○	○	○	Strongly Agree

12. The organization of the information and features of the cockpit were clear

	1	2	3	4	5	6	7	8	9	10	
Strongly Disagree	○	○	○	○	○	○	○	○	○	○	Strongly Agree

13. I liked using this VR training application

	1	2	3	4	5	6	7	8	9	10	
Strongly Disagree	○	○	○	○	○	○	○	○	○	○	Strongly Agree

14. This VR experience has all the functions and capabilities I expect it to have.

	1	2	3	4	5	6	7	8	9	10	
Strongly Disagree	○	○	○	○	○	○	○	○	○	○	Strongly Agree

15. I think that I would like to use this system again for similar learning purposes

	1	2	3	4	5	6	7	8	9	10	
Strongly Disagree	○	○	○	○	○	○	○	○	○	○	Strongly Agree

16. I found the VR application unnecessarily complex

	1	2	3	4	5	6	7	8	9	10	
Strongly Disagree	○	○	○	○	○	○	○	○	○	○	Strongly Agree

17. I would imagine that most people would learn to use this application very quickly

	1	2	3	4	5	6	7	8	9	10	
Strongly Disagree	○	○	○	○	○	○	○	○	○	○	Strongly Agree

18. I found the VR experience very awkward or uncomfortable

	1	2	3	4	5	6	7	8	9	10	
Strongly Disagree	○	○	○	○	○	○	○	○	○	○	Strongly Agree

19. I think there was not enough tutorial time or instructions to help to adapt to the tasks to be performed

	1	2	3	4	5	6	7	8	9	10	
Strongly Disagree	○	○	○	○	○	○	○	○	○	○	Strongly Agree

20. Overall, I am satisfied with this VR training experience

	1	2	3	4	5	6	7	8	9	10	
Strongly Disagree	○	○	○	○	○	○	○	○	○	○	Strongly Agree

Appendix G: VRFTE in-game video – Flow 3

Figure 19 - VRFTE Flow 3 Interaction

Appendix H: VRFTE Dataset of Usability Questionnaire

Age	ID	EXP	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q17	Q16	Q18	Q19	Q20	Positive	Negative
25-30	**VRYC1**	No XP	9	9	9	9	9	9	8	9	9	9	9	8	10	8	10	9	3	3	3	9	8.94	3.00
18-24	**VRSY2**	Flight XP	5	3	6	2	4	2	3	3	3	10	7	7	6	4	4	7	2	7	7	4	4.71	5.33
31-40	**VRTU3**	No XP	9	10	8	6	9	10	9	7	10	10	10	9	10	10	9	9	2	3	2	10	9.12	2.33
31-40	**VRTU4**	VR XP	8	10	9	10	10	10	9	9	10	10	10	10	10	10	10	9	4	3	4	9	9.59	3.67
31-40	**VRAK5**	No XP	5	7	8	8	7	9	9	6	8	10	10	9	10	10	10	9	2	2	2	10	8.53	2.00
18-24	**VRAI6**	VR XP	6	9	9	10	10	10	9	8	8	9	9	9	9	9	9	8	6	1	1	10	8.88	2.67
31-40	**VREC7**	VR XP	7	8	9	9	8	9	8	10	10	10	10	10	10	9	9	8	2	2	3	9	9.00	2.33
18-24	**VRAW8**	VR XP	9	10	9	10	10	9	9	10	5	10	10	9	10	9	10	10	1	2	1	10	9.35	1.33
25-30	**VRCD9**	No XP	10	10	10	10	10	10	10	8	10	10	10	10	10	10	10	10	1	1	1	10	9.88	1.00
25-30	**PTAK1**	No XP	8	7	7	8	9	7	8	9	9	10	7	9	9	10	9	7	2	3	8	8	8.29	4.33
25-30	**PTAK2**	VR XP	8	9	8	7	8	9	8	8	8	9	7	9	9	10	9	8	2	2	8	9	8.41	4.00
25-30	**PTAC7**	No XP	6	7	8	8	8	6	7	7	8	10	8	7	8	7	10	6	1	1	1	8	7.59	1.00
18-24	**VRTD10**	VR & Flight XP	8	8	8	10	8	8	10	9	10	8	8	8	9	9	7	8	2	3	8	10	8.59	4.33
31-40	**VRPM11**	VR & Flight XP	10	9	10	10	8	8	9	7	8	9	8	8	9	7	9	7	1	2	9	9	8.53	4.00
>41	**VRTA12**	Flight XP	6	6	7	10	7	7	9	5	8	9	7	9	9	9	8	7	3	3	2	8	7.71	2.67

EXP AVG	Positive	Negative
No XP	8.72549	2.277778
VR XP	9.047059	2.8
Flight XP	6.205882	4
VR & Fligh	8.558824	4.166667

Appendix I: VRFTE Interactions Logs – Flow 5

FLOW	Five
TRAINING OR EVALUATION	Training Activated
Participant ID	**Num. Interactions**
PTAK1	**7**
Iddle ThrottlePosition	1
KeySwitch Cooling Off Engine	1
MasterSwitch Off	1
Mixture Iddle	1
The Parking Brake Is On	1
ThrottleFull	1
WingsFlaps Up	1
PTAC7	**30**
KeySwitch Cooling Off Engine	1
MasterSwitch Off	12
MasterSwitch On	11
Mixture Iddle	1
MixtureLean	1
MixtureRich	1
The Parking Brake Is On	1
ThrottleFull	1
WingsFlaps Up	1
PTAK2	**19**
Iddle ThrottlePosition	1
KeySwitch Cooling Off Engine	1
MasterSwitch Off	7
MasterSwitch On	6
Mixture Iddle	1
The Parking Brake Is On	1
ThrottleFull	1
Trottle1/2	1
VRAI6	**15**
Avionics Off	3
Avionics On	3
FuelPumpOff	1
KeySwitch Cooling Off Engine	1
KeySwitch Turning On Engine	1
MasterSwitch Off	2
MasterSwitch On	2
The Parking Brake Is On	1
ThrottleFull	1
VRAK5	**10**
KeySwitch Cooling Off Engine	1
MasterSwitch Off	3

MasterSwitch On	2
Mixture Iddle	1
MixtureLean	1
The Parking Brake Is On	1
ThrottleFull	1
VRAW8	**14**
Avionics Off	4
Avionics On	3
Iddle ThrottlePosition	1
KeySwitch Cooling Off Engine	1
MasterSwitch Off	1
Mixture Iddle	1
Taxi Lights On	1
The Parking Brake Is On	1
ThrottleFull	1
VRCD9	**18**
Iddle ThrottlePosition	1
KeySwitch Cooling Off Engine	1
MasterSwitch Off	2
MasterSwitch On	1
Mixture Iddle	3
MixtureLean	2
MixtureRich	2
The Parking Brake Is On	1
ThrottleFull	1
Transponder ON	1
TransponderOFF	1
WingsFlaps 10 Degrees	1
WingsFlaps Up	1
VREC7	**22**
Iddle ThrottlePosition	2
KeySwitch Cooling Off Engine	2
MasterSwitch Off	5
MasterSwitch On	3
Mixture Iddle	2
Navigation lights On	2
The Parking Brake Is On	2
ThrottleFull	2
WingsFlaps Up	2
VRSY2	**8**
Iddle ThrottlePosition	1
KeySwitch Cooling Off Engine	2
MasterSwitch Off	1
Mixture Iddle	1
The Parking Brake Is On	1
ThrottleFull	1
WingsFlaps Up	1

VRTA12	**7**
Iddle ThrottlePosition	1
KeySwitch Cooling Off Engine	1
MasterSwitch Off	1
Mixture Iddle	1
The Parking Brake Is On	1
ThrottleFull	1
WingsFlaps Up	1
VRPM11	**7**
Iddle ThrottlePosition	1
KeySwitch Cooling Off Engine	1
MasterSwitch Off	1
Mixture Iddle	1
The Parking Brake Is On	1
ThrottleFull	1
WingsFlaps Up	1
VRTD10	**7**
Iddle ThrottlePosition	1
KeySwitch Cooling Off Engine	1
MasterSwitch Off	1
Mixture Iddle	1
The Parking Brake Is On	1
ThrottleFull	1
WingsFlaps Up	1
VRTU3	**13**
Avionics Off	3
Avionics On	2
KeySwitch Cooling Off Engine	1
MasterSwitch Off	1
Mixture Iddle	2
Strober Lights On	1
The Parking Brake Is On	1
ThrottleFull	1
WingsFlaps Up	1
VRTU4	**34**
Iddle ThrottlePosition	1
KeySwitch Cooling Off Engine	1
MasterSwitch Off	10
MasterSwitch On	9
Mixture Iddle	2
MixtureLean	1
Moving the Yaw Helps visibility	3
Strober Lights On	1
The Parking Brake Is On	2
ThrottleFull	2
WingsFlaps 10 Degrees	1

WingsFlaps Up	1
VRYc1	**18**
Iddle ThrottlePosition	1
KeySwitch Cooling Off Engine	1
MasterSwitch Off	1
Mixture Iddle	2
MixtureLean	1
MixtureRich	1
Navigation lights On	1
The Parking Brake Is On	1
ThrottleFull	1
WingsFlaps 10 Degrees	2
WingsFlaps 20 Degrees	2
WingsFlaps Full	2
WingsFlaps Up	2
Grand Total	**229**