University of Derby

Evaluation of Detecting Cybersickness via VR HMD Positional Measurements Under Realistic Usage Conditions.

Patrick Merritt

A submission in partial fulfilment of the requirements of the University of Derby for the award of the degree of Doctor of Philosophy.

April 2021

Table of Contents

List of Figures

List of Tables

Preface

The research and writing undertaken within this document is wholly the work of Mr Patrick Merritt.

Ethical approval for this study was granted by the University of Derby ethics board under project number ETH2021-0165.

Acknowledgements

I would like to thank the following people who have had a significant impact in creating this thesis.

I am extremely grateful to Christopher Windmill as the guiding hand for the thesis, providing consistent advice and challenging my thought process throughout the project. Additional thanks are needed for the support offered regarding keeping my mental health in check and going beyond the traditional support role when my anxiety issues got the best of me. My gratitude extends to the support you have shown me as both a teacher and a friend.

I would also like to thank Yong Xue for their continual support in guiding the technical aspects of thesis structure and content and helping with writing my first significant publication. Thank you for the kick up the arse when I needed it most.

Additionally, many thanks go to Rich Conniss for his excellent guidance with the project's maths and statistics areas and for opening doors to exciting places. For guiding me through this without asking anything in return.

I gratefully acknowledge the assistance of Bradley Davis for having my back when it mattered most.

I also would like to thank Danielle Turvill because lunch has kept me sane.

I also appreciate my friends and family's love and support over the last ten years, who without I probably would not have achieved this.

Additionally, my heartfelt thanks go out to all the experiment participants, who without their help, none of this would have been possible.

https://xkcd.com/722/

Abstract

With the resurgence of virtual reality, head-mounted displays (VR HMD) technologies since 2015, VR technology is becoming ever more present in people's day-to-day lives. However, one significant barrier to this progress is a condition called cybersickness, a form of motion sickness induced by the usage of VR HMD's. It is often debilitating to sufferers, resulting in symptoms anywhere from mild discomfort to full-on vomiting. Much research effort focuses on identifying the cause of and solution to this problem, with many studies reporting various factors that influence cybersickness, such as vection and field of view. However, there is often disagreement in these studies' results and comparing the results is often complicated as stimuli used for the experiments vary wildly.

This study theorised that these results' mismatch might partially be down to the different mental loads of these tasks, which may influence cybersickness and stability-based measurement methods such as postural stability captured by the centre of pressure (COP) measurements. One recurring desire in these research projects is the idea of using the HMD device itself to capture the stability of the users head. However, measuring the heads position via the VR HMD is known to have inaccuracies meaning a perfect representation of the heads position cannot be measured.

This research took the HTC Vive headset and used it to capture the head position of multiple subjects experiencing two different VR environments under differing levels of cognitive load. The design of these test environments reflected normal VR usage. This research found that the VR HMD measurements in this scenario may be a suitable proxy for recording instability. However, the underlying method was greatly influenced by other factors, with cognitive load (5.4% instability increase between the low and high load conditions) and test order (2.4% instability decrease between first run and second run conditions) having a more significant impact on the instability recorded than the onset of cybersickness (2% instability increase between sick and well participants). Also, separating participants suffering from cybersickness from unaffected participants was not possible based upon the recorded motion alone. Additionally, attempts to capture stability data during actual VR gameplay in specific

areas of possible head stability provided mixed results and failed to identify participants exhibiting symptoms of cybersickness successfully.

In conclusion, this study finds that while a proxy measurement for head stability is obtainable from an HTC Vive headset, the results recorded in no way indicate cybersickness onset. Additionally, the study proves cognitive load and test order significantly impact stability measurements recorded in this way. As such, this approach would need calibration on a case-by-case basis if used to detect cybersickness.

1 Introduction

Since 2015, virtual reality (VR) has seen a massive resurgence, transitioning from the gimmick-based systems of the ’90s to systems with the potential to contribute to many fields (Stein, 2016). Primarily used for entertainment, virtual reality devices are also beginning to appear in a broader range of setting, such as education, architecture, and staff training scenarios. As a result, the possibility of a person encountering VR technologies in their day-to-day life is increasing.

This type of immersive visualisation is not without its problems. Some users experience a severe form of motion sickness, called cybersickness, which induces symptoms ranging from headaches and dizziness to vomiting in extreme cases (LaViola, 2000; Rebenitsch and Owen, 2016; Stanney and Kennedy, 1997). This issue has been a factor in immersive display usage since the initial use of immersive displays in training simulators (Kennedy et al., 1992). However, as these technologies required expensive equipment and vast amounts of space, their usage was mostly restricted to niche scenarios where the expense could be justified.

In 2012, a consumer-grade VR HMD device was conceived in the Oculus Rift DK1. This device demonstrated vast improvements from the bulky systems seen in the 90s’ and, as such, demonstrated VR HMDs had the potential to be used in a wide range of entertainment and industrial applications. However, as the number of people using VR grew, so did the reports of cybersickness incidents. Therefore, serious work began on identifying the cause and a solution to cybersickness. Work has identified contributions from vection (Palmisano et al., 2015), field of view (Fernandes and Feiner, 2016), locomotion (Clifton and Palmisano, 2020) and vergence accommodation conflict (Kramida, 2016). Several studies propose detecting cybersickness onset using physiological measurements to identify the onset of sickness, including the concept of detecting postural instability during VR usage (Risi and Palmisano, 2019a).

1.1 Research Structure

The investigation conducted in this thesis focused on identifying and quantifying the relationships between the four factors presented in **Error! Reference source not**

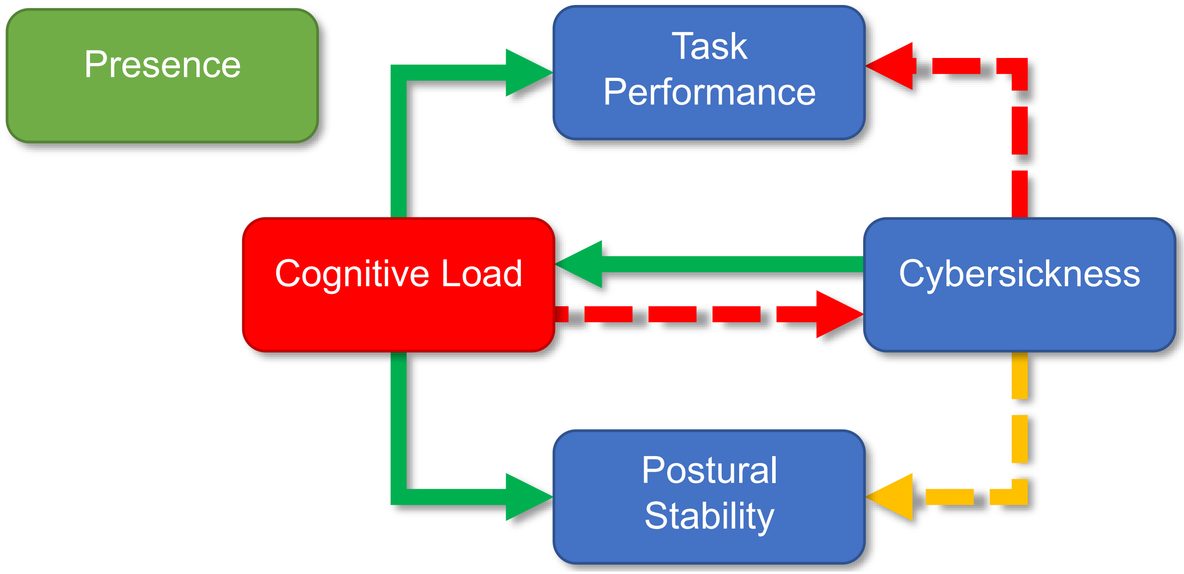


Figure 1-1:- Overview diagram for the structure of thesis identifying the major components and their relationships. In this diagram green arrows represent known unidirectional relationships, red dotted arrows indicate the relationships that are being analysed by this thesis and the yellow dotted arrow represents the attempted measurement pathway as a proxy for detecting cybersickness.

found. Cognitive Load, Postural Stability, Task Performance and Cybersickness, these relationships will then be used to answer the following four research questions.

1. Does an increase in cognitive load within a virtual reality task influence the severity of cybersickness symptoms?
2. Does an increase in the cognitive load experienced by a subject within a virtual reality environment impact postural stability measurement?
3. Can the alteration of task performance be used to identify the onset and severity of cybersickness in an individual?
4. Can the onset of cybersickness be practically identified mid-task using commercial off-the-shelf VR equipment?

1.1.1 Issues with measurement of cybersickness:

One problem with current research is the lack of standardised stimuli, making comparing results across different studies difficult. A standard method for capturing the severity of cybersickness symptoms exists in the Simulator Sickness Questionnaire (SSQ). However, the differences between each experience's different factors, such as vection make attributing any difference to a single factor difficult, even when the factors investigated are the same. As such, much disagreement exists as to whether postural instability is an indicator of cybersickness (Arcioni et al., 2019; Dennison and D’Zmura, 2017; Rebenitsch and Quinby, 2019; Risi and Palmisano, 2019a).

However, questions exist as to whether the accuracy of the data provided by the VR HMD is sufficient to make this determination. Niehorster et al. (2017) shows the HTC Vive HMD may have sufficient accuracy for this but identifies several ground plane orientation issues during data capture. This study theorises that perfect accuracy is not necessary for this approach to be successful; and is not a reasonable target within a real-world VR use scenario. This implies that there is a need to investigate the possibility of identifying postural instability using real-world/imperfect measurements.

1.1.2 Relationships within diagram

Individual stimuli used may contribute to cybersickness at a higher level than currently acknowledged and the current test environments used may hide some of this contribution. Significant differences exist in the cognitive load of different experiences used in research, riding a VR roller coaster and playing a VR game provide very different experiences. Most stimuli chosen for VR studies or the methods used to capture data are not good representations of actual VR usage. They either use exceptionally provocative stimuli such as a VR roller coaster ride (Davis et al., 2015) or significantly restrict the users' movement, for example, making them stand on a balance board to record postural stability (Dennison and D’Zmura, 2017). Neither model accurately represents normal VR usage either through extreme stimuli or unusual behaviour.

There is a known link between cybersickness and cognitive load (Nalivaiko et al., 2015; Nesbitt et al., 2017), however, the opposite is largely unquantified, therefore this restriction on cognitive load may impact cyber sickness onset, and as such, models for this kind of study should utilise a more realistic VR usage model. Studies such as Mittelstaedt et al. (2019) have shown weak evidence that cybersickness can impact perceived task performance, however, this did not consider the cognitive load of the task. This may therefore actually be a consequence of the link between increased cognitive load due to cyber sickness symptoms which has a known strong link to task performance.

The level of cognitive load has a known effect on postural stability (Andersson et al., 2002). If this holds true in in VR environments this may be able to be detected in positional data captured via an HMD. This leads to the question of does cognitive load affect cybersickness (a bidirectional link) and by proxy postural stability measurements. Therefore can cybersickness onset be measured through changes in postural stability.

The standard method of capturing postural stability data uses the quiet stand method, where the user adopts a known stance and looks straight ahead for a given period. This stance creates a stable point of head motion for collecting samples of the head position. This research theorises that this approach may not be necessary, as certain VR usage aspects create moments of predictable movement and therefore postural stability within the experience, such as walking down a hallway. While this represents a much smaller opportunity to collect data and additional noise will be found within the results, this study believes this will not impact the collection of reliable measurement and, as such, represents an opportunity for passive data collection.

From this we can analyse our four research questions:

1.1.3 Question 1) Does an increase in cognitive load within a virtual reality task influence the severity of cybersickness symptoms?

Currently, a significant problem with cybersickness research relates to using a myriad of different stimuli to induce the condition. These stimuli already have many differences, which impact the onset of cybersickness (outlined in chapter 2). This results in a significant amount of conflict in the reported results of studies. Most studies do not consider the task's cognitive load requirements on the point of onset and severity of cybersickness. Typical experiences used in this research field range from completely passive to ones with active user involvement. Passive experiences such as virtual rollercoasters, or 3D videos are devoid of user interaction and have no complex environmental factors to consider. In contrast, active experiences often require complex problem-solving and object manipulation tasks that induce an additional multidimensional cognitive load onto the subject.

This study theorises that additional cognitive load within the stimuli reduces the mental resources available to process the virtual environment. This additional load could significantly influence the subject's ability to process the virtual environment, altering disorientation and cybersickness symptoms. Answering this question would help address the cause of conflicting results within comparable studies in the field currently purely attributed to differences in the physical properties of the studies involved.

1.1.4 Question 2) Does an increase in the cognitive load experienced by a subject within a virtual reality environment impact postural stability measurement?

Postural instability theory attributes the cause of cybersickness to the mismatch between the stimulation induced by the virtual environment and the real-world environment. The resulting conflict of these stimuli causes cybersickness symptoms. Postural stability theory also states that while suffering from cybersickness symptoms, users also exhibit increased involuntary postural motion (referred to as postural sway). Currently, there is significant conflicting evidence as to whether this is true.

This study theorises that differences in cognitive load may influence measured postural stability, directly through increased cognitive load caused by the task; or indirectly through increased memory requirements for task processing taking memory resources away from environmental processing. This study theorises that this will be evidenced by significant differences in postural sway and subjective cybersickness measurements. By quantifying the impact of cognitive load on postural sway measurements, inconsistencies reported in other studies may be explainable.

1.1.5 Question 3) Can the alteration of task performance be used to identify the onset and severity of cybersickness in an individual?

Being sick and disorientated has a known impact on a user's ability to perform tasks. Theoretically, if this performance degradation is measurable and monitored in real-time, it should be possible to identify the onset of cybersickness in a subject as their performance varies from a baseline within an activity.

This study theorises a predictable variation of user task performance exists as cybersickness symptoms worsen. If this is true, it may be possible to effectively identify the onset of sickness symptoms by monitoring performance and alerting the user when their performance deviates, either below a certain threshold or as a deviation from the baseline performance.

1.1.6 Question 4) Can the onset of cybersickness be practically identified mid-task using commercial off-the-shelf VR equipment?

Cybersickness is a significant barrier to widespread VR adoption in industrial fields. Some studies have taken the existing theories of postural stability and have attempted to use the subconscious motion of subjects to identify cybersickness onset. While theoretically, these approaches have demonstrated some promise, they all seem to lack consideration for the practical implementation of the solution into a working VR environment using an unmodified HMD and no additional tracking system. Solutions relying on modified headsets or additional tracking solutions are impractical to deploy due to expense and availability of equipment and, as such, are unlikely to see a practical implementation. Most studies assessing this use the quiet stand technique, requiring the user to remove themselves from their task to adopt a specific pose to measure the instability. This breaking from the task is undesirable, as it breaks the user's concentration, and the inconvenience will likely result in this being difficult to implement into a normal workflow.

This study theorises that the onboard positional measurements of the headset may be sufficient to identify differences in the subconscious postural motion of the subject. While the measurements are not as accurate as dedicated external tracking systems, they may be “good enough” to detect these differences and allow the system to act upon them in the form of warnings to the user or make adjustments to the simulation to counteract the onset of sickness. If so, this would represent a significant step towards producing a practical method of identifying cybersickness without additional cost or resources.

This study also theorises that the quiet stand process may not be necessary for identifying instability. It may be possible to take these measurements by identifying periods of low or consistent activity during the experience and taking measurements during these periods. Successfully proving that this is possible would enable passive monitoring of users throughout the VR experience resolving the problem of intrusive measurement procedures and increasing the likelihood of adoption within a real-world usage scenario.

1.2 Novel Contributions:

As such, this study aims to identify any contribution to both cybersickness and VR HMD instability caused by an increase in cognitive load within the VR HMD environment. Additionally, this study aims to appraise the method of capturing VR HMD positional data using the quiet stand method and identifying potential stability areas during VR HMD usage. The value of this investigation comes in 4 parts.

1. Identifying the contribution of different cognitive load types on cybersickness onset will help compare results between different cybersickness studies, especially when severe differences exist between the stimuli used.
2. Identifying if the quality of headset positional data captured during usage of an HTC Vive VR HMD is sufficient to detect instability may identify a cheap, cost-effective way of measuring subjects' stability during VR HMD experiments.
3. Identifying the impact of cognitive load on stability measurements made via the positional measurements of VR HMD devices will help identify the robustness of this approach and if factors within the environment itself can affect instability, reducing the reliability of using this method to detect cybersickness.
4. Identifying periods of potential stability within a VR HMD experience may allow for the development of an entirely passive approach to identifying cybersickness onset without the need for a frequent quiet stand period.

By answering these four questions, the study aims to contribute by identifying a way to make cybersickness detection unobtrusive in a manner requiring no additional equipment beyond the standard VR setup that works for a realistic VR usage scenario.

This research's novelty comes from investigating the impact of cognitive load on cybersickness and is one of the most promising measurement techniques currently being developed. If the amount of cognitive load impacts cybersickness severity, it will prove the need to rethink task loads contribution when comparing multiple studies' results. Additionally, collecting instability data via the HMD positional data will prove this approach's practicality revealing any specific issues related to this approach in a realistic VR usage model. Finally, the impact of cognitive load on stability measurements made via the VR HMD device will identify the impact of a critically overlooked component of this technique, providing additional information towards developing a practical approach to applying this technique to mainstream VR usage.

1.3 Thesis Structure

Chapter two of this thesis explores VR technologies and cybersickness to identify the groundwork for this research. The investigation looks at cybersickness, its causes and how the severity of symptoms are captured and recorded. The chapter also looks at modern VR HMD devices, typical usage scenarios, how they work, and what VR HMD usage factors impact cybersickness. Finally, a review is undertaken into the cognitive load to identify what can cause a task to have a high load, the different cognitive load types, and how to assess a task's load qualitatively.

Chapter three outlines the design of the virtual environments used in testing. Describing how the environment allows the identified factors to be studied and their importance to the research. The design decisions made during the development of the environment are discussed to ensure the suitability of these decisions and the environment as a whole.

Chapter four outlines the experimental method. This outlines the process for conducting the experiments and the data collection methods used during testing. Hypothesis are also identified allowing for an answer to be formed relating to the research questions. Statistical models for proving these hypothesis are also identified.

Chapter five presents the study's results across 4 areas SSQ, TLX, HMD position and Balance Board data. This data is then used to validate the hypothesis proposed in chapter 4 and make a conclusion as to their validitiy. Chapter 6 sees a discussion of the results and the consequences of the study's findings and detrimines if each researchquestion can be validated; Chapter 7 presents the study's conclusions and advice regarding this approach's effectiveness.

2 Literature review

2.1 Virtual Reality

Since 2015 VR (Virtual Reality) technologies have seen significant innovation and investment from several sources such as the video games, architecture, training and education sectors. These projects aim to deliver on the promise of infinite virtual worlds to consumers. The various systems and alternative platforms such as AR (Augmented Reality) and MR (Mixed Reality) can make identifying what constitutes a VR system complicated. Therefore, a suitable definition defining exactly *“what is VR?”* is required. The Virtual Reality Society defines VR as…

> *“Virtual reality is the term used to describe a three-dimensional, computer-generated environment which can be explored and interacted with by a person. That person becomes part of this virtual world or is immersed within this environment and whilst there, is able to manipulate objects or perform a series of actions.”*
>
> *(Virtual Reality Society, 2017).*

VR systems enhance the link between display and interaction; one of the earliest essays on the subject, The Ultimate Display by Sutherland (1965), talks about how visualisations may constrain user interaction. With chairs accurate enough to sit in and handcuffs physically constraining the user, but also talks about visualisations identifying the purpose of the display to be.

> *“A looking-glass into the mathematical wonderland constructed in computer memory”*
>
> *(Sutherland, 1965).*

This quote acknowledges the ideal focus of VR. The visualisation should be “the looking glass” to another world, creating the belief the user is no longer in the real world but is instead replaced by the artificial. It seems more appropriate to define VR, not in terms of interactivity, which may not be necessary to cross through the looking glass; But more in terms of presence, the level of immersion or sense the user is within the new environment. Analysing the stereoscope, a Victorian would describe the sensation of viewing a 3D photograph as like being in the place they are viewing. As technology moves on, our requirement for immersion increases. Thus, while the stereoscope is at best considered a primitive VR experience by modern standards, due to the lack of immersion provided, the definition of what a VR system currently has not changed.

Can the experience successfully transport the user from their existing environment to another virtual one? If so, this is fundamentally a virtual reality experience replacing the user's belief in the real world with a virtual one. The limit of control a user has on the environment should have no impact on the definition of whether an experience is a VR experience or not. Therefore, the definition of VR should be revised to include systems where the user is a passive observer.

"Virtual reality is a system designed to replace a user's existing perception of reality, convincing them that artificial reality exists and they have a presence within it."

At this point, it is worth briefly considering where AR systems fall into this definition. A VR system aims to replace the world, whereas, in contrast, an AR system aims to enhance it, overlaying graphics into the users existing worldview highlighting areas of interest or inserting graphics to enhance the existing environment. AR systems do not create an alternative environment; they instead must enhance the existing one. Therefore, AR systems are not considered VR systems as they do not replace the users existing perception of reality and serve a different function.

2.1.1 A Brief History of VR

The first widely acknowledged VR system was the Sensorama (Heilig, 1962) (Figure 2-1). The Sensorama was a cabinet approximately the size of a vending machine designed to stimulate all senses (except for taste). It utilised a stereoscopic display and vibrating seats, speakers, fans, and a Smell generator. Heilig, its inventor, followed this in 1960 with the first virtual reality head-mounted display. The Telesphere Mask (Figure 2-1) utilised two miniaturised CRT displays to produce full-colour stereo 3D images with stereo sound.

Content removed due to copyright restrictions

Figure 2-1:- The Sensorama Cabinet (left) and Telesphere Mask (right).

1961 saw another significant leap in the concept of head tracking with the Headsight (Sutori, 2020), which facilitated head tracking via the usage of a magnetic motion tracking system; this is the genesis of what modern VR systems look like today. The Sword of Damocles (

Content removed due to copyright restrictions

Figure 2-2) followed in 1968, invented by Ivan Sutherland. The device provided stereoscopic wireframe images that responded to changes in the user's viewing point.

Content removed due to copyright restrictions

Figure 2-2 shows the framework required to facilitate the head-tracking via mechanical means and support the device's substantial weight. While the display was partially translucent, the device is not considered an AR device due to the lack of consideration or intent to place the graphics into the real world.

In 1987 Jaron Lanier introduced the term virtual reality to the field, which did not have a term to describe the collection of devices developed throughout the 70s and 80s. Jaron's definition gave uniformity and identity to the fledgeling field; from here in the early 90's the first wave of consumer virtual reality products entered the market, effectively triggering the first wave of virtual reality interest from the public, primarily in the games industry. Systems were bulky, expensive and the visuals provided were generally of inferior quality. Big-name companies such as Sega, with the Sega VR (Wiltz, 2019) and Nintendo with the Virtual Boy (Edwards and Edwards, 2015) invested heavily into virtual reality products, which never delivered the promises of next-generation gaming to consumers.

It is at this point that the problems of cybersickness rose into the public eye. While devices and simulators developed up to this point (primarily military simulations) had shown a tendency to induce motion sickness-like symptoms in users, these systems were never widely used; thus, in most cases, cybersickness went unreported. The ’90s saw the first commercial wave of VR technologies released to consumers. Testers and early adopters of these VR systems also reported cybersickness symptoms after relatively short usage periods. A much larger user base existed for these devices, which led to comprehensive, widespread reporting by specialist magazines of the day (Stein, 2016). While initially, the press for these devices was positive, the expense of purchasing these underperforming devices ultimately killed the publics’ interest in these systems (Edwards and Edwards, 2015; Fowle, 2015). This lead to a flat period in VR advancement due to lack of commercial investment (Jenkins, 2019), Specialist simulators continued to improve. However, no significant innovations emerged until 2010, when the Oculus project began.

Content removed due to copyright restrictions

Figure 2-2:- The Sword of Damocles VR System.

The Oculus project inspired a renewed interest in VR from an industrial perspective and the general public (Heaven, 2013). The DK1 prototype released in 2012 was a significant step up in quality from the headsets of the 90s, taking significant advantage of advances in display technologies to produce a lighter, more comfortable headset. While the DK1 headsets had significant issues, Such as the screen door effect (see section 2.1.2.2), the DK1 was sufficient to show the technology had matured to the point that VR could be a practical reality with significant potential in many fields. The Oculus' success drove many competitors to enter the market, most notably HTC, with the HTC Vive (HTC Corporation, 2018) backed by Valve, the company behind the most prominent pc digital marketplace (Statt, 2019).

The 2nd generation of modern commercial VR HMD devices entered the market with increased visual fidelity and improved ergonomics. These devices are the first generation of devices capable of delivering the promise of VR without the significant limitations of older systems and have seen a wave of consumer adoption. With this, the media has reported a significant increase in cybersickness incidents in adopters of the technology (A. B. C. News, 2019; Caddy, 2016).

2.1.1.1 Industrial Applications

VR is not limited to entertainment purposes and has applications in many additional fields. Today's VR systems are lightweight, portable and, most importantly, cheap enough for consideration in the workplace. Typical industrial VR applications utilise the same interaction methods as games, just generally at a slower pace and less complexity, often sacrificing simulation for ease of use. Typical example tasks include visualisation of building structures for architecture, video conferencing and training exercises.

2.1.1.2 Medical Applications:

The medical field represents a significant portion of the ongoing research in the usage of VR devices. Many potential applications in medicine, both in assisting health care professionals in patient education (Jimenez et al., 2018) and training (Riener and Harders, 2012), but also in a wide variety of physical (August et al., 2005; Carrougher et al., 2009) and phycological (Freeman et al., 2017; Riva, 2005) therapies.

2.1.1.3 Architectural Applications:

One field that has shown great success is the field of architecture. VR visualisations of projects for customers are a standard service and inform stakeholders about the final output from every perspective. VR allows a real-time walkthrough of the project, allowing for a better representation in scale than plan drawings and traditional 2D displays. Another prominent feature is seeing the utilities' position in scale and their position throughout the building. Traditional visualisation offers a limited perspective of these components, where small misalignments are not always apparent in traditional perspective views. A VR based walkthrough allows intersections with walls or misalignments to be observed in situ before construction even begins, allowing expensive mistakes with the plans to be corrected without reworking on site.

2.1.1.4 Training Scenarios:

Another big area for potential VR adoption is workplace training. Computerised training programs have been a significant part of workplace training for many years, providing an interactive platform to communicate knowledge to employees and validate the knowledge transfer quickly. These systems' interactivity is generally quite limited, mainly being multiple-choice questionnaires or identifying the problem within the scene activities. Virtual reality represents a way to improve this process by allowing the user to be placed in the working environment to perform the task in a safe environment without risk. As the VR uses motion controls to perform the task, training can go from pure theory to confirming the practical skills needed have been trained correctly. For example, the standard manual handling training package can be upgraded from a “question and answer” approach to a practical training approach. A VR approach would allow monitoring posture and avoidance of environmental hazards, informing users of errors being made and ensuring the correct procedure is adopted.

2.1.1.5 Teleconferencing and Remote Working:

VR could be used to replace traditional teleconferences in the future. Having a virtual meetup space with physical avatars could help business, allowing remote conferencing with physical avatars giving the advantages associated with in-person meetings. Traditional teleconference models lose essential information, such as body language, which significantly builds trust and confidence in an individual (Handford, 2010; Peleckis et al., 2016). With the COVID crisis, many jobs have moved to remote working, turning houses into offices and represents a significant shift in the working patterns of many industries, with platforms such as Microsoft

Teams (https://teams.microsoft.com/) and Zoom (https://zoom.us/) seeing widespread usage and taking over the role of interpersonal conversations.

Surveys of people remote working for the first time have shown a significant workforces desire to continue this practice in some form after the crisis (Ansorge, 2020), and this idea seems to be supported by major companies (Harper, 2020). Therefore, the remote working platform is likely to evolve rapidly as more significant portions of the workforce transfer most of their work activities to remote locations outside of the office. However, this changeover will not be immediate, as the immediate concern is getting a bare-bones solution to accommodate the organisations' immediate needs to continue to function. This change will serve as the acid test as to whether remote working is a feasible long-term solution for these organisations, potentially seen as a perk of the role. If so, improvements to these systems are likely, with VR being a probable avenue for immersive conferencing solutions.

2.1.1.6 Education:

VR experiences allow a virtual field trip (Harris, 2019), visiting sites of historical importance or visualising knowledge in an immersive way. As VR headsets' price comes down, VR usage in the classroom will become more commonplace, allowing teachers to immerse students in environments for intense, engaging learning experiences. The concept of virtual field trips may become commonplace in the future, sending the class to the titanic in VR for the day in a similar fashion to how a class would go to a museum.

2.1.1.7 Data Visualisation

Depth in terrain data is vital for many factors of geographic information systems data; stereoscopic outputs, such as those found in VR HMD systems, would facilitate depth in existing visualisation techniques. The benefits of 3D stereo object recognition in these systems could significantly enhance feature detection and analysis. However, in these fields, visualisation technologies generally lag behind mainstream visualisation. Comparing the workflows of geographic information systems and other more advanced fields such as architecture highlights a significant usability gap between the two fields. This gap exists because of the field's specialist nature, meaning that the user base is not as large and the people who work in it are incredibly specialised. Therefore, the field has fewer investment opportunities, only really evolving out of necessity when the benefit significantly outweighs the effort of changing the established workflows. VR has yet to prove useful enough in this field to see widespread adaptation; however, it has proven successful when implemented to solve a specific problem on a small scale (Haklay, 2001).

2.1.1.8 Health and Safety

These new systems pose new challenges in the workplace. Health and safety regulations are complex and highly detailed (Health and Safety Executive, United Kingdom, 1974), designed to keep workers on the job safe. Introducing Virtual Reality into the workplace is challenging, as the most common VR setups all involve effectively blindfolding the user and having them gesticulate within the space with motion controls. This interaction method is understandably quite dangerous (Baur et al., 2021; CNN, 2017), and most industrial applications are either performed in a dedicated space with limited access to other people or limited motion desk setups.

2.1.2 HMD Technologies:

To understand how the lenses in a virtual reality headset work, first, a basic understanding of how human vision works is required (Remington, 2012). Figure 2-3 shows a cross-section of the human eye. Light from the environment is focused first by the cornea, then by the lens onto the retina. The eye's retina is one of a series of layers at the rear of the eye, containing a series of photoreceptor cells, which send signals to the brain via the optic nerve when stimulated by light. There are two types of photoreceptive cells in the retina referred to as rod and cone cells. Rod cells serve to detect luminance and detect motion in objects. Cone cells are stimulated by specific frequencies of light and come in three varieties, sensing the intensities of

Content removed due to copyright restrictions

Figure 2-3:- A cross-section of the human eye (https://www.allaboutvision.com/resources/anatomy.htm)

red, blue, and green light, respectively; these cells allow humans to perform detail discrimination tasks.

The lenses in a virtual reality headset work similarly to how a traditional pair of glasses do. Light entering the lens is refracted in a carefully engineered manner to ensure the light emitted from the display enters the correct part of the user's eye, bending the light towards the correct part of the retina. Most VR headsets utilise a specific kind of lens known as a Fresnel lens (Davis and Kühnlenz, 2007), which uses a series of concentric rings of different angles engraved into a plastic lens to redirect the light into the eye. Fresnel lenses are much smaller and lighter than traditional lenses. Seeing how weight is a critical factor for VR comfort, this weight reduction is a desirable feature for VR headsets. As fresnel lenses have a flatter profile than traditional lenses, the displays can be mounted closer to the users' eyes, effectively making the headset's weight more comfortable to carry on the user's head. Figure 2-4 shows how as the rings step away from the centre of the lens, the ring's slope angle can be changed to focus the light from the display onto the correct area of the retina.

However, using fresnel lenses does have drawbacks. The ridges created in the lens create magnification bands in the image, possibly leading to blurring or noticeable

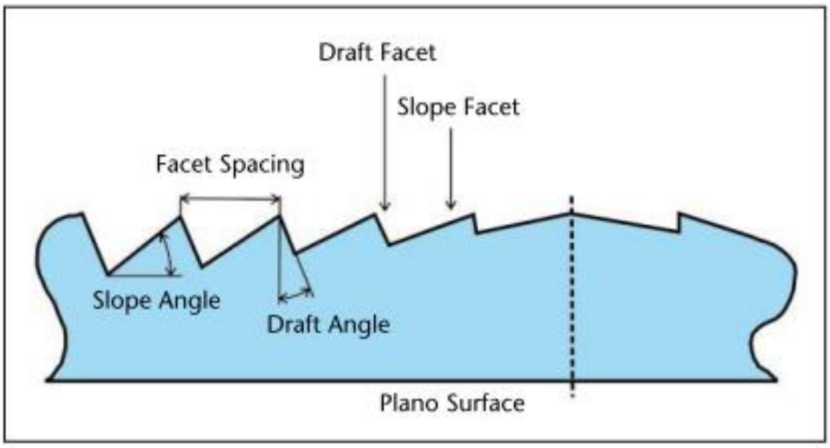


Figure 2-4: - Profile View of a Fresnel Lens (Davis & Kühnlenz, 2007)

artefacts where the magnification bands meet. Additionally, the refraction of light is not uniform across the colour spectrum resulting in chromatic aberration, where colours appear to bleed from the edges of objects. The solution to this would be to use smaller facet spacing but achieving this would introduce additional complexities into the manufacturing process. However, achieving this would also increase the number of draft facets and increase the amount of scattering from off-axis light hitting the lens, introducing artefacts such as god rays. As such, the optimal fresnel lens for VR must strike a balance between step number and space given to draft facets.

2.1.2.1 Field of View

Field of View (FOV) in VR headsets refers to the amount of the environment a user can view through the headset. For reference, humans have an approximate natural FOV of 210º Horizontal and 180º vertical (Mazuryk and Gervautz, 1999). Determining the FOV of a particular VR headset and comparing it to another can be difficult as FOV values for VR headsets are presented either as a single diagonal measurement in degrees or as the horizontal and vertical measurements as separate values. Different ratios in the horizontal and vertical lengths for headsets with identical diagonal FOV measurements can have very different FOV parameters in practice, comparing different headsets difficult.

The other consideration to be made with VR HMD devices is the different physical characteristics of the user's eyes. With traditional displays, the slight natural variation in human eyes' position from face to face has a negligible impact on viewing the display. With VR HMD devices, the eye position is critical and carefully considered during the headset design. The eye position's depth relative to the lenses needs to set correctly; otherwise, the lens will over or under-refract the light resulting in an unfocused image. Likewise, the lens's central point needs to be aligned correctly to focus the light through the iris in a comfortable manner. Otherwise, the eye muscles will need to hold an unusual position causing eye strain. Vertically adjusting the lens position can be achieved by moving the headset up or down on the user's face. However, setting the horizontal lens position requires adjusting the lenses closer together or further apart based upon the user's Inter-Pupillary Distance (IPD). Horizontal adjustment of the lens can be achieved either by mechanical means (physically moving the lenses further apart) or by digital means (adjusting the software display matrix).

FOV is vital for HMD Devices; studies have shown that increases in FOV generally lead to increases in Cybersickness symptoms (Lin et al., 2002), particularly during movement within the environment. Devices with smaller FOV's than that of the human visual system can lead to a system that feels restricted or boxed in, resulting in a reduction in environment presence (Seay et al., 2001). However, a blurring effect in the peripheral vision may reduce cybersickness (Lin et al., 2020).

2.1.2.2 Resolution and DPI

Early modern VR headsets (specifically the oculus DK1 and its prototypes) had a significant problem with the viewing experience known as the screen door effect, so-called as the viewing experience was precisely the same as if viewed through an old screen door. This effect occurs as the material used to separate the pixels from each other is large enough to be visible when viewed through the HMD's lenses. Figure 2-5 shows a closeup of an LCD panel. The black areas represent the framework (known as the substrate), holding each of the separate-coloured subpixels together. The bigger these areas are, the more noticeable the screen door effect will be. The screen door effect significantly affects a headset's perceived image quality (Dakers, 2020). A noticeable substrate interferes with the image presented and makes it harder to believe in the environment’s realism. Current generation headsets have minimised this issue by increasing the Dots Per Inch (DPI) of the displays, effectively fitting more pixels into the same viewing area, reducing the amount of visible substrate material between each pixel.

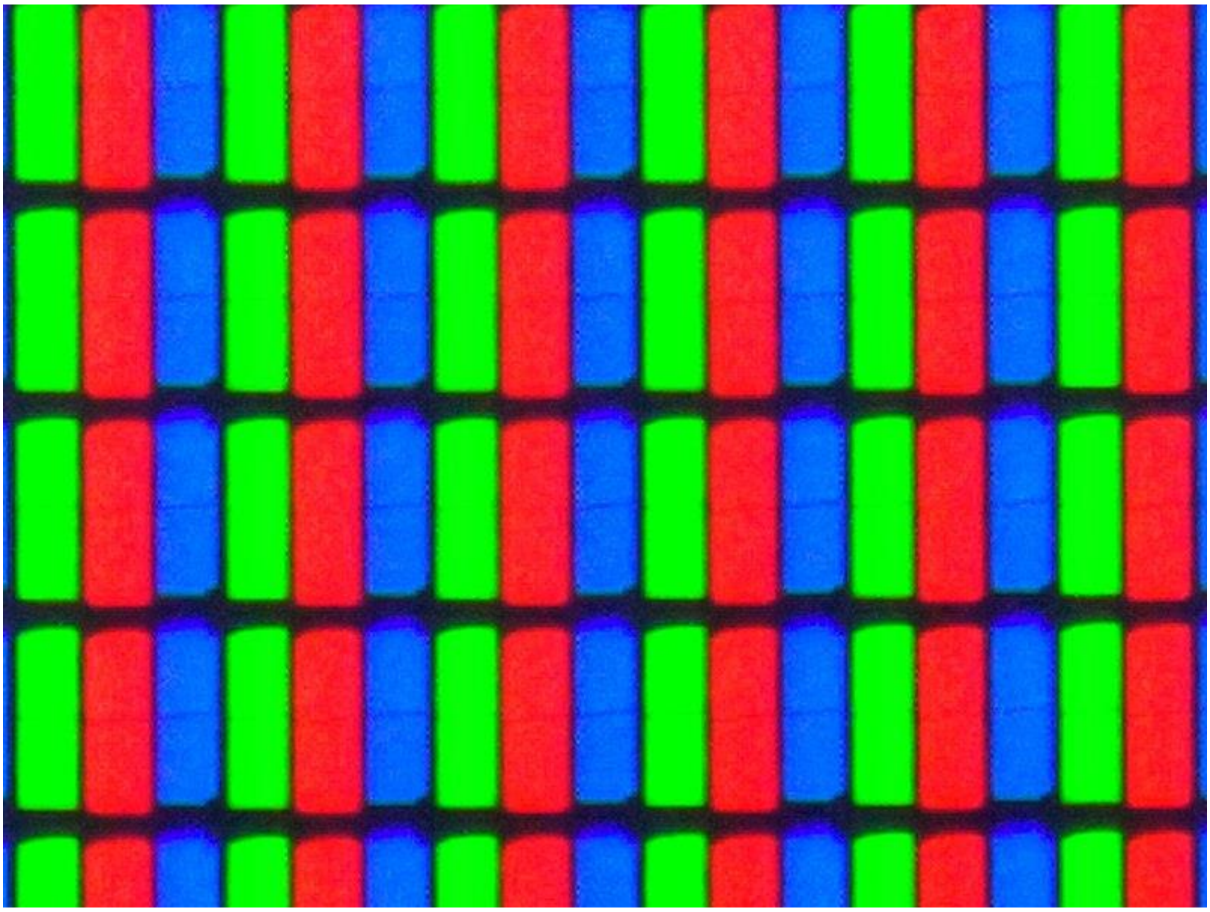


Figure 2-5:- Closeup of an LED panel (https://commons.wikimedia.org/wiki/File:Very_close_up_view_of_a_full_HD_LED_TV_screen.jpg)

2.1.2.3 Refresh Rate

The refresh rate of a device determines how quickly a display device can replace the displayed image with another, generally defined as a response time in milliseconds. The term refresh rate should not be confused with framerate, which is the rate at which the combined software and hardware can produce new images for display. To achieve smooth motion in digital media, the refresh rate and framerate must surpass the critical flicker fusion rate. The critical fusion rate is determined to be when human perception cannot distinguish modulated light from a stable field (Davis et al., 2015), Barten, (1999) identifies 72hz to be an acceptable rate to achieve this. The Occulus store enforces a standard framerate for applications on their store (Facebook Technologies, 2021a) between 72hz and 90hz, depending upon the target headset.

2.1.2.4 Vergence Accommodation Conflict

The accommodation reflex is the process by which the lens of the eye changes shape to adjust to the object's proximity. Figure 2-6 (a) shows the need for the lens to increase its curvature when examining close-range objects. The light reflected from closer objects is at a steeper angle than those further away; thus, the optimal viewing alignment requires more deflection to bring them into sharp focus. Convergence defines the process where the eyes move independently to focus on objects adjusting to ensure each eye is pointed directly at the target.

The optic systems in the VR headset place the display at a fixed distance from the user; however, virtual objects can have different depth cues closer or further away. Figure 2-6 (b) illustrates how the eyes attempt to accommodate the display's physical range while simultaneously attempting to converge on the displayed object's virtual depth. This mismatch in accommodation and vergence is impossible in the natural world and results in a mismatch between the accommodation and convergence distances, creating the vergence accommodation conflict (VAC) problem (Hoffman et al., 2008; Kramida, 2016; Reichelt et al., 2010). The result of this, particularly during scenarios of protracted usage, are headaches and eye strain.

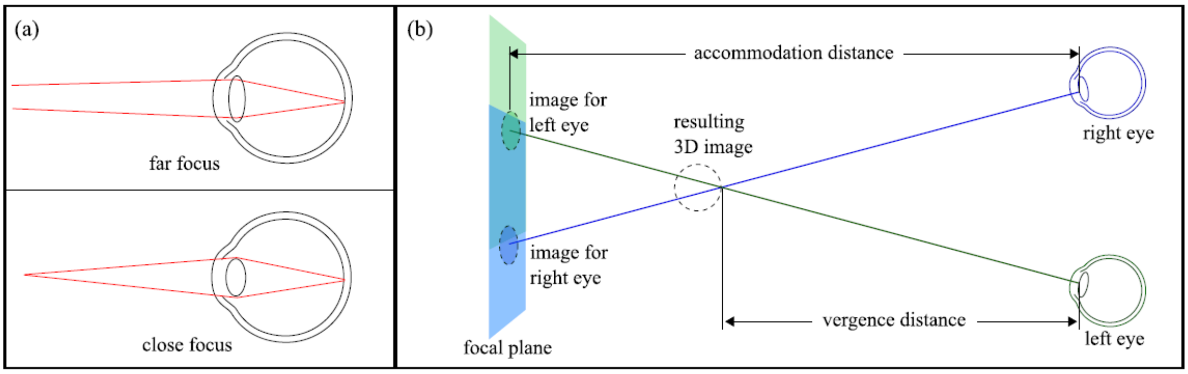


Figure 2-6:- Illustration of the Vergence Accommodation Conflict (VAC). (a) visual depiction of the accommodation reflex. (b) A visual depiction of the vergence conflict problem. (Kramida, 2016).

2.1.2.5 Rendering

Rendering to the displays of modern VR headsets is taxing on the hardware involved. Producing stereo visuals requires rendering the scene from two slightly different perspectives, once for the left eye and once for the right; this process requires the entire rendering process to be performed twice, effectively doubling the graphics hardware's workload. Effectively this doubles the rendering cost of high-quality full-screen effects such as anti-aliasing and bloom as all processes have to be applied twice. The increased cost of these techniques makes using them in VR environments an exercise in balancing the trade-off between quality and performance. As a side effect, the visual quality of VR games is often significantly lower than their non-VR counterparts. However, this limitation may be helpful Pouke et al. (2018) shows a weak correlation between scene realism and cybersickness severity, suggesting that a push towards higher fidelity graphics may increase sickness incidents.

One solution would be to determine the portion of the display currently being viewed and concentrate on rendering efforts there. The human visual system uses the centre of the retina to discern detail within the environment. The retina's outer edge (primarily comprised of rod cells rather than cones) only detects general luminance changes. Therefore, the majority of fine detail rendered into these areas is wasted and unobserved. Foveated rendering is a technique that utilises eye-tracking hardware within the headset to identify where the user is looking and optimise the rendering quality in that area of the display appropriately. Another variant of this technique called fixed foveated rendering forgoes the eye-tracking system, instead assuming the user has a fixed gaze point and optimising rendering around that. This technique cannot be applied as aggressively as non-fixed foveated rendering due to the need to accommodate slight gaze variations.

2.1.2.6 Tracking

Head tracking is a crucial part of most VR experiences, and early in its development, researchers identified its value to the immersion of VR experiences (Pao, 2020). Allowing the user control of their viewing perspective within the environment increased the sense of immersion within the environments enhancing the VR experience (LaValle et al., 2014). In modern systems, this tracking is also commonly extended to the controllers to simulate hands or tools for interacting with the environment. Two standard tracking methods exist for VR, referred to as the inside-out method and outside-in method. Inside out tracking uses sensors mounted on the headset to track its position, while outside-in tracking uses external sensors mounted around the play space to track the headset position. This tracking is achievable in various ways, such as physical markers recorded by cameras and accelerometer data (Heaney, 2019). However, no standard definition of the level of accuracy required to achieve this is available.

2.1.2.7 Headset Characteristics

As shown in Table 2-1, there is a wide range of headsets on the market, each with different physical characteristics; all of these impact the headset's performance, task suitability, and the likelihood of inducing cybersickness.

Headset	FOV	Refresh Rate	Pixels Per Degree	Resolution Per Eye	Price
PSVR	100	120Hz	9.6	960 x 1080	$348
Valve Index	130	144Hz	11.07	1440 x 1600	$1000
PiMAX 5K	170	90Hz	12.05	2560 x 1440	$699
HTC Vive PRO	110	90Hz	13.09	1440 x 1600	$1200
HTC Vive	110	90Hz	9.81	1080 x 1200	Discontinued
PiMAX 8K	170	80Hz	15.02	2560 x 1440	$899
Samsung Odyssey+	110	90Hz	13.09	1440 x 1600	$499
ASUS HC102	95	90Hz	15.15	1440 x 1440	$399
Oculus Rift S	110	80Hz	11.63	1280 x 1440	$399
Oculus Quest	110	72Hz	14.4	1440 x 600	$399
PiMAX 5K XR	200	85Hz	14.7	2560 x 1440	$899
Vive Cosmos Elite	110	90Hz	13.09	1440 x 1700	$549
Vive Cosmos	110	90Hz	13.07	1440 x 1700	$699
PiMAX 8k Plus	200	110Hz	19.02	3840 x 2160	$899
Star VR	210	90Hz	8.7	1830 x 1464	$3200

Table 2-1:- List of Current Generation Headsets (original chart by:- upset980ti https://www.reddit.com/r/virtualreality/comments/gl9p5a/made_a_chart_comparing_some_of_the_many_vr/)

2.1.2.8 Improving Headset Quality

The various characteristics outlined above do have a significant impact on cybersickness symptoms. However, there may be an upper ceiling to these effects.

Shafer et al. (2019) compared the cybersickness scores of users experiencing different VR experiences on two different Oculus headsets, the DK2 and the CV1. These headsets have similar display characteristics illustrated in Table 2-2. Shafer et al. (2019) monitored the cybersickness scores of subjects playing games of varying degrees of intensity (Minecraft, elite Dangerous and Lucky's Tale) on the two headsets. The study reported no significant difference between experiences viewed on the two devices, suggesting that when comparing modern headsets, these various factors have a much lower impact than comparing more severe changes in individual aspects would lead us to believe. This result suggests that modern HMD devices may have hit the ceiling of cybersickness induction; However, a few caveats exist with this conclusion. The study showed a significant number of participants (25 people, 15.6%) withdrew from the study due to extreme sickness and the results for those who completed showed a generally low SSQ score (CV1 (M = 8.61 SE = .91) and the DK2 (M = 7.79 SE = .89)). This fact suggests that these results represent the difference in cybersickness scores in mild sufferers, severely affected subjects may still prove sensitive to the differences between headsets.

Headset	Oculus DK2	Oculus CV1
Field of View	100º	110º
Display Technology	AMOLED (Samsung Note 3 Pentile)	Dual low persistence Samsung AMOLED (Diamond PenTile subpixel matrix)
Display Panel Size	5.7"	7"
Display Screen Resolution	1920 x 1080	2160 x 1200
Per Eye Resolution	960 x 1080	1080 x 1200
Total Pixels per eye	1,036,800	1,296,000
Display Refresh Rate	60Hz, 72Hz, 75Hz	90Hz
Head Tracking	6DOF Positional Tracking (based on the combination of 3-axis rotational plus 3-axis positional) USB connection using Near-Infrared CMOS sensor	6 DOF Constellation camera optical 360-degree IR LED tracking

Table 2-2:- Headset characteristics of 2 oculus headsets, the DK2 and CV1. (Shafer et al., 2019)

2.2 Cybersickness

Cybersickness is the term used to describe the condition where a user of an XR (Cross Reality) system begins to suffer undesirable symptoms akin to motion sickness. (Mazloumi Gavgani et al., 2018)

2.2.1 What is Motion Sickness

Traditional motion sickness has been well studied (Leung and Hon, 2019), but the fundamental root cause of motion sickness is still disputed (LaViola, 2000; Zhang et al., 2016). There are four main theories regarding the underlying causes of motion sickness: sensory conflict mismatch theory (Reason, 1978); postural stability (Riccio and Stoffregen, 1991); Poison theory (Treisman, 1977); and nystagmus theory (Ebenholtz et al., 1994). Of these, the former two are considered the most likely.

2.2.1.1 Sensory Mismatch Theory:

This theory attributes the cause to a mismatch in the visual and vestibular (the balance mechanisms in the ear) and other sensory systems providing conflicting information (Bos et al., 2008). Types of motion sickness in this theory are attributed to be caused by one of three scenarios.

- A motion that is felt but not seen. An example of this would be traditional seasickness, where the boat's motion does not match the surroundings' apparent lack of motion.
- A motion that is seen but not felt. It is commonly induced within a VR environment when the perspective within the environment shifts, creating visual motion, yet the subjects' vestibule systems register no mismatch.
- A motion that is both felt and seen but is conflicting. Commonly experienced during periods of zero gravity when the subject's vestibule responses do not match the expected motion represented by the visual motion experienced by the subject.

2.2.1.2 Postural Stability:

This theory suggests that the cause of motion sickness is a prolonged period of postural instability or inaccurate changes in a user's posture when responding to unusual external stimuli.

The postural stability approach (Riccio and Stoffregen, 1991) associates changes in a subject's stance and head movements with increases in the level of cybersickness a user is experiencing. Stoffregen et al. (2008) identified changes in posture of subjects who get motion sick playing traditional video games. The study demonstrated that subjects who suffered visually induced motion sickness playing games also exhibited more significant body position changes than individuals not exhibiting symptoms. Motion sickness in traditional displays is uncommon; however, sickness is a lot higher in VR gaming and applications (Sharples et al., 2008). Various experiments have shown that virtual environments induce postural instability alongside cybersickness (Murata, 2004; Rebenitsch and Quinby, 2019; Risi and Palmisano, 2019a). Postural stability is measured via pressure plates to measure the subject's centre of gravity over time, with variation in this measurement detecting increases in postural sway (Widdowson et al., 2019).

However, whether postural stability is indeed a guaranteed effect of cybersickness is not conclusive. Dennison and D'Zmura, (2017) show that cybersickness symptoms can occur with or without changes in postural stability and that increases in stimuli severity (in this case, vection) did not increase the amount of postural instability in the majority of subjects tested. This finding suggests that postural instability response is not guaranteed to accompany cybersickness. The effect may have a cap, meaning that it is an indicator a subject is suffering cybersickness but may not be useable to determine the severity.

Postural stability may also be useable in a predictive capacity. Arcioni et al. (2019) showed that users exhibiting cybersickness symptoms with increased severity of symptoms exhibit greater post-VR-exposure-test instability. However, the method of sickness induction used an inversely compensated model, which does not represent the expected scene motion and may not apply to general VR usage.

2.2.1.3 Poison Theory and Nystagmus Theory:

Poison theory suggests the unusual combination of sensory inputs triggers an evolutionary holdover as a response to being poisoned by some external source. (Treisman, 1977) Nystagmus theory suggests that stress on the ocular motor reflex may be another cause of motion sickness (Ebenholtz et al., 1994).

2.2.2 Categorising Motion Sickness Derivatives

Motion sickness symptoms are generally grouped into four broad categories: gastrointestinal, central, peripheral, and sopite, as described in Table 2-3. Sufferers of the condition describe the sensation in various ways, from general discomfort to full-blown physical reactions such as vomiting. Estimates from Rebenitsch and Owen, (2016) place the percentage of sufferers of the condition between 30% (Chen et al., 2011) and 80%+ (Kim et al., 2005).

Cybersickness is a form of (VIMS) Visually Induced Motion Sickness triggered by immersive displays (Rebenitsch and Owen, 2016) and has a proven impact on user engagement with content (Israel et al., 2019; Yildirim, 2019). These symptoms are generally unpleasant and, if severe enough, can lead to user rejection of immersive technologies.

Gastrointestinal	Generally described as an upset stomach, symptoms include nausea, stomach awareness, and vomiting.
Central	Internally derived disorders such as fainting, light-headedness, blurred vision, disorientation, dizziness, and the sensation of spinning.
Peripheral	Externally derived disorders such as sweating, feeling hot.
Sopite	Phycological-based symptoms such as annoyance, drowsiness, tiredness, and uneasiness.

Table 2-3:- Description of Symptoms of Motion Sickness.

Cybersickness is not unique as an identified subcategory of motion sickness as simulator sickness is a form of motion sickness specifically related to simulators (such as flight and driving simulators). The three conditions are linked, with similar symptoms; however, the intensity of symptoms is different for each Kennedy et al. (1992) and Stanney and Kennedy, (1997) illustrate the difference in intensity between the various symptoms. The symptoms of cybersickness can affect users in 3 ways and are categorised, based upon the systems they affect, as D (Disorientation), N (Nuerovegative), and O (Occularmotor). Symptoms may influence two categories at once. For example, the blurred vision symptom impacts both N and O systems. These responses are provoked in different intensities with different experiences.

XR environments provoke more disorientation-related symptoms when compared to traditional motion sickness. This profile for XR symptom intensity denoted as D > N > O, differing from traditional motion sickness expressed as N > D > O due to its higher contribution from the nausea component (Stanney and Kennedy, 1997). Mazloumi Gavgani et al. (2018) Argues and presents evidence that the two conditions, cybersickness and motion sickness, are clinically identical. This evidence means that while the methods of inducing the two conditions may differ, treatment of the symptoms will be identical. Thus, cybersickness should be considered a type of motion sickness.

2.2.3 Useful Sickness Measurements

Currently, a common way to identify if a user is suffering from sickness symptoms, is to ask them. This assessment can occur either during the process using the fast motion sickness scale (Keshavarz and Hecht, 2011) or post-test using methods like the Simulator Sickness Questionaire (SSQ) (Kennedy and Lane, 1993). In a practical sense, these methods are not very useful. The SSQ only provides information after the sickness has occurred. While this may be informative to future experiences, it does not help identify sickness during induction and cannot identify users experiencing the onset of sickness. The fast motion sickness scale tracks the level of sickness a user is experiencing during the experience but has the disadvantage of requiring the user to break concentration with the task in VR to answer questions about how sick they are frequently. While this process could easily be automated using text-to-speech software, the method would remind the user they are continually using a VR experience, reducing the presence experienced within the environment and increasing cybersickness symptoms (Israel et al., 2019).

Content removed due to copyright restrictions

Figure 2-7:- Examples of Electroencephalogram (Left) (https://www.nhs.uk/conditions/electroencephalogram/) and Electrocardiography (Right) (https://www.adinstruments.com/blog/correctly-place-electrodes-12-lead-ecg).

What is needed is a method of passively determining how sick a user is getting in the experience. Measuring physiological properties would be a suitable method of doing this, using sensors to monitor the user's physical characteristics. Some methods, such as EEG (Dennison et al., 2016; Jeong et al., 2019; Pane et al., 2018) (**Error! Reference source not found.**), have proven to have some ability to predict cybersickness and would be perfect for this task. However, these methods suffer from the need to have expensive equipment to record the measurements and need expert knowledge to use. This fact rules out using them in a practical scenario, as it is doubtful a majority of users would have access to the equipment and skills needed to make these measurements. A cheaper approximation of the measurements may come from other sources such as EEG headsets like the Emotiv Epocx (https://www.emotiv.com/epoc-x/). While evidence exists to state these devices may be useful (Taylor and Schmidt, 2012), they are very bulky and incompatible with VR HMD devices.

Heart rate may be another option Garcia-Agundez et al. (2019), Lin et al. (2018) and Nalivaiko et al. (2015) all show a correlation between increases in heart rate and severity of cybersickness symptoms using an ECG measurement. Once again, this causes a problem by the invasive nature of the data capture measurement, using a significantly complex piece of machinery with sensors stuck directly to the user's chest. However, unlike the EEG, measuring heart rates can be done using a less invasive method. Several devices exist in the marketplace for measuring pulse in the form of fitness monitors. However, accuracy is lower than that of the EEG (Nelson and Allen, 2018) and inconsistent from device to device. Whether this is suitable for usage in detecting cybersickness is unknown. This solution still requires additional hardware. Not every user will have access to a Fitbit or similar device for monitoring fitness. Ideally, our cybersickness detection should be entirely passive, as requiring no additional effort on the user's part to implement, adoption becomes automatic rather than an option, therefore increasing uptake.

2.2.4 Induction Mechanisms

The primary observed difference between cybersickness and traditional motion sickness is the method of induction. If we accept sensory mismatch theory as accurate, the cause of motion sickness-like symptoms is the disagreement on the reported state of the environment between different sensory systems (Bos et al., 2008), a sense of motion is induced in the subject by the visually represented environment. However, the subject feels no associated motion in the environment to accompany this resulting in a mismatch and induction of sickness symptoms (Bos et al., 2008). The visual stimulation may also be unnatural, such as severe rolling or flight simulation, further enhancing the mismatch and resulting in cybersickness.

One of the most prominent methods of establishing an individual's susceptibility to motion sickness is to utilise the Parabolic Flight Static Chair Test (Miller (II) and Graybiel, 1969). This test rose to public infamy when the popular tv show MythBusters ("Mythbusters Seasickness: Kill or Cure," 2005) utilised a form of the test to determine whether traditional and over-the-counter remedies for curing motion sickness work (seen in Figure 2-8). The Parabolic Flight Static Chair Test involves seating the subject in a chair attached to a rotating platform. The subject is then blindfolded and rotated in the chair at a constant speed. While rotating, the subject must touch their head to points suspended at four points around them (Forward, Backward, Left, and Right) at a regular pace in random order. This process creates a state of confusion in the bodies' sensory systems, resulting in motion sickness in susceptible subjects.

Such a test environment does not exist in VR because it is unnecessary for inducing cybersickness in modern environments; it is a simple process to place subjects within a VR environment and see if they get sick or not, and many studies do this (Gavgani et al., 2017; Stanney et al., 2020). When doing this, provocative stimuli, such as VR roller coaster rides, are chosen to maximise the possibility of inducing symptoms. This process is comparable to the Parabolic Static Flight Chair test as both are exposing the subject to deliberately provocative stimuli to induce symptoms for study.

Content removed due to copyright restrictions

Figure 2-8:- MythBusters Implementation of the Parabolic Static Flight Chair Test. https://sharetv.com/shows/mythbusters/episodes/300513

2.2.4.1 Validity of Induction Mechanics:

Cybersickness is not exclusive to VR experiences or computer visualisations (Naqvi et al., 2013; Solimini, 2013; Yang et al., 2012). However, the incidence and severity of symptoms are more pronounced in stereo visualisations, especially in HMD environments (Cobb et al., 1999; Yildirim, 2020). Initially, when 3D movies rose to prominence in the last decade, many reports arose of people feeling sick when watching the films. Most famously, Avatar (*Avatar*, 2009) generated a large amount of discussion (Brooks, 2009) over the film's apparent ability to make people nauseous.

2.2.4.2 Physiological / Biometric Approach

Collecting physiological symptoms has shown promise for identifying cybersickness in subjects (Dennison et al., 2016; Magaki and Vallance, 2019). The usage of electroencephalogram (EEG) (Lin et al., 2007; Pane et al., 2018), electrogastrogram (EGG) (Cheung and Vaitkus, 1998), and electrocardiogram (ECG) (Garcia-Agundez et al., 2019) have all shown the potential to identify CS symptoms in subjects. However, the methods required to collect, record and analyse the data require specialists to administer and are pretty invasive to the subjects and thus are not suitable for widespread usage. Sweating is known to be a symptom of cybersickness, (Gavgani et al., 2017) has shown that a measurable increase in skin conductance correlates with an increase in sickness. These factors all provide qualitative data regarding a subject's experience of cybersickness but the relationship between the biometric parameters and cybersickness severity is not understood well enough to make definitive decisions about the severity of user's symptoms.

2.2.4.3 Predictive Model Approach

Predictive modelling involves the usage of mathematical formula and machine learning to analyse data collected during VR usage to determine the confidence level of whether a user is experiencing symptoms or not. Studies such as Jin et al. (2018) collect data from the headset, such as head motion, and virtual environmental factors such as scene motion, and feed this information into machine learning models. These models' results are then compared to other qualitative methods such as the SSQ to establish the model's accuracy. This method is still in its infancy and thus requires additional time to develop.

Each of these methodologies has its pros and cons. Postural stability is the method with the most research currently available, although several studies dispute its validity. Nevertheless, as a method, it is non-invasive and passive throughout the experience requiring no engagement from the user to monitor. However, the usage of a force plate limits the user's ability to move around during VR usage and can be an expensive addition to a VR setup, as an example (Perform Better, 2022) sells an entry level force plate for £750 which is significantly more than most headsets (see Table 2-1) . Physiological and biometric approaches such as EEG are promising in a scientific setting and provide reliable data that can be correlated directly to cybersickness symptoms. However, these sensors are often invasive and require specialist knowledge to interpret the data correctly.

In some cases, they also require additional hardware, which can be expensive. Finally, predictive modelling shows promise as a passive monitoring system. However, to effectively train the model, large amounts of data must be gathered, depending on how specific the model is tailored to and may be unfeasible to obtain in the real world (Porcino et al., 2022).

2.2.5 Standard Measurement Techniques of Cybersickness

Silva and Fernando, (2018) categorised these methods into four broad approaches Questionnaire, postural instability, physiological/biometric and predictive.

2.2.5.1 Simulator Sickness Questionnaire (SSQ)

The established methods for quantifying cybersickness symptoms come from those used to quantify motion sickness. Early studies used the Pensacola Motion Sickness Questionnaire (MSQ) (Kellogg et al., 1965), a series of 25 – 30 questions that assign a numerical value, indicating the severity of each of the parameters associated with the symptoms of motion sickness. These values are then processed using a weighting system to give an overall score indicating the severity of motion sickness a subject is currently experiencing.

In 1993 Kennedy and Lane, (1993) identified the inefficiencies in the usage of the MSQ when studying simulator sickness, identifying the fact that the MSQ includes several factors which are not strictly relevant, such as drowsiness and alterations in appetite. Kennedy and Lane, (1993) also identified significant overlaps in some of the questions resulting in impact from the same symptom being included in the results twice. To address this, they introduced a new methodology termed the Simulator Sickness Questionnaire (SSQ); the SSQ Reduced the 25 questions from the MSQ down to 16. The SSQ measurement produces four scores, three of which refer directly to groups of symptoms, SSQ-O for oculomotor, SSQ-D for disorientation, and SSQ-N for nausea weightings (shown in Table 2-4). Finally, the SSQ-T score is a summation of all the effects considered for the SSQ. Studies quickly adopted the SSQ as the standard metric for collecting data relating to simulator sickness and has persisted to this day.

	Nausea	Ocular Motor	Disorientation
General Discomfort	9.54	7.58	
Fatigue		7.58	
Headache		7.58	
Eye Strain		7.58	
Difficulty Focusing		7.58	13.92
Salivation Increasing	9.54		
Sweating	9.54		
Nausea	9.54		13.92
Difficulty Concentrating	9.54	7.58	
Fullness of head			13.92
Blurred Vision		7.58	13.92
Dizziness with eyes open			13.92
Dizziness with eyes closed			13.92
Vertigo			13.92
Stomach Awareness	9.54		
Burping	9.54		

Table 2-4:- Symptoms and Weightings for the Simulator Sickness Questionnaire.

2.2.5.2 *VR Sickness Questionnaire:*

The study of Cybersickness in VR environments utilises the SSQ as the standard measurement for qualitatively assessing the severity of sickness symptoms experienced by a subject. Kim et al. (2018) presented a potential update to the SSQ called the Virtual Reality Sickness Questionnaire (VRSQ), citing the increased prevalence of VR Head Mounted Display (HMD) usage since the inception of the SSQ as justification for this change. The VRSQ eliminates the contribution to the overall score from the nausea component, citing observations in Drexler, (2006) showing that the nausea component of the SSQ contribution was less significant in

VR HMD studies than in VR Cave Automatic Virtual Environment (CAVE) or Binocular Omni-Oriented Monitor (BOOM); suggesting the VR environment generates smaller amounts of symptoms relating to the nausea component of the SSQ. so far, this method has yet to be adopted by the wider community.

Kim et al. (2018)'s proposal to remove the nausea component from the SSQ to create a more effective methodology does not seem to make sense. The justification used that nausea symptoms contribute less to SSQ scores in VR HMD environments is flawed. While evidence exists to support identifying nausea-related symptoms as having a lower impact on cybersickness symptoms (Mazloumi Gavgani et al., 2018), no study has shown that its contribution is non-existent or insignificant, justifying its removal from the scale. Also, the VRSQ implies that it is suitable for all forms of VR. Studies by Drexler, (2006) and Stanney and Salvendy, (1998) have demonstrated that various methods of displaying VR produce different severities of O, N and D symptoms. If the SSQ is not suitable as a generic case, then the derivative is unlikely to be either.

2.2.5.3 Cybersickness Questionnaire:

Another alternative is presented by Stone Iii, (2017) as the cybersickness questionnaire (CSQ). The CSQ is a refinement to the SSQ, aiming to remove the effect of symptoms not directly attributed to the onset of cybersickness, such as sweating, which is a common side effect caused by the headset's heat and physical exertion during VR usage. A better approach may be to quantify the headsets' impact on these effects rather than removing them from consideration altogether. By calculating the headset's contribution to these factors, identifying their additional contribution to these measurements from cybersickness could occur. As all the components contained within the SSQ are known symptoms of cybersickness, removing any of them because a secondary source may induce them does not make sense. Unless a factor can be proven not to be a symptom of cybersickness, measuring it should continue.

2.2.5.4 Discussion

Sevinc and Berkman, (2020) shows that both the VRSQ and CSQ can improve cybersickness identification over the traditional SSQ method. This report indicates that using a sickness scale tailored explicitly to modern VR HMD devices would likely be a better approach to the unique sickness profile of VR HMDs. The study does note that the sample sizes for these studies were too small to draw a definitive conclusion. Therefore, further validation of this method is needed to verify its suitability to replace the SSQ for VR HMD-focused cybersickness studies.

However, in many cases, no formal method is applied, with studies such as Paroz and Potter, (2018) using a 5-point scale instead of the 16 questions required by the SSQ. Many studies' small sample size (Farmani and Teather, 2018; Pane et al., 2018) also results in many ambiguous or inconclusive results (Faber and Fonseca, 2014). As an example of this, Paroz and Potter, (2018) produce inconclusive results as to the impact on cybersickness symptoms. This result is due to the relatively small sample size (6 people for each of the two samples) and a lack of depth in the questions used to replace the SSQ. The study uses a single five-point scale in place of the SSQ, which generates no insight into exactly how a subject may be getting sick and misses symptoms that are a factor in cybersickness but not considered relevant by subjects.

2.2.5.5 Motion Sickness Susceptibility Questionnaire (MSSQ)

Developed initially by Reason and Brand, (1975), the Motion Sickness Susceptibility Questionnaire (MSSQ) is commonly performed in tandem with the SSQ to establish baseline susceptibility to motion sickness. The MSSQ is often used to determine if a participant is susceptible to motion sickness and validate if a subject falls into the category of people under investigation. However, more commonly, the MSSQ is used to identify subjects who are likely to have an extremely adverse reaction to the stimulation and remove them from the study. Modern MSSQ studies can use the more compact version of the study, the MSSQ-Short (Golding, 1998).

2.2.5.6 Fast Motion Sickness Scale

The fast motion sickness scale was developed by Keshavarz and Hecht, (2011) to continually evaluate the severity of a subject’s motion sickness symptoms. This test contrasts with the MSSQ and SSQ methods that only collect data pre-and post-testing. The fast motion sickness scale asks the user to rate their motion sickness symptoms severity each minute verbally. This test allows a constant review of a subject’s condition during testing. While this information is not as granular or detailed as the SSQ, it does provide an in-the-moment measurement correlating sickness levels with various points of the experience. Modern automated options do exist for collecting this data; however, implementing them with modern VR headsets, where the subject is effectively blindfolded, can be complicated, especially if the user is already performing other tasks in the environment. Adding a visual scale to capture the information also breaks up the subject's workflow impacting presence. Therefore, the verbal method of communication remains prevalent. Other methods use a similar approach to this, such as the Nausea Rating Test (NRT) (Lo and So, 2001) with a different set of questions than the fast motion sickness scale. However, the fast motion sickness scale is by far the most prevalent. Table 2-5 shows an alternative to the fast motion sickness scale. The Misery Score (MISC) (Wertheim et al., 2001) is an alternative scale that yields a result between 0 and 10, each associated with a descriptive term describing the severity of the subject's symptoms. The MISC scale provides more detail than the fast motion sickness scale but still provides less detail than the SSQ.

Symptoms	Severity	MISC Score
No Problems		0
Some discomfort, but no specific symptoms		1
Dizziness, cold/warm, headache, stomach/throat awareness, sweating, blurred vision, yawning, burping, tiredness, salivation, but no nausea.	Vague	2
	Little	3
	Rather	4
	Severe	5
Nausea	Little	6
	Rather	7
	Severe	8
	Retching	9
Vomiting		10

Table 2-5:- Misery Scale, Derived from Bos, (2015)

2.2.5.7 Presence Questionnaire

The Presence Questionnaire (PQ) (Whelan, 2008; Witmer and Singer, 1998) measures the amount of "presence" within a virtual environment or the extent to which a subject believes they are within the environment. Version 3 of the presence questionnaire contains 32 questions evaluated on a 7point Likert scale (Witmer et al., 2005) (See Appendix A for full details). Implementation of this is inconsistent; if asked in the original order outlined in Witmer et al. (2005), some of the questions are out of order; for example, questions 5, 11 and 12 relate to sound and should, therefore, be grouped rather than spread out throughout the questioning. Additionally, asking questions in different orders in subjective tests can affect how a subject answers them, making a comparison of results more challenging.

2.3 Factors Influencing Cybersickness

2.3.1 Framerate and Refresh Rate

Latency in virtual environments has a well-documented effect on task effectiveness in computerised systems. Stauffert et al. (2018) shows that positional tracking latency in VR HMD systems increased cybersickness symptoms. Meehan et al.

(2003) has shown that latency also harms presence, which leads to a less enjoyable experience and, therefore, may significantly impact user acceptance of VR devices.

Framerate and refresh rate impact cybersickness in a similar manner and the two terms are often confused. Refresh rate describes the speed at which the display can replace one image with another. Framerate describes the rate at which the application can produce new frames for the display to show. As a minimum, the framerate and refresh rate should both be above 72hz to present flicker-free motion of the environment (Barten, 1999), but ideally, the refresh rate should aim to match the refresh rate of the HMD commonly 90hz (steantycip, 2020). Achieving this represents the maximum possible update rate the systems can physically present to the user. Exceeding this will provide a performance buffer to the system, allowing for momentary dips from the baseline performance, possibly caused by small pockets of graphical or processing intensity.

Failing to achieve this will result in poor performance, whereby the system fails to provide the next drawn frame to the display of the HMD before it is needed (steantycip, 2020). In this situation, the HMD must accommodate this. Suppose the system's performance is significantly lower than needed. In that case, meeting the minimum refresh rate will be impossible. Therefore, the system will inconsistently fail to achieve the desired response times resulting in many missing image frames. The user identifies these dropped frames as stuttering; when there is no replacement image available, the old one remains, giving the appearance of slight pauses in the environment's motion. The HMD can reduce its refresh rate to accommodate this. However, this induces latency (steantycip, 2020). Slower image refresh rates mean longer times between images reducing the perception of fluid movement, perceived as flickering and a reduction in accuracy in objects' positions.

Inadequate framerate leads to latency within the environment and increases the mismatch between the visual stimuli and the user's vestibule systems (Ng et al., 2020). If sensory mismatch theory is correct, this is one of the most significant contributing factors to cybersickness. Latency has a significant impact on sickness, as demonstrated by (Stauffert et al., 2018). Therefore, it can be stated that inconsistency in framerates and the variability in latency it creates have a significant impact on sickness. In reality, it is impossible to guarantee a consistent framerate.

Headset manufacturers are looking for solutions to resolve this issue; for example, Oculus has developed the Asynchronous Spacewarp technology (Oculus, 2016), aiming to detect movement in a scene and automatically generate missing frames within an underperforming system. However, this technology is not perfect as it does not have time to fully re-render the image. Otherwise, there would not have been a problem in the first place.

2.3.2 Vection

Vection is the sensation that the body is moving in space despite no motion taking place, the problem of vection is not exclusive to VR, Vection research has been ongoing since the 1900s (Wood, 1895), and most people have experienced some form of vection, an example given by Riecke, (2010)

> *"when sitting in a train waiting to depart from the station and looking out of the window where a train on the adjacent track starts moving, many people experience a rather convincing illusion that their own train started moving." (Riecke, 2010)*

In VR, the changing viewpoint or non-physical navigation through the environment often induces this, with the visual systems of the user detecting motion within the environment, causing the belief that the user is moving through the environment. Traditional video games also introduce vection to help pull the user into the game world. This presence must be balanced as too much vection triggers cybersickness symptoms (Stoffregen et al., 2008). Many senses can induce Vection, although it seems easier to induce via visual rather than auditory stimuli (Keshavarz and Hecht, 2012; Riecke, 2010).

The type of motion also has a significant impact, as demonstrated by Bonato et al. (2009), where rotating a virtual room on a single axis generated lower SSQ scores than rotating the room on multiple axes. This finding suggests that more complex motion increases the severity of cybersickness symptoms. This study is open to uncertainty, though, as the higher SSQ scores could be attributed solely to the roll rotation, which was never tested in isolation to identify its impact. Bonato et al. (2009) defends this by citing work in an earlier study (for which the results are unavailable) and the findings presented in Joseph and Griffin, (2008). The work of

Joseph and Griffin, (2008) does not support this as it induces sickness via the vestibule systems, whereas Bonato et al. (2009) induces the sickness by visual stimuli without presenting evidence that the stimuli are equivalent.

Diels and Howarth, (2011) tested the combination of simulated forward-backwards motion and roll motions induced via a field of white dots. The study reported a slightly lower SSQ-T score for the forward, backward motion (mean = 31.2; SD = 26.1) than the roll only (mean = 33.7; SD = 37.1) and combined motion (mean = 33.3; SD = 35.1) stimuli. Comparing the roll and combined motion stimuli results showed no significant difference in SSQ scores, contradicting Bonato et al. (2009)'s results. The different motion could explain this result (Posterior/Anterior motion vs pitch) and the non-uniformity of the vection. The method of inducing the motion may also affect the result, with the dot field being a much more unnatural effect than that of the rotating room. However, neither study presents enough evidence to dismiss the other completely.

Content removed due to copyright restrictions

Figure 2-9:- Examples of the stimuli used in Diels and Howarth, (2011) Left (dot field), and Bonato et al. (2009) (rotating cube).

2.3.3 Presence

The objective of a VR experience is to replace a user's existing reality with a virtual one. Presence is the term used to describe the feeling of "being there" within the environment (Sylaiou et al., 2008). Increasing presence within the environment is desirable for VR HMD environments. Belief in the environment helps remove the user from the real world and immerse them in the virtual one. Measuring presence is entirely subjective, as such fluctuations in what each individual determines as "real" will always exist. Schwind et al. (2019) compares three methods for measuring presence in virtual reality environments, including the Slater-Usoh-Steed questionnaire (SUS) (Slater and Steed, 2000), the Igroup presence questionnaire (IPQ) (Regenbrecht and Schubert, 2002) and the Witmer and Singer questionnaire (Witmer and Singer, 1998). Schwind et al. (2019) concludes that these tests are best administered within the VR environment to prevent a break-in presence caused by removing the headset and separating the subject from the environment. Schwind et al. (2019) acknowledges the variance of various scores for the different methods as helpful in detecting breaks in presence but recommends using the IPQ method as the best trade-off between reliability and administration time.

Israel et al. (2019) shows that increased presence also may influence cybersickness, showing slightly lower sickness scores reported in subjects with increased telepresence, resulting in increased enjoyment. It is worth remembering considering the subjective nature of the result collection methods. Generally, people tend to respond more positively to questions when enjoying themselves, so whether this represents a reduction in CS symptoms or an increase in tolerance affecting scores is currently undetermined.

It is essential to acknowledge that presence does not entirely come from the visuals of the headset. Audio Collins et al. (2014) plays a significant role in establishing the physical properties and location of objects in the environment, even when not directly observed by the user. Haptic technologies are another major forefront of VR research (Magnenat-Thalmann and Bonanni, 2006), adding the sense of touch into the environment, giving objects a solid physical presence within the environment.

The environment and its contents can significantly impact immersion and are contextual from one environment to the next. For example, the inclusion of modern computer terminals in a high fantasy environment may be jarring if deployed into the environment without justification for their existence in the environment's narrative (Marcus Andrews, 2010). Believability and realism often go hand in hand. Pouke et al. (2018) shows a weak correlation between graphical realism and an increase in cybersickness symptoms. Participants were exposed to the VR environment for 15 minutes as an on-rails experience. The comparison, shown in Figure 2-10, shows a decrease in texture quality between the high-quality and low-quality environments. The experiment consisted of two groups, one exposed to the high-quality environment and the other to the low-quality environment. However, the evidence provided here is not conclusive due to a low sample size being insufficient to prove the statistical significance, and environmental factors, such as navigating sets of stairs in the environment, possibly influencing the result.

Content removed due to copyright restrictions

Figure 2-10:- Illustration of the Environments Used in Pouke et al. (2018) The top left image represents the high-quality environment. The bottom left image represents the lower-quality environment.

However, the question of *"what is a realistic virtual environment?"* is challenging to resolve. Simulator research conducted before the Occulus Rift's existence had significantly lower environment quality than what is achievable today. For example Nichols et al. (2000) present a study where participants participate in a "shoot the duck" arcade game, illustrated in Figure 2-11. by today's standards, the graphics in this project are incredibly dated, and the framerate (10fps) and refresh rate (4hz) is extremely slow. However, at the time, these graphics were much closer to the cutting edge than today, but as people's expectations for the hardware were much lower, this performance was likely acceptable. Thus, running the same experiment today with the same hardware and software would likely result in a significant difference in SSQ-T results despite identical test conditions due to perceived differences in acceptable graphical quality. Therefore, there is a significant possibility that a characteristic relating to user expectations exists, affecting cybersickness tolerance. With time making what was considered high performance in the past obsolete and even jarring to users, this will inevitably affect the cybersickness response it generates.

Content removed due to copyright restrictions

Figure 2-11:- Shoot the duck game used in Nichols et al. (2000).

2.3.4 Stereo Visualisation

Stereo visualisation allows the perception of depth in virtual environments. Palmisano et al. (2019) show that removing the stereo vision component of VR HMD visualisation protected subjects from the effects of cybersickness induced by head motion and display latency. Therefore, It seems the depth component of the headset plays a significant role in developing cybersickness, providing another vector for the sensory mismatch. Alternatively, the increase in presence offered by the stereo rendering may be impacting the results, and the lack of vergence conflict issues may also be having a significant impact. Removing stereo visualisation from the system does not grant cybersickness immunity as the effect is also present in monovision scenarios (Ujike et al., 2008). Ultimately while removing the stereo vision component of the modern VR HMD may reduce cybersickness symptoms, the loss of depth cues in the virtual environment is likely to be an extremely unfavourable trade-off for most headset users.

2.3.5 Locomotion & Control

Locomotion is one of the biggest challenges facing VR development (Al Zayer et al., 2020; Boletsis and Cedergren, 2019; Steinicke et al., 2013). VR HMD devices' interactions should aim to replicate the actual movement used to navigate the environment's real-world equivalent to minimise the conflict between visual and vestibule systems. The bigger problem is navigation within the environment.

One solution is to provide no options for the user to control their movement within the environment, either having a stationary environment to observe and interact with or having a predetermined path through the environment such as a virtual roller coaster ride as an experience “on rails.”. This approach may be suitable for certain kinds of experiences; however, forced motion can, in cases, add motion sickness if it is unexpected or sudden (Dennison and D’Zmura, 2018). Finally, most experiences desire to give the user some freedom regarding their movement within the environment, restricting this limits potential interaction with the environment.

Headset tracking is a staple feature in modern VR. Tracking the headset position allows the automatic adjustment of the users' viewpoint as they look around the environment (Aukstakalnis, 2017). This tracking system also allows limited navigation via natural motion within the environment space. This form of motion is generally less jarring than other VR locomotion types. However, this locomotion type's limiting factor is the amount of physical play space available, which varies from setup to setup. Once the user reaches the local environment's boundary, another form of locomotion is needed to continue beyond this. Another issue is ensuring the user does not accidentally wander beyond the edge of the play space. Most VR HMD manufacturers have systems such as the guardian system (Facebook Technologies, 2020) to provide digital representations of the edge of the play space to the user, warning them when they stray too close to the edge. However, this is not foolproof, and accidents still may occur. Work in this area is ongoing, and an alternative system is currently in development that transitions the user to a passthrough camera when exceeding the boundary to see the physical world when in danger (Hawthorne, 2021).

Another approach is to attempt to mimic the physical actions used to move within the environment. This mimicry is an attempt to tie the information the senses are providing together into a coherent picture. Arm swinging involves swinging a pair of motion controls by the user's sides to mimic the actions of running. Head bobbing has the user bob their head up and down to move forward, similar to how the head would move during movement (Martindale, 2017). Other physical motion methods do not attempt to mimic realistic motion focusing on functionality over realism. One example of this method is Superman locomotion (Colgan, 2015), where the user extends one or both arms to fly through the environment in the controller's direction. Another currently popular experimental approach is the omnidirectional treadmill (Melnick, 2020), allowing users to simulate walking without moving. While the approach seems viable, the space required for these systems makes them very unfeasible for home use.

Content removed due to copyright restrictions

Figure 2-12:- Visualisation of Oculus Rift Guardian System. The Yellow Fence Represents the Boundary of the Physical Play-space. (https://www.vrfocus.com/2020/05/new-oculus-quest-guardian-updates-will-detect-objects-within-play-area/)

Redirected walking (Razzaque et al., 2001) utilises the principle that blindfolded individuals cannot determine the difference between a slightly curved path and a perfectly straight one known as veering (Kallie et al., 2007). Redirected walking works by continually adjusting the environment's rotation to ensure the user is, wherever possible, walking towards the furthest wall within the environment. This method has two significant issues. Firstly a significant amount of space is required for the algorithm to have time to take effect; at a minimum, the user needs to take a fair few steps before being artificially turned Razzaque et al. (2001) uses a 4x10 meter space, for example. Secondly, the algorithms needed to make this process work assume time is available to reorient the environment between movement periods. Suppose the activity has the user rapidly moving about the environment in different directions. In that case, the rotational algorithm can lag, affecting presence, and vection, resulting in incorrect orientation and users walking into the play space's edges.

Slide locomotion is the technique that relates closest to traditional locomotion from existing 3D game environments. By moving a control stick, touchpad or similar analogue device in a chosen direction, the user slides around the environment. Rotation is achieved either by physical motion within the environment or by "snap turning", moving the stick left or right to turn the user around in the environment in known increments. The two techniques are commonly used together and represent one of the most common locomotion options for VR navigation in use today (Porter and Robb, 2019). Farmani and Teather, (2018) proposed a variation of the snap turn method that eliminates the intermediate frames during rotation, thus reducing vection, with no apparent impact on presence.

Teleportation locomotion operates by specifying a location within the environment and moving the user to that point. The instant teleportation variant does this immediately with no pause. This sudden change in position can be very disorientating for users and reduces presence as the movement is unnatural. Several variants do exist that aim to reduce the impact of the transition. Blink teleportation fades the screen to black before relocating the user and fading back in rapidly. Dash teleportation combines slide locomotion with teleportation by sliding the user to the new location.

2.3.6 Dynamic FOV

Studies have shown that reducing the FOV of a headset can reduce cybersickness symptoms (Lin et al., 2002); however, the associated reduction in presence is undesirable. Fernandes and Feiner, (2016) explored the idea of dynamic FOV changes during gameplay. Dynamic FOV changes are performed either by narrowing the field of view or by blurring the outer edges of the image, thus reducing stimulation of the user's peripheral vision.

2.3.7 Race, Age and Gender

Arns and Cerney, (2005) evaluated 450 visitors to the university's VR lab who experienced their immersive CAVE setup. Having collected SSQ results from participants after viewing, they noticed a significant trend for older users to report higher amounts of sickness than their younger counterparts. This research contradicts findings for traditional motion sickness (Reason and Brand, 1975), which identifies that younger users tend to be much more susceptible, further supporting the argument that cybersickness and motion sickness should be considered separate conditions with the difference in induction methods.

Women are proven to be more susceptible to the effects of cybersickness than men (Stanney et al., 2020). This study attributed the increase to the women having slightly higher FOV than men and thus increased sensitivity to flicker in images. Kolasinski, (1995) also identifies a link between ethnicity and cybersickness Klosterhalfen et al. (2005) has shown that the interaction between race and gender is not linear and rather complex regarding motion sickness, with MSSQ responses being a more reliable indicator of predicting susceptibility.

2.3.8 Individualism

All of the factors discussed here vary from individual to individual. Clifton and Palmisano, (2020) show that while slide locomotion generally generated more cybersickness, a significant portion of participants showed that teleportation locomotion generated increased sickness responses. Therefore, implementing a singular approach to tackle cybersickness is inappropriate. A better approach would be to follow guidelines like those used to develop accessible technologies, providing options to the user that allow them to configure the environment for optimal usage for the individual (Dealessandri, 2020).

Ensuring the HMD fits correctly is also essential. Stanney et al. (2020) shows how a study's results can be easily misinterpreted as a difference between the sexes when the contributing factor was proven to be the headset's inability to support an accurate IPD for a significant number of female participants. This result is concerning as many studies fail to report any measurement of the IPD of subjects, particularly in pre-oculus era studies, showing a significant impact on cybersickness caused by an ill-fitting device.

2.3.9 Experience Comfort Ratings

There is no standard metric for establishing the intensity (likelihood of inducing sickness) of VR experiences which could be an on-rails tour of historical sites or an action shooter.

The two major distribution platforms have provided an indicator system for VR intensity though these are both subjective. Oculus (Facebook technologies, LLC, 2020) gives a comfortable, moderate, intense, or unrated comfort rating. These scores are derived partly by the developers and partly by Oculus themselves. An apparent conflict of interest exists here; both developers and distributors are motivated to sell products and are thus more likely to select a rating desirable to the target audience. The unrated category also provides concern as it allows the circumnavigation of the system. The steam platform has no rating systems; however, a community of users (VR comfort rating, 2020) attempt to rate VR experiences. As with the Oculus system, this group categorises experiences into one of four groups: Green Comfortable, Yellow Moderate, Orange Intense, or Red Extreme. However, the method used to determine these categories is not known.

The term comfort is subjective; however, it refers to a very measurable output; the likelihood and severity of sickness a user is likely to experience. Given that the SSQ appraises sickness symptoms in users of virtual environments, it may be possible to derive a metric from SSQ results collected post usage of several users going through the experience. In practicality, this is unfeasible to do. To ensure the results are reliable enough to predict comfort accurately, the recruitment of a large sample size would be needed to accurately represent the general population and cover all the potential variances in susceptibility. A better approach would be to establish an automated process for capturing data from a VR experience, and then deriving a score for each factor from data automatically collected during the VR experience. This score would provide consumers with a consistent point of reference for determining the intensity of VR experiences. In addition to this, a VR test environment could allow users to subject themselves to stimuli of varying intensities until they reached their threshold for sickness, identifying the score that is acceptable to that user.

2.3.10 Techniques for Reducing Sickness Symptoms

2.3.10.1Adjusting Exposure Time

Research has shown that repeated exposure and gradual exposure to VR stimuli can increase resistance to the effects of cybersickness (Gavgani et al., 2017; Hill and Howarth, 2000; Taylor et al., 2011). In an industrial setting, usage of the devices may become a mandatory part of the job. Eventually, a set of standards for VR usage in the workplace will be released. P2048.5 (https://standards.ieee.org/project/2048_12.html) seems a potential source for this; this standard is still under development. However, in the real world, standards are often poorly enforced. Visual display usage in offices has a host of health and safety instructions associated with their usage (Health and Safety Executive, 2019). However, in practicality, it is rarely convenient to stop the current task and switch to another, particularly as a larger number of office activities become digitised for convenience. Modern VR setups generally utilise motion controls requiring more postural motion than just sitting in a chair and effectively becomes exercise with small weights for long periods and transition from a purely visual display issue to a manual handling one. The risks of manual handling tasks are well documented (Health and Safety Executive, 2020) and represent an additional risk factor when considering the technology's widespread adoption.

2.3.10.2Medication

Several medical remedies exist for the treatment of the symptoms of motion sickness. However, most of these have undesirable side effects such as drowsiness and place restrictions on the type of activity the person taking such medicines can perform (Pollak et al., 2010). Colloquial remedies, such as ginger and magnetic wrist straps, may help but are wildly inconsistent from person to person in their success. With a specific focus on VR HMD, vibration stimulation (Plouzeau et al., 2015; Weech et al., 2018) may prove to be a method of reducing or eliminating cybersickness's effects research into how to implement this effectively is still being evaluated.

2.3.10.3Other Techniques

Wienrich et al. (2018) showed that the simple addition of a virtual nose into the scene significantly reduced cybersickness symptoms without impacting the environment's enjoyment. The study states that the addition of the nose within the environment provides a continuous static point of reference or “rest frame” within the environment (Hettinger and Haas, 2003). While the paper reports this to be effective, this method's comical nature means it has yet to see widespread implementation in commercial environments.

2.4 Cognitive Load

Cognitive load is the term used to describe the amount of working memory currently used by an individual during a problem-solving activity (Miller, 1956; Sweller, 1988).

2.4.1 Types of Memory

Cowan, (2008) identifies different categories of memory that exist.

2.4.1.1 Short Term Memory

Short-term memory is one of the two main types of memory. Wickens and Carswell, (2021) defines short-term memory as the portion of human memory responsible for holding information for the immediate duration until it is either transferred to long-term memory or discarded as unneeded. The duration of the information held in short-term memory is very short, only lasting only a couple of minutes.

2.4.1.2 Long Term Memory

Long-term memory is the other main type of memory. Wickens and Carswell, (2021) defines long-term memory as the portion of human memory that stores information about things that happened more than a few minutes ago. Long-term memories vary in intensity, with common facts used every day being the strongest and easily accessible and weak facts observed once or twice, requiring significant effort to recall. Additionally, long-term memory is not permanent and can fade over time.

2.4.1.3 Working Memory

Working memory is the third type of memory and is a form of short-term memory. Working memory performs decision-making based upon the information passed from the sensory system. Working memory retrieves and stores information into long-term memory as needed in the decision-making process. Miller, (1956) Shows working memory to have a limited capacity for processing information of approximately seven pieces of information (plus or minus two elements) and can be affected by the type of information. Additionally, specialised types of memory exist within these categories.

2.4.1.4 Sensory Memory

Sensory memory processes all incoming information from the human sensory systems such as touch, sound, and vision. This information is processed, and essential elements are sent to working memory to be processed. Anything deemed unimportant is discarded.

2.4.1.5 Relational memory

Relational memory manages relationships between facts in memory (Koenig, 2017). These relationships manage associations between places, things and order. Information is stored as schemas (DiMaggio, 1997) which identify and connect the components of information.

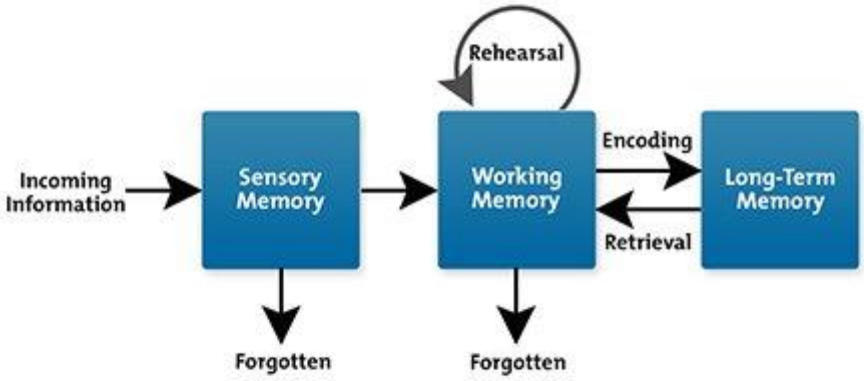


Figure 2-13:- Information Processing Model. Adapted from (https://tinyurl.com/mfbad3zk)

Cognitive load theory (Sweller, 1988) suggests that the designer should control the rate at which information flows to the user when designing tasks not to overwhelm working memory. Halarewich, (2016) shows that controlling information flow protects the user from being overwhelmed with information, which would otherwise negatively affect task performance.

2.4.2 Types of Cognitive Load

2.4.2.1 Intrinsic

Intrinsic load refers to the difficulty of a task (Chandler and Sweller, 1991). Appraising task difficulty is complex as this process must consider the individual's experience with the task, the task's context, and some measure of the objective difficulty. Some tasks are easy to compare on the surface, such as counting to 10 being less complicated than solving Pythagoras theorem. However, if the context is counting 10 moving objects that regularly overlap, the task becomes much more difficult. Comparing task difficulty for two experts in their specific fields is incredibly difficult, considering that brain surgeons consider rocket science difficult. In contrast, rocket scientists would consider brain surgery difficult, as their knowledge and experience do not cover the activity.

2.4.2.2 Extraneous

Extraneous load refers to the method used to communicate information to the user. Chandler and Sweller, (1991) states there are four main methods of communicating information (Indeed, 2020)

- **Verbal: -** The usage of language to communicate. The verbal category includes non-spoken languages such as sign language.
- **Non-Verbal**: - The usage of gestures to communicate, such as facial expressions and body language. It can be intentional or unintentional.
- **Written**: - The usage of written symbols and numbers to communicate.
- **Visual**: - The usage of pictures and diagrams to communicate.

2.4.2.3 Germane

Germane load is induced by constructing and interacting with schemas and refers to the effort required to transfer information from short-term to long-term memory.

Software designers have limited control over cognitive load within the environment. Some forms of cognitive load are controllable by design principles, and some are not. The intrinsic load of a task is not controllable as the task to be performed controls its difficulty. However, by applying appropriate design principles, such as providing contextual help and reducing unnecessary extraneous load by communicating information through common established formats, more memory resources can be dedicated to processing the task increasing the chance of success (Krug, 2021).

Extraneous load is mainly within the designer's control as the designer has the choice of methodology for information communication. However, certain extraneous communication types may not be available due to a variety of environmental constraints. Users lose a significant amount of non-verbal communication abilities within a VR environment due to most modern HDMs, not tracking eye position and facial expression. Many projects are attempting to introduce these forms of communication back into VR environments, such as introducing Tobii eye-tracking systems into the Vive Pro Eye, the hand and gesture tracking in the Oculus Quest or the automated generation of facial movement from voice samples by Facebook (Wei et al., 2019). The designer can also control the germane load by altering the flow of information to provide better and more obvious linkages through sequencing or contextualising information.

2.4.3 Methods for Establishing Cognitive Load

Buettner, (2013) has demonstrated that tracking various physiological factors of the human eye can be used to derive cognitive load. However, two flaws exist with using this approach; firstly, it requires expert knowledge to interpret. Secondly, while eye tracking is a feature of some VR headsets, it is far from the standard and does not track the biometric qualities needed to assess a user's cognitive load. It is not to say VR HMD technologies could not implement sufficient eye-tracking technologies to detect this. However, eye-tracking of sufficient quality is unlikely to appear in mainstream headsets anytime soon due to a lack of mainstream demand.

Alternative methods are more likely to be appropriate for gauging task difficulty. However, these methods do not collect data in real-time. For a subjective measurement, the NASA Task Load Index (NASA TLX) (Hart, 1986; Hart and Field, 2006) is an after task qualitative assessment method that provides an individualised determination of a task's cognitive load. The NASA TLX method sees the subject scoring their experience in 6 categories: mental, physical, and temporal demand, and performance, effort and frustration (seen in Table 2-6). These factors are then presented in a pairwise manner to establish an individual weighting for each category used to determine a final score for the subject's cognitive load. Additionally Harris et al. (2020) proposes a specialised version of the NASA-TLX method for simulation tasks, adding several categories such as perceptual strain and presence to the task assessment process. The results seem promising; however, the method's effectiveness has yet to be widely verified.

Category	Description
Mental Demand	The level of mental activity required by the subject to complete the task.
Physical Demand	The level of physical activity required by the subject to complete the task.
Temporal Demand	The pace of work required by the subject to complete the task.
Frustration	How insecure, discouraged, irritated, stressed and annoyed the subject gets.
Effort	How hard did the subject work to complete the task.
Performance	How successful does the subject think they were.

Table 2-6:- Summary of the NASA TLX Categories.

2.4.4 Causes of High Cognitive Load?

Ferdig, (2009) defines four factors that can induce additional cognitive load.

- Split Attention
  The user's requirement to associate data separated by space (Spatial) or time (Temporal), large gaps in either of these cases can lead to stress demands on working memory (known as the split attention effect) (Yeung et al., 1998). Evidence suggests that working memory may utilise a central pool of general resources (Vergauwe et al., 2012). Having to keep information within working memory longer than is necessary reduces the amount of working memory available for other tasks. For example, requiring the user to memorise a code and enter it into a console at the end of an environment would induce a spatial and temporal load. A temporal load, in contrast, could be induced by displaying the code then removing it before allowing the user to input the code.

- Excessive Information

  Introducing too many elements into the environment can exceed the working memory capacity, potentially resulting in discarding the excess or parts of the previous sample. An example of this would be introducing many objects into the user's view that they must differentiate in some manner.
- Search

  Insufficient knowledge about a task or environment requires the user to add additional load to "learn" how to work within the environment, increasing demands on working memory for this inference. Providing an unfamiliar control scheme within the virtual environment, e.g., an inverted y-axis that requires them to think about movement in a non-natural way would induce this kind of load.
- Redundancy

  Adding information to the interface or environment that overlaps with the user's existing knowledge base can create confusion. The user knowledge base overlaps with provided external guidance, essentially defining a mismatch between what the user already knows and the instructions as to how to complete a task. A typical example of this would be in a game tutorial where a user already knows how to complete an action but is prevented from doing so as it would break the sequence of events in the tutorial; as such, the action performed does not work as expected.

2.5 Issues Regarding Comparing Studies

The biggest issue appears when attempting to draw comparisons between studies. Fortunately, when it comes to reporting results, we have the SSQ. While some have questioned its suitability when related to VR (Kim et al., 2018), it allows an easy and quick general comparison of the severity of cybersickness across stimuli and to some extent across studies. However, inconsistencies in the application of the method may make comparisons more difficult. As illustrated above, many factors potentially impact cybersickness on some level. Studies have shown that the display medium significantly impacts cybersickness frequency and severity (Cobb et al., 1999). However, this issue does not seem to be as prevalent when comparing results from modern VR HMD (Yildirim, 2020). Nevertheless, frequently cybersickness studies fail to consider the differing display modalities when comparing results to older studies.

Risi and Palmisano, (2019b) determined that subjects demonstrating increased spontaneous anterior/posterior motion were more likely to report feeling unwell after exposure to a virtual environment using an HMD. Unfortunately, the conclusion is questionable as rather than using the SSQ, the study opts for a simple metric of verbal response to the question “do you feel well?” to establish if a user is suffering sickness effects or not. This metric is flawed as it provides no granularity or information about the severity of symptoms, forcing subjects to pick one extreme of the scale; sick or well. Also, in subjective studies, what one individual describes as sick another may describe as well; the question “do you feel well” will get vastly different responses depending upon whom you ask. While the SSQ is also a subjective measurement, it is far more granular than this, asking the subject to quantify individual symptoms rather than the overall effect. In this case, looking at the SSQ-D, SSQ-N, and SSQ-O scores in conjunction with the SSQ-T could have potentially shown a link between the separate elements of cybersickness symptoms. The SSQ shows the severity of symptoms, not if a subject is well or not, although a standard threshold for determining “sick” participants is often defined as an SSQ-T value of > 20.

The significant problem is the underreporting of environmental conditions for the experiment. Dennison and D’Zmura, (2018), Dennison and D’Zmura 2017 and Weech et al. (2018) all report the refresh rate of the various devices used in the experiments, for example Dennison and D’Zmura, (2018) reporting the refresh rate of the Oculus headset. This description does not tell us the framerate the application used in the experiment was running at, which is a significant oversight considering how easy developing a poorly performing VR experience is, especially when utilising hardware considered entry-level for VR operation. Ideally, framerate drops should not occur and should, as a minimum, stay above the 72hz range to present smooth environment motion to the subject.

Reporting hardware used in studies helps to give confidence in the results. If the environment's technical details are known, the performance of the environment can be estimated. A better approach would be to have some standard method for reporting environment characteristics. Having this information would allow researchers to instantly remove this element as a variable in their consideration of results. If the experiment uses an Oculus headset, this is readily available via the debug tools packaged with the device (Facebook Technologies, 2021b).

Another source of frustration is uncontrolled stimuli within the studies unless investigating a specific issue such as vection. There is a tendency to use an extreme VR experience such as a roller coaster ride (Bruck and Watters, 2011; Mazloumi Gavgani et al., 2018), which causes issues with consistency between studies. As each study uses a different ride, the two stimuli in the studies are not equal and thus, comparing them is difficult as we cannot determine which is more severe. Also, the fact that rollercoaster experiences are generally rated as extreme experiences and are not recommended for novices due to the extreme speed and rapid changes in direction and are very likely to make users sick. Additionally, these stimuli are wildly inconsistent in the amount of vection a subject is experiencing, with undetectable peaks and troughs throughout the experience.

Two potential solutions exist for this problem. A method for assessing the amount of vection in a VR environment would be an excellent tool as the severity of the VR experience could be quantified on some level. Achieving this would require developing a tool capable of analysing a VR experience's visual output and identifying vection within the environment. Alternatively, the development of a tool integrated directly into the virtual environment recording motion in the environment. This alternative would be a great solution in scenarios where the environment is bespoke developed. It could be just dropped into the environment and provide all the data collection needed. This method would not be feasible for commercially developed products (such as rollercoaster rides).

Another solution would be to nominate an experience (or experiences) as the standard platform for testing, preferably one with the automated test data collection tools embedded in it. This approach would address the issue with consistent amounts of vection within the environment and measure any of the environmental factors associated with cybersickness within the environment at any given point in the test. These measurements are essential as a standard feature of cybersickness studies is to allow the user to stop the experiment should they feel they cannot continue. Therefore, correlating the amount of vection experienced at this point (or slightly prior) would allow for a more accurate picture of the exact stimuli that terminated the experiment, rather than using duration purely as a measure of susceptibility. However, as discussed earlier by Nichols et al. (2000), VR technologies are evolving rapidly. As such, any experience nominated runs the risk of quickly becoming outdated in terms of visual quality. Additionally, constant maintenance of the experience would be needed to support new headsets creating a project with extremely high maintenance costs.

2.5.1 Reliability of Study Comparisons

The vast majority of experiments related to cybersickness are currently very narrow in their focus, looking only at the aspect they are investigating. As already discussed, a wide variety of factors potentially impact cybersickness symptoms and isolating these factors can be problematic. Most studies tend to attribute any increase in cybersickness symptoms observed during testing to their specific factor and then attempt to describe how other factors could influence the results. Problems arise when comparing studies as the potential influence of environmental factors from each study needs to be considered. Additionally, unconsidered factors may pollute the accuracy of the comparison of the results.

The inconsistent environment could be potentially overcome by having all studies utilise the same testing platform. This consistent environment would remove all inconsistencies between stimuli in each study. One way to achieve this would be to nominate a VR experience for usage in all experiments as a baseline. However, picking an off-the-shelf experience, such as a roller coaster or video game, would be insufficient mainly due to such systems' inflexibility. For example, if testing for the impact of vection in VR environments, tight control over the environment's motion amount would need to be maintained. This control is unlikely to be achievable in an off-the-shelf solution unless the stimuli selected had multiple versions with apparent differences in the desired effect, without affecting any other parameters.

Appendix D highlights another problem with study comparisons in the sample sizes. Appendix D shows a sample of multiple studies referenced in this paper. The number of subjects in each paper is often less than 30 and peaks at around 60. With low sample sizes comes the issue of outliers in the data. A sample size of 15 means a single outlier can represent a significant portion of the sample and significantly influence the results, potentially explaining the difference and variance between studies. A few studies focusing on volume may help remove this influence by showing trends in a large sample size; however, recruiting participants for a study with a significant number of samples (100+ participants) will be difficult due to the nature of the research. Few people are willing to volunteer for experiments liable to induce illness, especially without compensation. Additionally, achieving a diverse sample size as VR technologies tend to interest young adults more than their older counterparts.

A possible alternative may be the development of an open-source testing platform for cybersickness experiments. The development of a baseline experiment would provide a consistent point of comparison between studies. The adjustment of the environment's features would allow the investigation of individual factors while allowing a more accurate comparison with other studies' results and the initial baseline result. Additionally, having access to the environment source would make integration and assessment of experimental techniques easier. Finally, this process would standardise the data capture of these projects, and as such, the quality of information gained from the investigation would likely rise. Additionally, re-evaluating factors that were initially considered irrelevant would be possible as a complete picture of the data would be available after the fact.

One problem with an open-source solution is that it requires maintenance to remain relevant. This maintenance creates a need to constantly update the software to support new headsets and graphics technologies as they come onto the market. Additionally, as VR technologies are an emerging technology (Ludlow, 2015), the possibility exists of a significant change in the method of displaying VR content, such as the adoption of light-field displays (Lang, 2020). As such, the maintenance effort is likely to snowball quickly on this type of project to unmanageable levels.

Additionally, modifications to the environment require expert knowledge to implement. Generally, VR environments utilise game engines, such as Unreal or Unity, due to their proficiency in rendering 3D environments. This fact means that some knowledge of these tools is required to modify these environments. Visual scripting is becoming more prominent in these environments, with Unreal’s blueprint system (Epic Games, 2018) and Unity's Bolt (Unity Technologies, 2021). These systems' usage may reduce some of this burden as only simple changes would be needed to change common factors such as FOV and movement speed. However, more complicated issues such as environment realism and novel locomotion methods will require expert knowledge to achieve.

2.5.2 Cognitive Load and VR

In 2019 Mittelstaedt et al. (2019) demonstrated the impact of utilising different VR display methods (Large VR TV display and VR HMD devices) and control methodologies on task effectiveness after VR exposure. Mittelstaedt et al. (2019) They observed that the reaction times of subjects were significantly lower after VR exposure than before it. The study demonstrated a low correlation between this effect and the severity of cybersickness symptoms experienced. Previous studies such as Nalivaiko et al. (2015) and Nesbitt et al. (2017) have suggested that cybersickness causes detrimental cognitive performance. This finding is reinforced by Mittelstaedt et al. (2019), which shows that the utilisation of the VR headset impacts cognitive performance in many areas such as reaction time (Nalivaiko et al., 2015) and mental rotation (Levine and Stern, 2002). However, without knowing the effect of the environment on cybersickness symptoms, it is impossible to state for certain that the effect observed is entirely attributed to VR HMD usage.

When comparing results between studies, it becomes difficult to provide an accurate comparison, as each study uses a stimulus that induces cybersickness differently. Looking at Mittelstaedt et al. (2019) and Nesbitt et al. (2017) as an example, the former uses a virtual bike ride while the latter utilises a virtual roller coaster and reaches different conclusions about the impact of cybersickness on cognitive performance post-VR experience. Determining which study is correct is complicated due to inconsistent stimuli. This difference led to significant differences in average sickness scores for each study. Interpreting these results is further complicated by the severity of the stimulation being non-uniform. Therefore, cybersickness effects may have been introduced by the different severity transitions. With traditional motion sickness, exposure to the stimuli over an extended period leads to reduced symptoms (Graybiel and Lackner, 1983; Zhang et al., 2016) as habituation occurs. During habituation, the human brain is learning to ignore the conflicting stimuli and adapts to the new situation. The same process occurs during VR usage, and exposure is known to induce habituation effects (Gavgani et al., 2017; Heutink et al., 2019; Hill and Howarth, 2000). These studies show that habituation applies to the test environment. However, both studies failed to demonstrate if this habituation transfers to different environments and scenarios as it would with traditional motion sickness.

Habituation is known to occur in VR experiences. However, no evidence exists to identify if this habituation applies to different VR setups and experiences. Testing this issue requires a new way of thinking as isolating factors is impossible due to the real and virtual variance within each experimental environment. At its most superficial level, this problem could be analysed by merely having users experience a single VR experience until evidence of habituation appears, then switching to a different environment and seeing if sickness levels return to an average level. The problem with this approach would be that the two environments would probably induce cybersickness in different intensities. Thus, changes in the SSQ ratings cannot be determined to be solely caused by unfamiliarity with the environment as differences within the environment could affect cybersickness symptoms.

2.6 Discussion of Relationships

2.6.1 Cybersickness and Postural Stability Measurement

Therefore ideally, monitoring of cybersickness should require nothing more than the equipment that comes in the box. One possibility exists that may achieve this. Postural stability theory suggests that instability occurs in response to motion sickness. Multiple studies (Dennison and D’Zmura, 2018; Murata, 2004; Rebenitsch and Quinby, 2019; Risi and Palmisano, 2019a; Villard et al., 2008) have all demonstrated in various ways that additional instability is present in subjects experiencing sickness symptoms.

A common approach to this is to measure the user’s centre of pressure. Risi and Palmisano, (2019a) shows how sick and well participants have differences in their pre-test spontaneous postural motion, suggesting susceptibility to cybersickness may be detectable by measuring balance before exposure. Risi and Palmisano, (2019b) demonstrates a correlation between cybersickness severity and spontaneous postural measurements. However, the lack of any validated method for determining sickness makes these results questionable. Arcioni et al. (2019) shows a link between spontaneous postural instability and cybersickness in users. While this finding is useful Arcioni et al. (2019) acknowledge that the study's motion only affects the head. Common VR usage involves head and body motions. Therefore the motion mismatch effect is not fully represented in this study.

Content removed due to copyright restrictions

Figure 2-14:- COP measurement made using the Nintendo Wii Balance board (Widdowson et al., 2019).

Dennison and D'Zmura, (2017) presents evidence that cybersickness can occur without any associated increase in instability measured either via the headset or the balance board. Further evidence from Dennison and D'Zmura, (2018) shows that instability may come from unexpected motion and still show no correlation with cybersickness. The experiment used a commercial rather than scientific device – the Wii balance board – which utilises four load cells in the device's corners to approximate a pressure measurement (Bartlett et al., 2014). As such, the accuracy trade-off made when using the Wii Balance Board may mean the device can no longer accurately detect the subtle motions needed to identify sickness. However, the Wii Balance Board is an appropriate substitute for study areas where absolute precision is not required (Weaver et al., 2017).

Using a pressure plate in a real-world scenario is unfeasible and unrealistic due to the hazard it introduces to the play area and its requirement for the user to remain standing in a small area. Multiple studies (Dennison and D'Zmura, 2017; Jin et al., 2018; Rebenitsch and Quinby, 2019) attempt to use or express an interest in using the headset's data to identify instability and sickness. Rebenitsch and Quinby, (2019) demonstrates that it may be possible to detect cybersickness from HMD position alone, showing that participants exhibiting cybersickness symptoms also showed an increase in the amount of HMD roll motion they exhibited. The experimental setup for this utilised an external system (InterSense IS900), which has questionable accuracy (Gilson et al., 2006), to track the HMD movement and provided evidence that increases in head movement correlate to the onset of cybersickness symptoms. Jin et al. (2018) used machine learning to identify cybersickness in subjects from HMD position and external video data, using a wide range of stimuli to model responses and provide an annotated training set. The sampling frequency was low (2hz), indicating that this may not be appropriate for detecting instability that a pressure plate would. Also, the test scenario saw the Test subjects were seated throughout, likely impacting the subject's stability compared to a standing scenario. Jin et al. (2018) did demonstrate some success with this approach; however, the video data requirement makes this method unlikely to be feasible in the real world due to additional hardware performance requirements.

3 Virtual Environment Design

The proposed virtual environment aims to provide a platform to assess the issues identified during the literature review. The environment must be capable of isolating the methods of cognitive loading during VR HMD usage. By isolating this factor, an insight into the interaction different types of cognitive load have on the onset of cybersickness can be gained. The test platform must present scenarios as two separate components: one with a high mental load and the other with a low load and allow for the separation of cognitive load sub-components. Each condition will need to capture subjective data relating to the subject's perceived levels of task performance and cybersickness symptoms.

Additionally, the environment must capture data from the headset relating to the headset position to assess changes in motion during VR usage. Importantly this measurement must be conducted in such a way that does not impact the regular usage of the VR HMD. This factor is critical to maintaining a realistic usage scenario of VR usage currently not represented during most postural stability studies assing the impact of cybersickness on postural sway.

Finally, the environment must collect all of the necessary data without requiring expensive additional equipment, equipment requiring expert knowledge to interpret the results or measurement techniques unable to be performed in real-time alongside a realistic VR program. These limitations aim to prevent the development of a scenario that could not be realistically and practically implemented into the majority of VR HMD setups.

3.1 Test Environment Considerations:

To produce a test environment that is appropriate to determine if this VR habituation is true, we need to be able to alter independently:

- The environment in which the test is performed in terms of:
  - Visual complexity
  - Temporal complexity
  - Spatial complexity
- The task the individual undertakes in terms of:
  - Physical complexity

    - Mental complexity
    - Repetitiveness
- The aspects of memory controlled for:
    - Use of short and long-term memory
    - Task flow
    - Task presentation
    - Schema supporting components

One way to achieve this may be through the adaption of the difficulty of the task, the inclusion of a variant of an existing task, or only requiring completion of a task more quickly. An example of this would be a simple target selection task: starting with targets of a single colour to be picked from a selection of two then increasing the number of colours to four. This will increase the task complexity while leaving cybersickness-inducing factors and most memory aspects identical between them. This method would create two distinct test scenarios with identical cybersickness induction.

Cognitive load's impact on cybersickness is not fully understood, and the assumption that cognitive load does not impact cybersickness would make separating components more difficult. Armougum et al. (2019) demonstrated no difference in cognitive load between identifying information in a train station and a virtual simulation (Figure 3-1). This study used an environment designed to minimise cybersickness symptoms, and subjects exhibiting cybersickness symptoms were withdrawn and replaced. This study suggests that virtual environments are comparable to their real-world counterparts in terms of complexity and would seem to suggest no major impact of an average task on cybersickness symptoms.

Content removed due to copyright restrictions

Figure 3-1:- Virtual vs Real-life Train Station used in Armougum et al. (2019)

However, the task performed in this experiment is at no point mentally complex, just requiring perceptual observation of the environment and navigation of the train station. Introducing complexity into the environment can be achieved in multiple ways, such as used in Andersen et al. (2018), which evaluated the impact of distributed simulation on cognitive load on subsequent training using a surgical simulator. In this case, the experiment provided load through the task's complexity rather than information search within the environment. Andersen et al. (2018) also demonstrated that the cognitive load in a virtual reality environment could be reduced with repeated exposures to an environment, thus suggesting users can learn an environment. However, a significant number of participants had prior virtual reality experience and, therefore, may have already acclimatised to VR HMD usage. Additionally, the study does not measure cybersickness symptoms in subjects. Therefore, we cannot determine if the lower load evidenced in this study also reduced incidents of cybersickness.

3.2 UI Design in VR

These factors are all primarily controlled by software design. In an application setting, split attention issues can be solved by placing associated data together, regulating the flow of information and interactive elements to the user to a reasonable rate, providing sufficient initial and contextual instruction within the environment and ensuring consistency of interaction methodologies between environments. The benefit of adopting this approach is shown by Paas and van Merriënboer, (2020).

However, when considering VR environments, the ideal positioning for information is more difficult to achieve. The user interface is the primary way of providing information to the user in a prompt manner. User interface elements fall into four categories, primarily dictated by their physical presence in the virtual world and its integration into the fantasy. Fantasy, in this case, refers to the environment the virtual world is trying to represent.

- Diegetic
  Diegetic user interfaces are incorporated into both the game world's fantasy world and have a physical presence within the game world. An example would be an ammo counter on a futuristic rifle or a watch on the character's wrist showing the time. Diegetic interfaces aim to communicate the information realistically and seamlessly. Figure 3-2 shows an example of a spatial UI element highlighted in area A, an ammo counter integrated into the player's weapon's body.
- Non-Diegetic
  Non-diegetic interfaces have no connection to the fantasy of the game world or any position within it. Non-diegetic interfaces are the traditional UI sources, including text boxes, sliders, buttons and combo boxes, to name a few. Non-diegetic interfaces aim to communicate information as effectively as possible. Figure 3-2 shows an example of a spatial UI element highlighted in area B, a radar representing the location of enemies around the player.
- Spatial
  Spatial UI elements are positioned in the game world but have no connection to the game world's fantasy. Typical examples include racing line indicators and navigation points. Spatial UI elements aim to provide information about a specific place in a prompt manner. Figure 3-2 shows an example of a spatial

UI element highlighted in area C, a name tag hovering over a player's character to identify the player.

- Meta

  Meta UI elements are connected to the environment's fantasy but have no physical presence within it. Common examples include the overlay of blood effects and dirt on the screen to indicate damage. Meta interfaces communicate essential information passively and realistically to the user. Figure 3-3 shows an example of a meta-UI element with a red haze surrounding the player's view, representing the player being close to death.

Content removed due to copyright restrictions

Figure 3-2:- An Example of a Traditional Video Games User Interface. (https://venturebeat.com/2015/10/25/halo-5-guardians-takes-master-chief-and-his-pursuer-down-a-very-strange-path/)

Content removed due to copyright restrictions

Figure 3-3:- An Example of a Meta UI Element. The red haze around the edge of the display is caused by the player taking damage (https://support.activision.com/call-of-duty--wwii/articles/call-of-duty-wwii-campaign)

Content removed due to copyright restrictions

Figure 3-4:- Example of Doom Eternal UI (https://guides.gamepressure.com/doom-eternal/guide.asp?ID=53665)

As gaming is an everyday use for VR, comparing traditional game interfaces to VR game interfaces will highlight the difference. Modern traditional video games tend to rely on non-diegetic UI elements. Using Doom Eternal as an example (see Figure

3-4), most of the information the player needs, such as health, ammo and shields, are provided in an immediately accessible format in the corners of the display.

Non-diegetic interface elements simply do not transfer to VR environments very well at all (JiaxinWen, 2017; Salomoni et al., 2017; Winestock, 2018). As they have no depth, and the display encompasses the whole environment. If applied in the same manner, these elements are essentially “stuck to the users’ eyeballs. This effect makes them floating barriers, mostly unreadable and too uncomfortable to view. A common compromise is to convert these elements into spatial elements floating the UI elements at a certain distance from the user but in the same configuration. Developing interfaces, however, is easy to do poorly. Placing UI elements at an inappropriate distance from the user makes them difficult to focus on, causing eye strain and discomfort. Additionally, as UI elements move further away from the user, the amount of screen space they occupy is reduced. As such, significantly less detail is resolvable within the UI element (Purwar, 2019). In traditional UI design, placing information far away from where it is needed increases the task's spatial load, but in VR, not only must this be considered but also the position of these elements relative to the user greatly impacts their useability. Figure 3-5 shows the range for comfortable viewing in VR HMD. Placing elements beyond this range places stress on the user's neck and is uncomfortable to view.

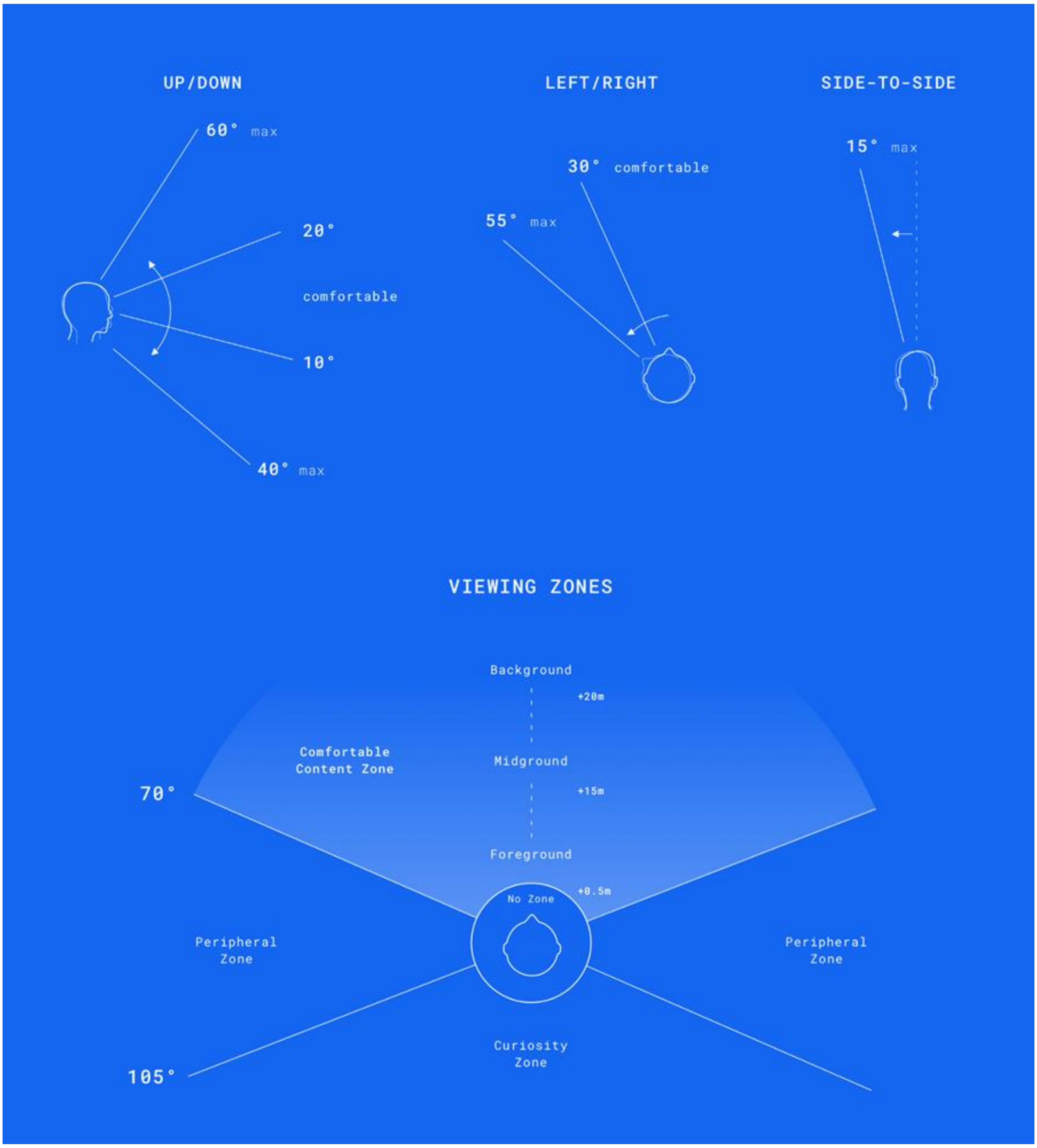


Figure 3-5:- Comfortable Range for VR Elements (Purwar, 2019).

This approach of having floating UI elements hovering around the player is better but mostly unrealistic and often causes other problems. However, this is often a quick fix for porting non-VR games to VR and often results in a poor user experience (See Borderlands 2 VR as an example (Lang, 2018)). Generally, UI in modern games is scalable to screen size, ensuring they do not intrude into the game area. Position of VR, the UI elements are embedded within the environment. As such, the environment dictates the scale.

Additionally, if the element's position is not relational to the player, then the display space occupied will change as the user moves around the space. As such, the developer must strike a balance between readability and physical space occupied. This approach can be successful with things like a floating health bar in front of the user, which does not occupy excessive screen space. Working elements requiring large amounts of space, such as mini maps, often obstruct too much of the screen.

Good UI design can also lower the cognitive load on the user (Yablonski, 2019). Minimising unnecessary elements and tasks can yield significant reductions in the effort required to use a program or piece of software. Additionally, the usage of common design patterns can reduce cognitive load by providing familiar constructs to reduce the task's extraneous load. Achieving this is difficult in VR as the field is relatively new. Identification of these standard workflows has yet to occur. Therefore, each application tends to have a unique method of interaction. Research such as Smith et al. (2019) is working on methods for evaluating VR experiences' useability, hopefully, to guide the development and assessment of the suitability of VR User interfaces and their usability.

Another consideration to make is the usage of central and peripheral vision in humans. Generally, when designing interfaces, the aim is not to keep the central area clear to display content and move controls off to the edge. Doing this leaves a clear area for the user to focus on while leaving controls easy to access and able to draw the user’s attention when necessary through the usage of motion or colour (Ilieva, 2016). In VR, the display quality is not uniform across the viewing area, especially if utilising techniques like foveated rendering are used, resulting in a non-uniform distribution of image quality across the display area (Patney et al., 2016).

When done right, diegetic interfaces in VR feel natural and provide convenient and logical access to information. For example, communicating health is commonly integrated into the hand or arm of the character, where it can be observed at a glance comfortably and consistently (see Figure 3-6 for an example).

3.3 User Experience in VR

Diegetic elements enhance immersion, and the purpose of most objects are communicated naturally via their function. For example, picking up a hammer should feel like picking up a hammer and it should behave like a hammer when swung. Therefore, if the user interface is correct, the hammer should be used as a hammer. For example, if swung at a fragile surface (such as glass), the surface should break. Actions and interactions should be as natural as possible. Where this is not possible, the information should be presented contextually at the point of integration. For example, gesturing at a key could provide a popup, illustrating details about the key, such as the doors it opens. This process lowers the cognitive load and interfaces clutter while simultaneously allowing immediate delivery of the information when and where it is needed.

Content removed due to copyright restrictions

Figure 3-6:- Example of Diegetic Health Meter in Half-Life: Alyx. The highlighted area shows the health meter integrated into the technology of the glove. (https://www.theverge.com/2020/2/13/21136678/half-life-alyx-launching-march-23rd-date-valve-announcement)

3.4 Environmental Design

3.4.1 VR Rollercoasters

Virtual roller coasters are commonly used in experiments as they generally provide a very provocative stimulus for those who will suffer from cybersickness. While most of these similarly provoke cybersickness on the surface, there are many environmental differences. Generalising, virtual roller coasters use fast simulated motion to induce the sensation of movement within the environment. However, each experience utilises the environmental characteristics, e.g., speed, drop height, loops, inversion time, and drop angles differently. The inclusion of a factor in one ride which is absent in another may induce cybersickness in different ways, making it difficult to attribute the induced sickness to the investigated factor alone. Environmental presence (Pereira Junior, 2021) may also induce emotional states such as fear, resulting in a unique response from each participant. In these cases, it becomes difficult to separate the effects of VR exposure from environmental effects and cybersickness symptoms from mental/emotional responses such as anxiety.

Davis et al. (2015) investigated the difference in sickness levels between two VR roller coaster rides using the Oculus Rift platform. Observations highlight differences between the two environments, including visual fidelity, optical flow, and required interaction which may have affected the subject’s cybersickness severity. The difference in cybersickness severity found in Davis et al. (2015) indicates that experiments not using the same environment must take the baseline cybersickness induction level into account when verifying their results. Factors such as Graphics, Draw distance, tempo, smoothness and interaction can all differ from one experience to another.

To validate the considerations in section 2.5.2, two rollercoaster experiences were selected for comparison. Rollercoaster VR Universe (OVERBUILDED, 2020) and Summer Funland (Monad Rock, 2018). These two roller coaster experiences were chosen from popular roller coaster titles on the Steam platform. Their features are highlighted in Table 3-1.

Content removed due to copyright restrictions

Figure 3-7:-Summer Funland (https://store.steampowered.com/app/780280/Summer_Funland/)

Content removed due to copyright restrictions

Figure 3-8: - RollerCoaster VR Universe (https://store.steampowered.com/app/1070690/RollerCoaster_VR_Universe/)

Therefore, while the task is the same, the experiences are very different, with Summer Funland having greater visual fidelity and frequent cinematic effects creating a higher sensory load. In contrast, Rollercoaster VR Universe has lower visual quality but increased vection from the unexpected motion from the jerky nature of how the coaster changes direction within the environment.

Therefore, despite the rollercoasters appearing identical on the surface, clear differences exist between the environments, which affect cybersickness (shown in section 2.3). As the selection of these two rollercoasters was random, it is unlikely that any two roller coaster ride experiences will be completely identical. Therefore, consideration needs to be given for studies using these as their stimuli for inducing cybersickness.

		Rollercoaster VR Universe (City)	Summer Funland
Environment	Visual Complexity	Low Realism and Fidelity	High Realism and Fidelity
	Temporal Complexity	Long pauses, with Intense Bursts	Uniform Complexity
	Spatial Complexity	Left Turn Bias Jerky turns Open space Low Peril	Slight Right Turn Bias Smooth turns Frequent Cramped Spaces Frequent Peril
Task	Physical	No difference	No difference
	Mental	Preparation via smooth movement	Objects dropped in path,
	Repetitiveness	Ability to predict movement	Random jerks
Memory	Short and long term	No difference	No difference
	Task Flow	No difference	No difference
	Presentation	No difference	No difference
	Schema	No difference	No difference

Table 3-1: - Comparison of Rollercoasters, Rollercoaster VR Universe and Summer Funland

3.4.2 Case Study:- Pavlov Vs Zero Caliber

Having used the framework from 2.15 to compare two primarily motion-driven static experiences, there is a need to consider interactive task-driven experiences. Two VR first-person shooter games were selected as they require spatial navigation, temporal activities, and physical interaction. The selected games Pavlov VR (https://store.steampowered.com/app/555160/Pavlov_VR/) and Zero Caliber (https://store.steampowered.com/app/877200/Zero_Caliber_VR/) were chosen as two of the most popular games in their genre.

On the surface, these games appear very similar in terms of what they do. Both:

- are VR first-person shooters
- use realistic weapons
- use real-world interaction mechanisms, e.g., scopes and reloading
- make use of slide locomotion
- have the same task: shoot the “bad” guy.

Content removed due to copyright restrictions

Figure 3-9:- Zero Caliber (https://store.steampowered.com/app/877200/Zero_Caliber_VR/)

These games though superficially similar present differences in many areas.

- Graphics

  Appraising the difference in graphical quality, in this case, is not as easy as it was for the VR rollercoasters. Comparing graphical quality is easy when looking at similar experiences on separate platforms, as a series of side-by-side comparisons can be performed (Swider et al., 2021). However, in the case of comparing two different games, this becomes difficult without identifying obvious deficiencies in the graphics quality of one. Graphics style (Keo, 2017) can give different games vastly different graphical appearances yet still be acceptable to the subject. Looking at the games in terms of factors that induce cybersickness, no obvious differences can be identified between the two games.

1. Locomotion

   While both games implement slide locomotion, the actual application of this is significantly different in their practical application. Considering the movement systems as shown in Table 3-2, differences exist between the two games in areas known to impact cybersickness (Al Zayer et al., 2020; Clifton and Palmisano, 2020; Step, 2017).

Factor	Pavlov	Zero Caliber
Locomotion Method	Physical Turning	Snap Turning
Locomotion Speed	Fast	Medium
Movement system specialities	Non-dominant hand dictates direction.	Slows movement while aiming
Vertical Travel	Slopes, ramps and stairs	Free climbing

Table 3-2:- Movement Characteristics of Pavlov Compared to Zero Caliber

2. Game Task

   Comparing the default experiences, Pavlov sees users fighting real people that are incredibly unpredictable compared to the AI-controlled enemies in Zero Calibre, which are very predictable, filing out from specified points, standing and shooting at the player. The design of levels between the two games shows significant differences. Levels in Zero Caliber are linear in design, not providing many options for attacking a target. Enemies generally attack from a single direction, and opportunities for surprise attacks are limited. On the other hand, Pavlov has levels designed to provide choices in how the player attacks the enemy team or defends an incoming attack from multiple possible avenues. This likely results in significantly different approaches from the player (Lubin, 2016). Pavlov is a very reactionary game with potential attacks coming from all directions, requiring the player to "keep their head on a swivel", always observing all areas of the environment, likely adding a significant amount of vection. In Zero Caliber, by contrast, AI (Artificial Intelligence) opponents funnel out from predefined points in the environment in front of the player and run in predictable lines to attack the player.

This difference in gameplay results in less need to check other areas of the environment, thus providing less vection and reducing the task's spatial search compared to Pavlov. Additionally, as attacks can come from any point in the environment, creating schemas for opponents' attack vectors is much more complex as strategies vary immensely from player to player.

3. Weapon Interaction

The games have reload mechanics that follow different interaction paradigms but follow a similar sequence:

- Remove the magazine
- Acquire the new magazine
- Insert the magazine
- Charge the weapon.

Pavlov provides a more realistic approach requiring physical actions to perform each action in the sequence, while Zero Caliber Short cuts this by allowing the player to place a new magazine into the gun while automatically ejecting the old one, skipping step 1. This skipping of step one highlights how an interaction system can require different thought levels and provide error tolerance. Other examples of differences exist. For example, muzzle climb is handled by detaching the weapon's front point from the motion controller's point in Pavlov. While this is a more realistic choice for a gun, it introduces the potential for disassociation from the controller and may impact immersion.

Thus, the Zero Caliber system includes fewer tasks with lower variance than those in Pavlov. Thus, cognitive load during reloading tasks is significantly different, with Pavlov having a significantly higher Intrinsic load than Zero Caliber. Additionally, aiming and handling Pavlov weapons is much more complex than Zero Caliber, increasing the task's intrinsic load.

In conclusion, by comparing the two games, significant differences between the tasks are evident (see Table 3-3), and some do have a significant impact on factors affecting cybersickness. However, differences in the complexity of seemingly identical tasks are also present in significant quantities. Knowing that an increase in cognitive load reduces task effectiveness (Sweller, 1988). We know users will generally perform worse in high cognitive load scenarios. However, the influence of cognitive load on cybersickness has not been confirmed. This fact aside, the differences in cognitive load may impact user behaviour, introducing differences in behaviour that impact cybersickness symptoms, resulting in high sickness levels.

		Pavlov	Zero Caliber
Environment	Visual Complexity	None	None
	Temporal Complexity	Aggressive movement on attack Waiting and observing on defence	Constant stream of targets, consistently aggressive
	Spatial Complexity	360º attack angle	High number of targets
Task	Physical	Additional complexity in mechanics	Button shortcuts for many actions
	Mental	Complex reloading system	Simplified reloading system
	Repetitiveness	Unique task each encounter	Similar task, repeated multiple times
Memory	Short and long term	Memorise common attack and defence strategies over multiple playthroughs.	Memorise enemy positions over multiple games
	Task Flow	None	None
	Presentation	None	None
	Schema	Large variety of possible attack vectors from opponents	Easily predictable attack vectors

Table 3-3:- Comparison of VR FPS (First Person Shooter) Games, Pavlov and Zero Caliber

In contrast to the differences between the VR rollercoasters investigated earlier, Zero Caliber and Pavlov's differences do not come from the environment. Although there are some differences in the environment, most of the difference between the two games comes from interactions and gameplay. However, to what extent do these differences in cognitive load required to interact with these environments influence cybersickness? Furthermore, how different do these factors need to be; before an influence on cybersickness symptoms can be identified?

3.5 Cybersickness Impact

This experiment aims to isolate the contribution of cognitive load on cybersickness symptoms. Achieving this will require minimising the contribution of know factors of cybersickness.

3.5.1 Environment Realism

Pouke et al, (2018) demonstrated a link between environmental realism and increasing cybersickness symptoms. While the link is weak, there is a possible link between the two factors, therefore, to remove the possibility of influence from this factor, a simple environment made up of simple shapes and basic textures.

3.5.2 Latency

Latency occurs in two areas, rendering performance and headset tracking. Headset latency has a significant and demonstratable impact on cybersickness symptoms (Stauffert et al., 2018). Latency primarily occurs when the machine's performance is inadequate, or the workload required to represent the scene is too large. Latency can also exist in the headset's tracking performance and requires sufficient processing power to represent the headset's position during usage accurately. Therefore, extensive testing is needed to ensure the performance of the hardware is sufficient for the task.

3.5.3 Framerate

The framerate of the application is a significant contributing factor to cybersickness (Stauffert et al., 2018). If the hardware used for the experiment failed to match or exceed the headset's refresh rate, the result would be the introduction of latency and dropped frames, both of which significantly affect the environment's presence and contribute to cybersickness. Therefore, the framerate will be monitored throughout the experiment, discarding any areas where the framerate drops below 90fps during data analysis. This framerate was chosen as it the maximum refresh rate of the headset. Achieving anything lower would result in an inconsistent level of performance.

3.5.4 Presence

As shown in research, presence negatively correlates with cybersickness symptoms (Sylaiou et al., 2008). Therefore, to reduce cybersickness, our environment should be believable and engrossing in its approach. For an environment to be believable, the environment does not necessarily need to be a realistic representation of the real world. It must be convincing enough to make the subject believe they exist within the environment.

3.5.5 Vection

Vection is motion within the environment; a limited amount is desirable and positive; however, too much induces cybersickness (Stoffregen et al., 2008). The induction of vection into the environment occurs in two forms. Firstly, the number and speed of moving objects in the scene influence the amount of vection and can be controlled by not utilising large numbers (30 +) of fast-moving objects. The second source is through the headset motion; rapid rotation of the headset effectively applies vection to all objects in the scene and thus should be avoided.

3.5.6 Field of View

The headset primarily dictates the field of view. Simultaneously, artificial restrictions could be placed upon the headset's field of view to bring all headsets to the same field of view. In this regard, the exercise of making the headsets equivalent becomes a snowball requiring matching of all parameters to remove their impact on the experiment. The simplest way of ensuring consistency is to pick a single model for this research exercise. Therefore, picking a standard headset would be necessary to ensure access to the environment for other testing purposes.

3.5.7 Locomotion Methods

Locomotion is a complex issue with many different solutions and has a complex interaction with cybersickness symptoms. Several options exist for locomotion within the environment, and no matter the choice, there will be some impact on cybersickness. Therefore, the experiment will only offer a single locomotion method, specifically slide locomotion, which is the most widely used.

3.6 Designing a consistent and appropriate stimulus

The stimuli developed for this experiment need to represent the full range of tasks performed within a VR environment. Therefore, a complete survey of VR technologies' current applications is required to determine the tasks required to represent a full and diverse VR environment accurately.

3.6.1 Virtual Reality Games

Virtual reality games are the biggest usage case for virtual reality technologies. The steam hardware survey conducted in June 2020 (Valve Corporation, 2020) shows that 1.67% of users have a VR headset available to them. This value represents just over 1.5 million users with a VR headset. Table 3-4 shows a list of game categories available from the 237 VR games on Steam. Presented below is a review of the typical features of each game type.

3.6.2 Action

Action Games incentivise physical activity and primarily test player's reaction times and hand-eye coordination, with examples in traditional games series such as Super Mario, Halo and Grand Theft Auto. The nature of VR lends to action games almost intrinsically; with motion controls requiring physical movement, the vast majority of games require some physical interaction. Action games in VR such as Beat Saber (https://beatsaber.com/) and Pistol Whip (https://cloudheadgames.com/pistol-whip/) make this movement the focus of the game, encouraging physical movement and making the player part of the physicality of the game.

Many action games utilise realistic tools, commonly firearms and weapons. Which are generally interactable elements that can be picked up, dropped and thrown around the environment. Interaction with tools within the VR environment should mimic their real-world counterparts, e.g., the user swings a sword into the enemy rather than pressing a button. Enjoyment is the game's typical focus, so tools are generally simple to use and easily identifiable in their function, allowing a pick-up and play game style.

Movement is generally fast-paced, having fast-moving objects to be dealt with or requiring the player to navigate the environment rapidly. This movement commonly uses slide movement, although other locomotion options are usually available such as teleport. Player orientation is either achieved via snap rotation or real-world physical rotation.

Category	Quantity	Game Examples
Action	137	Beat Saber, Blade and Sorcery, BONEWORKS, Pavlov VR, STAR WARS^{TM}: Squadrons, X-Plane 11, Assetto Corsa Ultimate Edition, VTOL VR, Superhot VR
Simulation	136	Assetto Corsa Competizione, Blade and Sorcery, BONEWORKS, STAR WARS^{TM}: Squadrons, Tabletop Simulator, X-Plane 11, Assetto Corsa Ultimate Edition
Adventure	85	Blade and Sorcery, BONEWORKS, X-Plane 11, Half-Life: Alex, Five Nights at Freddie's, Elite Dangerous, Vacation Simulator, L.A. Noire: The VR Case Files.
Horror	47	BONEWORKS, Half-Life: Alex, Five Nights at Freddie's, Hellsplit: Arena, The Walking Dead: Saints & Sinners, Budget Cuts
Sports	46	Beat Saber, Assetto Corsa Competizione, Dirt Rally 2.0, Tower Tag, The Thrill of the Fight – VR Boxing, VR Regatta – The Sailing Game, Box VR
First Person Shooters	41	BONEWORKS, Pavlov VR, Half-Life: Alex, SUPERHOT VR, Hotdogs Horseshoes and Hand Grenades, Pistol Whip, Tower Tag, The Walking Dead: Saints & Sinners
Driving	17	Assetto Corsa Competizione, Assetto Corsa Ultimate Edition, DIRT Rally 2.0, City Car Driving, Derail Valley, Project CARS 2

Table 3-4: - A List of Categories of VR Games Currently Available on Steam.

3.6.3 Simulation

Simulation games aim to accurately represent a task, with the most common application being racing games, but any task, imaginary or real, can be simulated. Non-VR examples include Harvest moon, EVE online, and the Forza racing series. VR simulation tends to focus more on introducing realism via actions. For example, in the shooter subcategory, weapon reloading is traditionally handled via a button press. VR simulation shooters force the player to actively engage in the reloading process, ejecting the magazine, inserting a new one and charging the weapon.

Simulation games are focused on accurately representing the environment and the task. These games have complex interactions and are focused on generating immersion rather than the practicality of control. Therefore, interactions are designed to enhance immersion and replicate what the user would be doing in the environment and how they should be doing it.

Locomotion in simulations varies widely by the implementation. Many experiences, such as rollercoasters and driving simulations, require no environment navigation, mealy tracking the headset position and tracking the user’s viewpoint within the environment. In any case, navigation in simulations is generally slower than that of other game types, with rapid location changes not being needed focusing more on the environment's complexity and the tactical challenge it presents.

3.6.4 Adventure

Adventure games focus on exploration, allowing a user to explore and discover an environment. Non-VR examples include The Elder Scrolls: Skyrim, Fable and Minecraft. VR adventure games offer several different types of experiences. In sandbox-style experiences, users are given a set of interactable elements in an environment and expected to interact with those items to amuse themselves. Other environments lean more closely into traditional game design casting the player in a role and providing a narrative and an objective for the player to engage with and complete. The key separator is that the player has the freedom to accomplish that goal with the tools given in any manner they see fit, with the user deriving enjoyment from the discovery process.

3.6.5 Horror

Like their movie counterparts, horror games provide suspense, anticipation, and fear to a user. Non-VR examples include Dead by Daylight, Amnesia: The Dark Descent and the Resident Evil series. VR horror games aim to take full advantage of VR systems' environmental immersion and adds new fronts for interaction. As the user now has a presence within the environment, the game designer can exploit this, invading the users' personal space to increase discomfort deepening the experience's intensity.

3.6.6 Fighting

Fighting games pit the user against enemies in martial combat, simulating melee fighting between two or more participants. Non-VR examples include the Super Smash Bros, WWE 2K and Street fighter series. Fighting games in VR are significantly different from those in traditional games. In traditional games, the user has a full view of the combat; In VR, fighting games restrict the players' view to a first-person perspective. This restriction results in attacks coming from opponents outside the users' field of view. Fighters are also assaulting the users' physical space increasing the immersion and intensity of the experience.

3.6.7 Racing

Racing games see the user piloting vehicles around courses commonly to achieve fast race times either alone or against opponents. Non-VR examples include Wreckfest, Crash Team Racing and the DIRT series. Racing games in VR are often simulation games, often positioning the user as the vehicle's pilot, seated in the cockpit in a first-person perspective of the vehicle.

3.6.8 First Person Shooters

First-person shooter games see the user utilising firearms to defeat groups of opponents. Non-VR examples include the DOOM, Call of Duty and Battlefield series. VR first-person shooter games generally are the most similar in terms of gameplay to their non-VR counterparts. Both types of game use a first-person perspective in both cases. The primary difference comes from aiming the weapon. In traditional games, the gun’s firing line is locked to a rigid axis, targeted at a crosshair on the screen. In VR environments, the user has complete freedom to aim the weapon in all directions but loses the artificial aiming of the rigid crosshairs, thus making aiming a more challenging and engaging task.

3.6.9 Puzzle

Puzzle games tax the mind and challenge the user's intellect. Non-VR examples include Portal, Tetris and Superhot. Puzzle games can often have a basis in board games; games like Klondike solitaire, chess, and checkers are digitised versions of their real-life counterparts. Games like bejewelled, slay the spire, and the Eternal Trading Card Game have their roots in board game design but exploit digital enhancements offered by the computer to increase the play's complexity and speed. VR puzzle games are generally not this kind of game, although a specialist game (https://www.tabletopsimulator.com/) does exist for playing traditional games in VR. Although some of these techniques are used in the puzzles within the games generally, the VR element would add very little to the game's enjoyment. VR games tend to be environmental puzzle games where the puzzle is part of the environment, requiring thought and trial and error to identify the correct solution.

3.6.10 Common VR Games Tasks

From the above review, a list of common VR tasks can be identified. These are typically performed using natural motion rather than simple button presses or mapped to inputs that simulate a natural motion, e.g., grip buttons.

The user can pick up, Manipulate, drop and throw objects within the environment. Usages for these systems are wide and varied, such as

- Carrying a weight to a switch.
- Moving a plank to build a bridge.
- Swing an axe at a tree.
- Discarding a no longer needed object.
- Pulling a lever
- Throwing a grenade.

3.6.11 Identified Task List:

Task	Rollercoaster	VR FPS
Head movement	Slow and steady, generally forward	Constant motion observing whole environment
Locomotion method	On rails motion	Default movement method
Physical movement	None	Rotating environment
Pick up	None	Grabbing weapons and tools
Point/interact	None	Menu selections, buying weapons.

Table 3-5:- Comparison of Interaction Methods Between Roller Coaster Rides and Pavlov

Table 3-5 shows a comparison of the interaction methods between VR rollercoaster rides and VR FPS games. This table shows a distinctive difference between the number of tasks the user must complete to interact with the environment successfully. Although differences will exist between VR FPS and other VR experiences, this model is a closer approximation of realistic VR usage. Most experiences involve interaction and locomotion and do not feature stimuli deliberately designed to induce nausea.

3.7 Experiment Environment

The experiment plan is to identify the impact of high cognitive load on cybersickness detection during VR usage. This goal will be achieved by developing three experimental environments, each having the ability to represent functionally identical tasks with different levels of cognitive load. Achieving this means creating tasks with changeable difficulty, with minimal impact on the user's actions to complete the task.

The tasks identified as illustrated in Table 3-6 are:

1. Remember number
2. Navigate Maze
3. Enter Number
4. Identify target locations

5. Shoot target
6. Remember list
7. Identify objects
8. Sort objects
9. Quiet Stand
10. TLX

	1	2	3	4	5	6	7	8	9	10
Visual Complexity				X	X	X	X	X		X
Temporal Complexity			X			X				
Spatial Complexity		X				X	X	X		
Physical		X			X			X	X	
Mental	X	X	X	X	X	X	X	X		X
Repetitiveness	X			X	X		X	X		
Short/long-term memory	X		X			X				
Task Flow/schema						X		X		

Table 3-6:- Comparison of Loads per Experiment Task.

3.7.1 Environment Design

3.7.1.1 Assessing Cognitive Load in Stimuli

As discussed in section 2.3, increasing the cognitive load in a task can occur in four ways: splitting attention, overstimulation, poor instruction, and inconsistent experiences. Ideally, our stimuli should isolate these factors to explore the individual effect on cybersickness. However, due to the nature of cognitive load, it is impossible to separate these issues entirely. Therefore, the experimental setup will isolate factors as far as possible and uses combined measures where this is not feasible.

Splitting attention comes in two forms, spatial and temporal. Spatial assessment will involve splitting critical pieces of information required to complete the task into different places in illogical association patterns requiring extra effort to collect and process. Temporal assessment will require the user to remember a piece of information for a long duration of time and recall it after completing several other tasks in between. While these additional tasks in between will like introduce load on their own, they will remain constant between experiments, thus inducing a consistent amount of load, leaving the temporal load as the only variation between the tasks.

Inducing overstimulation can be achieved by increasing the number of objects a user must engage with and/or the complexity of the tasks they must perform. Additionally, increasing the amount of information associated with the objects, such as introducing movement or variable options such as allowed and disallowed targets, will also increase stimulation.

The focus of this study is to evaluate the impact of task difficulty generated by the virtual environments. Stimuli design should focus on intrinsic methods of adding cognitive load, mainly split attention, and excessive information. Stimulations relating to inconsistencies could introduce additional sensory mismatch as objects do not behave as expected. Additionally, poor instruction quality can add additional extraneous load to a task. As experiment time is limited, this study will focus on the load introduced during tasks within the VR environment (memory, spatial and overload). These factors represent the most significant issues relating to comparing cybersickness studies. All instructions will be provided within the environment; Additional advice can be given verbally to confirm subjects understand how to complete the test before starting, removing issues related to instruction quality.

Implementing these factors into a testing platform will require care to ensure that all cybersickness-inducing factors are as similar as possible. Achieving this can be difficult in many cases. For example, adding movement to a target in an environment will induce unwanted additional vection into the environment. As the experiment plans to isolate cognitive load, it is essential to minimise other contributing factors. Where this is not possible, it is essential to acknowledge the potential influence of these factors and plan to identify other areas of the experiment which prove or disprove their influence.

3.7.1.2 Environment 1: - Maze Task, Memory Load

The maze task aims to provide a platform capable of altering the mental load within a task. While this also affects the temporal and spatial load, these are kept constant between iterations.

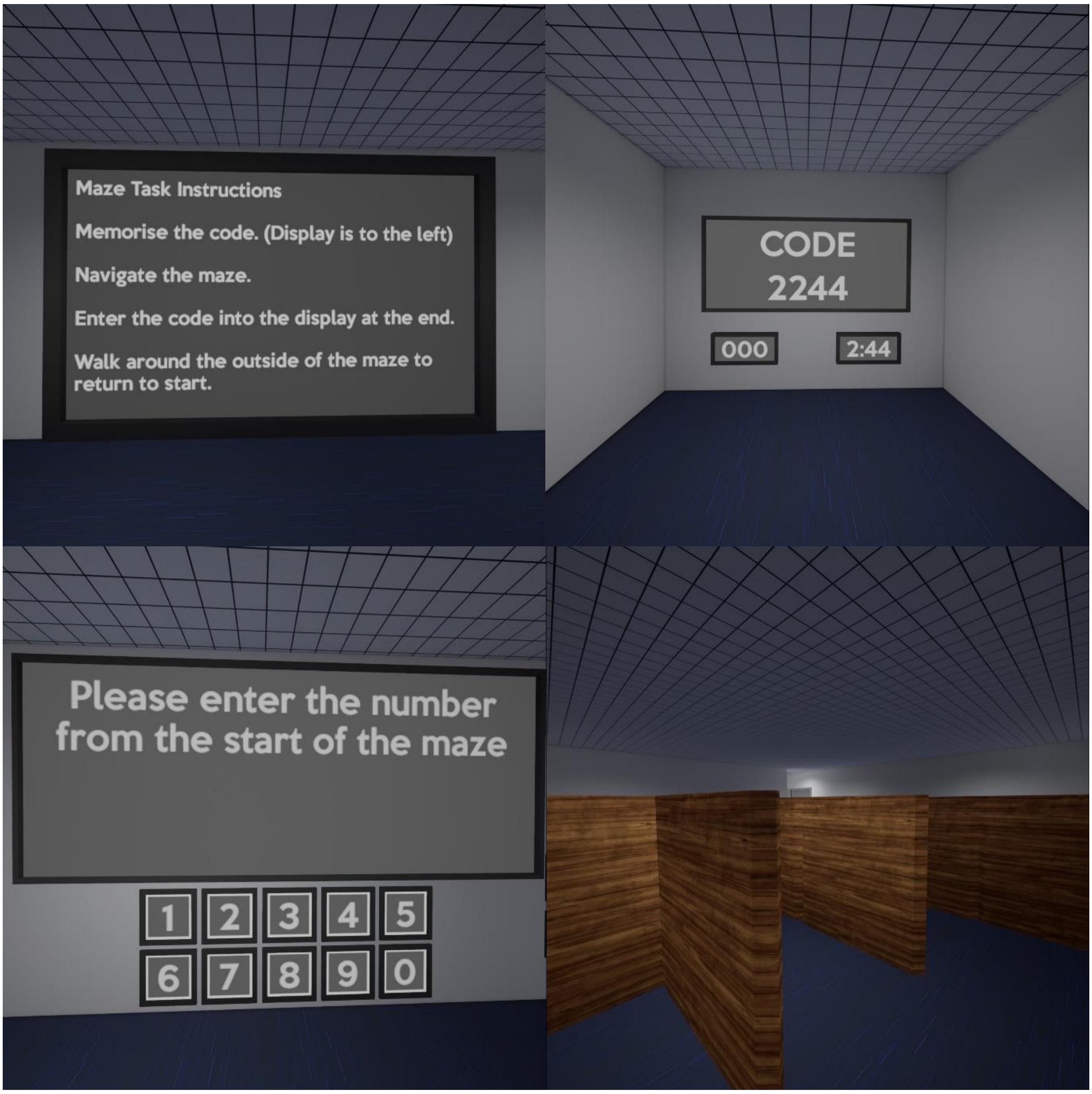


Figure 3-10:- The Maze Task. Showing the Instructions (top left), Memorised Code (top right), The Maze (bottom Left) and the Number Entry Console (bottom right).

The maze task presents the subject with a sequence of numbers to memorise. The subject must commit these numbers to memory and is allowed as long as they need to do this. Once memorised, the user then navigates a short 6x6 cell maze to the opposite corner. The maze acts as a simple spatial and temporal task, separating the code's acquisition and its usage, thus creating a mental load due to the separation.

At this point, the user inputs the number memorised into a console on the wall. A scoring system of eight points for a correct digit and minus four points for an incorrect one. After entering the number into the console, the subject must navigate the maze's outer edge and return to the task's starting point. The subject then repeats the task until the time limit expires.

The task also generates a unique path for each lap of the test and utilises a recursive backtracking algorithm to ensure a solution is always possible. This process is critical to ensure that a consistent maze does not introduce a different mental load when solving the maze between iterations. While this maze generation method guarantees that the maze is solvable, it introduces different mazes with different path lengths and direction changes for each solution. However, this will not impact results as subjects will take sub-optimal paths and make mistakes leading to a roughly equal time to solve in most cases. Adjustment of the task's memory load is achieved by increasing or decreasing the number of digits in the code. The high load task will use four digits; the low load task will use eight.

It is important to note that this task also impacts other areas of cognitive load and memory. Solving the maze introduces intrinsic load by introducing a problem-solving element and creating a mental map of the maze's layout. Randomizing the layout of the maze keeps this load uniform throughout the test. As such, the additional load introduced by this factor will be consistent in each test, requiring the same amount of mental effort each time.

3.7.1.3 Environment 2: - Shooting Task, Spatial Load

The shooting task aims to provide a platform capable of identifying the impact of spatial load within the test environment.

Figure 3-11:- The Shooting Task. Showing the Instructions (top left), Gallery (top right), Task A Test Scenario (bottom left) and Task B Test Scenario (bottom right).

The shooting task replicates a traditional shooting gallery environment found in fairgrounds and arcades around the world. Three rows of targets appear in the gallery, with targets spawning on the playfield's extreme left or right and scrolling to the opposite side of the playfield. Each row always moves in the same direction (the front and back rows scroll left to right, and the middle row scrolls right to left). The targets' speed was kept constant at 1 meter per second to keep vection contributions to a reasonable level. Targets appear in the range at different distances from the user (front row 2.5 meters, middle row 3.5 meters and back row 4.5 meters) and at different elevations (front row 1.5 meters, middle row 2 meters and back row 2.5 meters). Distances of the targets were determined based on the recommended viewing range for the headset.

A set of sample targets appear above the back row of shootable targets showing one of five coloured targets (Red, Green, Blue, Yellow or Pink). The subject must find a matching target in the rows of targets and shoot it, then a new sample target is created, and the process repeats. This process provides a change in spatial load between the two test iterations by having multiple positions for the sample target to be shown in the increased load scenario instead of a single position in the low load scenario. This reduction in the sample's position variation requires less visual search time, lowering the task's spatial demand similarly to that performed in other spatial search tasks (Longstaffe et al., 2014). Two points are awarded for each target successfully shot, and one point is deducted for each target incorrectly shot.

To remove unfamiliarity with pickup and shooting mechanics, a small training section for this task was provided. This process involves letting the user pick up the gun and take a few practice shots at the back wall before pressing the button to start the test. The shoot also contains a model not downloaded from an asset repository (the handgun model obtained from:- (https://free3d.com/3d-model/usp-45-44387.html)).

3.7.1.4 Environment 3: - Sorting Task, Overstimulation

The sorting task aims to provide a platform capable of identifying the impact of overstimulation within a task.

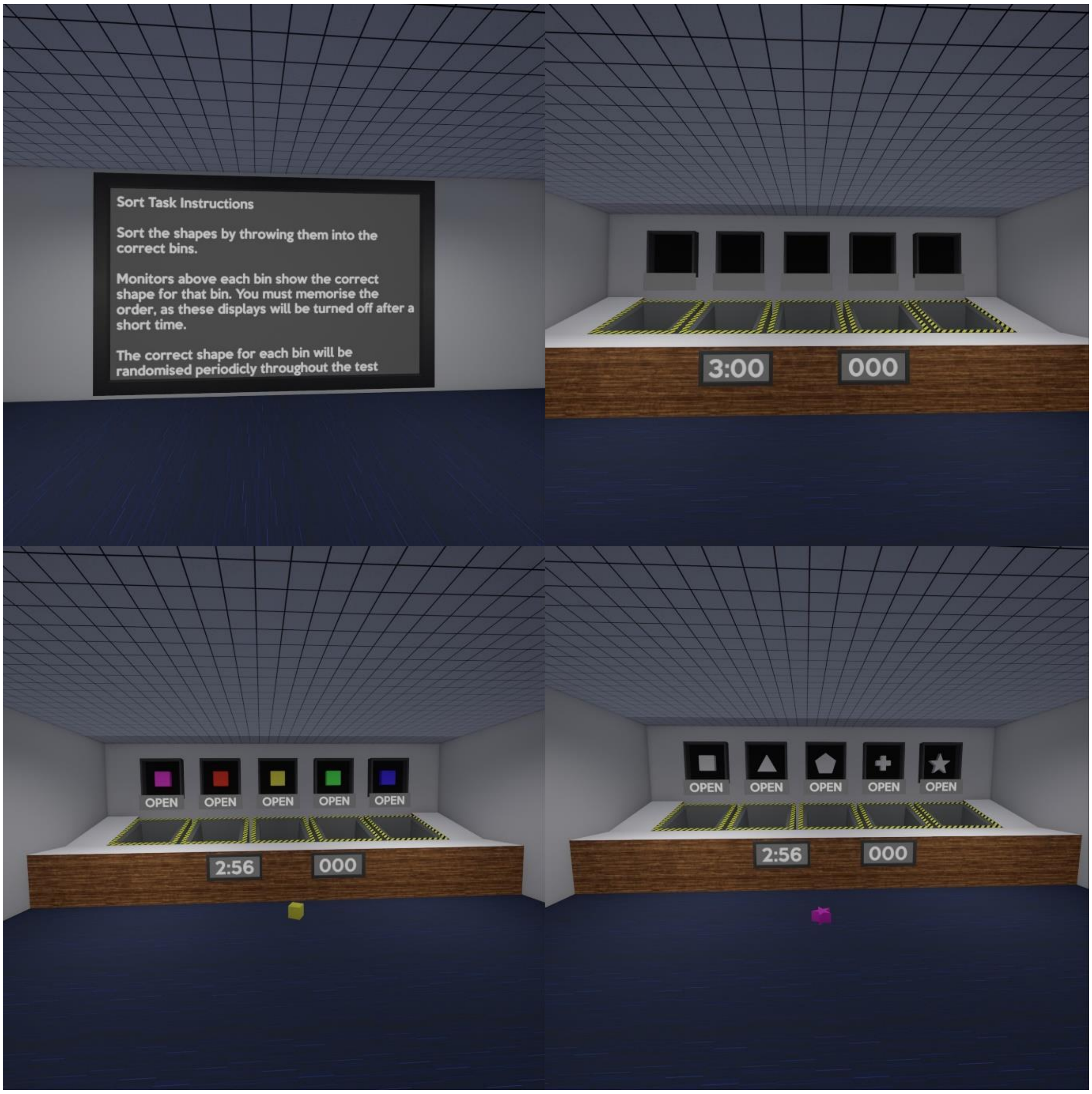


Figure 3-12:- The Sorting Task. Showing the Instructions (top left), Bins and Monitors (top right), Task A Test Scenario (bottom left) and Task B Test Scenario (bottom right).

The Sort task involves having the subject sort objects into bins. The environment contains five bins with sample displays above them designating which bin accepts which object. After a short delay, each display will no longer show its contents; therefore, the subject must memorise which bin is associated with each object. Various objects will then appear on the floor in front of the bins, and subjects must throw each object into its corresponding bin. Correctly sorting an object into its proper bin awards two points; sorting an object into the incorrect bin awards minus one point.

Two versions of the test are needed to test for the overload condition, a low load stimulation and an overload stimulation. Significant differences exist in the profile of each of these two tasks to achieve the two different scenarios. The low load simulation changes the samples every 60 seconds providing 40 seconds to memorise each sample before hiding the samples. The objects are all the same shape with different colours (Red, Green, Yellow, Blue and Pink). The spawning of new objects is set at a rate starting at one object per five seconds, rising linearly to 1 object per three seconds at the end of the test. The overload simulation significantly increases the task's complexity; the rate at which samples change is increased to once every 30 seconds, with only 10 seconds to remember the sample order for each sample set. Identifying different objects is done by shape (Circle, Triangle, Cross, Start and Pentagon). The spawn rate is also significantly increased, starting at one object per five seconds and rising linearly to one object per second at the end of the test.

This test aims to create a scenario by which it is impossible to keep up with the information presented to the subject, overwhelming them and identifying if this change in spatial and memory load impacts user stability and cybersickness. Training for this task involved having an object placed on the ground before entering the test room, which subjects could practice picking up and throwing before the test begins.

4 Experimental Methodology

4.1 Hypothesis

If the cognitive load a subject is under affects cyber sickness symptoms, and this link is bidirectional. It will become essential to establish the importance of using comparable stimuli rather than the myriad of stimuli present in most current research. Further, the impact of cognitive load may result in VR being inappropriate for long-term usage in complex scenarios. As evidence of cognitive load effects on cybersickness is limited, there are two major and one minor theory relating to this assessment with very different outcomes.

1. A user's cognitive load has no impact on cybersickness levels
   . A possible outcome that would mean integrating VR technologies into the workplace could proceed without concern.
2. Additional cognitive load reduces cybersickness.
   . This result is a possibility as the current theory that additional cognitive load in an environment leaves less working memory could result in the brain discarding conflicting information between posture and the visually perceived environment due to insufficient mental capacity to process the stimuli.
3. Additional cognitive load increases cybersickness symptoms
   . This result is a possibility, however minor, that the additional cognitive load from an environment results in the user having increased dysphoria from the conflicting stimuli as they have less working memory to support separating them.

From these we can produce the following hypotheses:

1. Cognitive load has a significant impact on cybersickness onset and severity during the usage of virtual reality head-mounted display environments.

Cognitive load's impact on cybersickness symptoms is not fully understood. As such, its contribution to cybersickness in experiments is currently unknown. As the variety of experiences used in cybersickness, studies vary massively, and differences exist in the cognitive load of even seemingly very similar experiences. Therefore, identifying its contribution to sickness symptoms will allow more insight into the variance of results when comparing studies.

This hypothesis is test by capturing subjects perceived mental loads through various tasks under high and low mental loads. Variations in the subjects SSQ-T scores will then correlate to the high and low load conditions. Differences will be observeable if the hypothesis is true.

2. A model representing actual VR usage severely impacts stability in subjects when recorded via COP pressure measurements or VR HMD positional data.

As the vast majority of cybersickness studies utilise engineered stimuli or heavily restrict user movement within the environment. Common VR usage is often unrestricted, allowing the user some free movement to move around the environment. For these approaches to be practical, they must be tolerant of the variance in actual VR usage.

By capturing positional data from the VR HMD, differing levels of subconcious motion may be observeable during the quiet stand measurement after performing different task with varying mental loads. These variances should not be directly linked to the SSQ-T score to establish proof that the varieance is not wholly attributed to the onset of cybersickness.

3. Cognitive load has a significant impact on the stability of subjects during VR usage sufficient to make stability measurements made via tracking of VR HMD position inappropriate as an indicator of cybersickness.

As the impact of cognitive load on cybersickness is indeterminate, an impact on subconscious stability may also exist. As such, positional data recorded from the HMD position may be influenced by this. If this is true, then an increase in head motion would be evident without an associated increase in sickness symptoms. If this is the case, this would suggest a need to consider the mental load of the experience while assessing instability as an indicator of cybersickness.

By varying the mental load placed on a subject during VR HMD usage, differences in the subconcious sway of subjects may be observeable. To prove this hypothesis the high load conditions should exhibit differing levels of sway in comparison to the low load conditions for identical tasks.

To validate the research questions we need to show the following links between our hypotheses as shown in table Table 4-1

Research Question	H.1	H.2	H.3
Q1	X		
Q2		X	X
Q3			X
Q4		X	X

Table 4-1:- Links showing the research questions which must be validated to answer each research question.

4.2 Approach to study

The study will have two test scenarios, each comprising 3 test tasks. Each test scenario will see subjects interact with the VR environment under differing levels of cognitive load. Test A will be the low load condition with the mental demands of each task set at a low level not to require significant effort for the participant to complete. Test B will be the high load condition. It will comprise the same tasks as the ones completed in Test A under a significantly higher cognitive load, increasing the mental demand on the subject.

Pre and post-test subjects were administered COP balance tests via the Wii balance board in line with the methods outlined in Weaver et al. (2017). This data will allow an analysis of pre and post-stability before and after each test, identifying if a difference in post-test stability exists between the two conditions. Identifying this will determine if a VR application's cognitive load significantly impacts user stability measurements. Additionally, subjects will complete an SSQ questionnaire (Kolasinski, 1995) post-test to identify the severity of cybersickness symptoms the subject is experiencing post-test.

Each test condition saw four quiet stand measurements (see section 4.5.1.2) made during the testing period, one at the start just after entering the VR environment and one more after each of the three tasks. (Maze, Shoot, and Sort tasks) each of these tasks placed a different mental load on the subject. Therefore, observing motion during the quiet stand periods after each task should give insight into how mental load influences subject instability in VR HMD environments. After completing each task, a survey captured NASA TLX results (see section 4.5.1.4) to validate the increase in task load and, by proxy, the cognitive load on the subjects during each test case. Section 3.7.1 (environmental design) provides descriptions of the tasks performed during each test.

4.3 Experiment design

4.3.1 Independent Measurements

Two independent measures existed in the study. Cognitive load is measured via two comparative test scenarios with differing loads, one high and one low. The impact of the type of cognitive load (memory, spatial and overload) is also assessed via the different tasks completed during each test. This combination of factors results in one test of each cognitive load condition for each type of cognitive load. Validation of differing cognitive loads is achieved via NASA TLX tests post each test scenario.

4.3.2 Dependent measurements

The position of the VR HMD is measured throughout the test to identify if any of the independent variables alter the subject's stability. Similarly, positional data was also recorded from Controllers and the centre of the back; however, the data produced was considered unreliable and discarded. COP measurements were gathered pre and post-test to identify any changes to instability using a validated methodology to confirm our findings' accuracy and identify any stability changes throughout testing. SSQ results were captured post-test to identify any onset of cybersickness symptoms brought on by either of the test scenarios. Various factors regarding the subject were also captured, such as gender at birth, identified gender and experience with VR technologies.

4.4 Equipment

4.4.1 VR Equipment

4.4.1.1 Data Collection

Headset positional data was acquired approximately 90 times per second via its reported position within the Unreal 4 game engine (https://www.unrealengine.com/en-US/). The data recorded included the relative world position (The real-world space position of objects) in X, Y and Z coordinates and the orientation of the HMD in roll, pitch, and yaw Euler angles. The capture of the in-game positions of the devices, such as the motion controllers, was also performed in the same manner.

The following objects were tracked and recorded.

- HMD, Recorded the position of the HMD in real space
- Left- and right-hand controllers, the position of the left- and right-hand controllers.
- Back Crown, A crown tracker placed in the small of the subject's back to provide an additional reference point for user stability.

The experiment incorporated a sensor worn on the back via a crown tracking device. In theory, the subject would wear the device like a belt around the chest, holding the tracking crown in the small of the back. In practicality, this setup proved incredibly inconsistent in its implementation. Getting a consistent position for the device proved extremely difficult, primarily due to the belt's elasticated nature holding the crown device. Firstly, the belt was insufficient to accommodate all subjects, mainly not contracting to a small enough size to allow position around the chest without slipping. In cases where the belt did fit the participant, slippage over time was evident during testing, and the orientation of the crown would change. Tracking issues were also prevalent due to the sensor's position commonly being occluded by the user's body from the light houses. While the crown's position generally remained stable during postural recordings, loss of tracking was very common. These issues made the results generated from this source unreliable and intermittent and unsuitable for further analysis.

The experiment environment captured data throughout the testing. However, this portion of the analysis only considers measurements captured during the quiet stand periods. This study considers three possible methods for detecting instability via HMD data. The first method involved calculating the path length to achieve a similar measurement to that of the balance board representing the total distance travelled by the user's head during the quiet stand period. The rotational distance travelled was also considered as a sum of all rotational movement recorded. The second method involved counting the number of times the subject's motion changed direction on each axis as an approximation of subconscious positional adjustment. The final method counted the number of times the motion changed from positive to negative or vice versa (referred to as zero crosses), representing a broad change from moving in one direction to another.

While analysing positional data, a significant number of results exhibited sudden massive spikes in movement, particularly at the start of data recording. These results are possibly explained by the user settling into the stance for recording. However, this issue should not have occurred as users had a significant period to settle into the position required for recording (10 seconds) before the recording started. Another explanation is a deliberate correction of stance being made by the subject, possibly due to discomfort or loss of balance. The possibility exists that subjects made a deliberate correction to their posture during the test, returning to a stable measurement in the new position after the correction. Finally, a loss of tracking during the recording process could also explain the errors in the data. While the examiner observed no tracking losses during the quiet stand period, there is a possibility it occurred.

In any scenario, these results are erroneous, but interestingly, they do accurately represent the kind of problems making these measurements will encounter in the real world. Furthermore, as data capture for this experiment was under laboratory conditions, this represents the best-case scenario that could exist in a real-world scenario, with ideal environmental conditions, hardware setup and an observer directly controlling the experiment and correcting any erroneous behaviour observed. None of these things will be true in a real-world scenario, and incidents of deliberate correction and tracking loss will likely increase. This issue shows the need for this method to accommodate these errors as the likelihood of achieving a perfect recording is unlikely. Another interesting observation seems to be that individuals who exhibit these erroneous measurements in earlier recordings also exhibit them in later recordings and alternate tests. This finding suggests the erroneous results recorded may themselves be an indicator of instability if they correlate with an associated increase in SSQ-T scores.

In either case, the removal of erroneous results needs to occur before results can be effectively processed. To achieve this, any result generating a moment-to-moment directional magnitude on a single axis of greater than 0.1cm per frame (90th of a second). This value was selected as it was approximately four times the standard deviation of the data set, making it incredibly unlikely this filter would remove any valid results. Upon detecting a possible error, half a second worth of data on either side of the result was removed to remove any possible influence on the results.

4.4.2 Balance Board (pre and post-stability measurements)

The participant's postural stability was measured pre and post-test using the Wii balance board. The measurement is the total distance travelled by the subject's centre of pressure throughout the test. The centre of pressure represents the direction vector generated by the sum of all forces acting on a supporting structure. The subject makes subconscious changes in posture; the force on the balance board shifts and changes in this distribution of force can be measured and used to determine the amount of subconscious motion.

The Wii Balance Board uses four load sensors to record these changes. As the users' weight shifts around, each sensor records a different amount of load. Widdowson et al. (2019) provides an algorithm for calculating the centre of pressure measurement for a given frame from these measurements. By measuring the distance from one moment to the next, we can establish the total distance travelled by the centre of pressure throughout the experiment. Recordings were 30 seconds long with a sample resolution of 40 samples per second.

4.4.3 Experiment Environment

The physical environment used for the experiment is the universities virtual reality development pods. These pods can run room-scale VR applications, consisting of a flat lino floor of 4 meters by 3 meters. Each pod contains a state-of-the-art desktop PC designed to run demanding VR applications (see Table 4-2). Pods also contain a cable suspension system designed to keep the cable that attaches the headset to the PC off the floor and above the user, reducing the trip hazard possibility and providing minimal interference with the user. The pods have three solid walls and a curtain acting as the fourth wall. This setup allows for privacy while using VR while preventing accidents caused by users wandering out of the pod while using VR (or the occasional absent-minded students wandering in). Finally, each pod is equipped with two displays to allow external observers to see what the user is doing within the VR environment.

Feature	Specification
Processor	Intel Core 19-9900 3.10GHz
RAM	32 GB
OS	Windows 10 Enterprise
Graphics	GeForce 2080 8GB

Table 4-2:- PC Specification for Experiment.

4.5 Procedure

- Briefing

  Subjects initially are briefed on the nature of the experiment (see Appendix B: - Briefing), identifying aspects of risk during the experiment, including the likelihood of experiencing nausea, headache, eye strain, dizziness and in sporadic cases vomiting. Procedures relating to mitigating risks from COVID-19 within the environment, including the need to wear a mask during the experiment and the sanitation procedures undertaken, were communicated to the subject. Subjects were informed of their right to withdraw and the procedure for withdrawing from the study. Subjects signed a consent form (see Appendix B: - Consent form) before commencing the study.

- Screening

  Subjects had their visual acuity verified via a Snellen eye test (Snellen, 1862). Additionally, subjects were asked about other visual problems they may have that may affect the test (such as colour-blindness and astigmatism). Any subject with visual problems was removed from testing. Subjects will also perform the MSSQ to identify and remove any individuals with severe susceptibility to motion sickness. Finally, the subject's interpupillary distance was measured, and the headset was calibrated to this distance.

- Baseline stability

  Wii balance board captures a baseline stability measurement for the subject. The process for this is outlined in section 4.5.1.1

- Task explanation

  Subjects were given a short presentation on how to use the environment. This presentation demonstrated how to pick up and drop items within the environment, locomotion via the slide locomotion method, rotation via snap turning or physical turning of the subject's body. An opportunity to ask questions was provided.

- Experiment

  Test tasks proceeded in the following order:

  - Quiet Stand
  - Maze Task -> Quiet Stand -> TLX Data
  - Shoot Task -> Quiet Stand -> TLX Data
  - Sort Task -> Quiet Stand -> TLX Data

  Instructions are provided for each test by a panel on the wall detailing the instructions for the test. Instructions given to a subject were identical to avoid any influence on the task's extraneous load; additionally, subjects could ask as many questions as they liked about what to do during the test before starting.

  The experiment order did not change and was the same for both load conditions. Tasks lasted for 3 minutes, and quiet stand recordings took 30 seconds to complete. Two runs were completed by each subject separated by at least 24 hours to allow residual effects from the first experiment to subside. The order of high and low load tests was alternated for each subject to remove the effect of order from the experiment.

- Post-test stability

  Wii balance board captures a baseline stability measurement for the subject in a similar way to the pre-test stability measurement.

- Debriefing

After the second test, the collection of the TLX pairwise comparison will complete the data set. After this, subjects debriefing will occur to inform the subject of the experiment's nature (See Appendix B:- Debriefing).

4.5.1.1 Pre / Post-test Postural Stability Measurement

Before and after testing, users will have their baseline stability measured using a Nintendo Wii (Nintendo, 2019) balance board in line with the method from Widdowson et al. (2019). Widdowson et al. (2019) performed the test while the subject is within the VR HMD environment, they utilised an entirely passive test, observing an environment and not requiring any interactions. As this test requires significant interaction and locomotion, users' likelihood of adjusting their stance during the testing process is significant. This possibility of motion creates a significant hazard to the user deliberately or accidentally stepping off the board and falling. Introducing the balance board to the VR environment during the quiet stand periods was considered. However, no safe method of performing this task could be derived, which did not require removing the subject from the VR environment or close contact with the subject, which is currently unacceptable due to COVID concerns. Therefore, balance measurements were made pre- and post-test as to establish a baseline measurement for instability outside of the VR environment. This process would allow for any significant differences in postural instability at the start and end of the experiment to be detected.

Subjects stood on the balance board in a quiet stance (feet slightly apart, legs and back straight and looking straight ahead). Once the subject is comfortable, they verbally confirm their readiness to the instructor who starts the recording. The recording period lasts for 30 seconds. This process is repeated twice, once with the user’s eyes open and again with the user’s eyes closed.

4.5.1.2 VR Stability Data, Quiet Stand

Figure 4-1:- The Quiet Stand Recording Environment

Effectively tracking user stability throughout the test requires multiple quiet standing periods, during which time the subject's head position will be monitored and recorded. Collecting this data involves having the subject enter a virtual room and navigate to a central platform. Once on the platform, the user will be given 10 seconds to ready themselves for the quiet stand measurement, adopting the same posture as used in the balance board test, standing up straight and looking straight ahead for the duration. A 30-second recording is then made of the headset position and rotation to monitor user movement within the environment during this process.

Slight variations in this movement should be detectable (van der Veen et al., 2019) and provide a metric of user instability. Four recording periods will occur, One upon entry to the VR environment to serve as a baseline for VR instability, and three more one measurement made after completing each task. Comparing these results to the SSQ data collected at the end of each test will determine whether instability detected during these quiet stand periods correlates to the intensity of sickness symptoms.

Data values for the headset’s position and orientation are captured through the Unreal 4 engine (https://www.unrealengine.com/en-US/) at a rate of 90 samples per second, recording the reported position of the headset and the game world representation of the headset. Capturing the positional data was challenging. This challenge came from the fact that occasionally significant differences in the headset's rotation would occur alongside a change in the playfield’s axis orientation. The x-axis would become y and vice versa. Manipulating this data to reorientate the headset back to the same axis showed no difference in the position recorded. Re-orientating the data to accommodate was not problematic as the application already performs this task to align the users' viewpoint. Observing the player's calculated headset position showed a margin of error of no greater than 0.0001 cm between the reported position and the headset’s actual position. As such, this measurement was deemed acceptable for gathering data. Similar tests revealed a similar level of accuracy in rotational data.

4.5.1.3 Alternative recording of quiet stand equivalent

The usefulness and practicality of applying the quiet stand measurement during actual VR usage is a significant question. The practicality of having the software user stand stationary for 30 seconds, looking straight forward in the application is problematic. Firstly, without proper instruction and outside of a laboratory scenario, it is doubtful a user will adopt the proper stance needed for the recording, generally choosing to be comfortable rather than accurate.

Secondly, having a user forced to stand still and record posture data for 30 seconds is frustrating and impractical. Taking the user out of the virtual experience for 30 seconds to see if they are sick or not breaks immersion and is frustrating if performed too frequently, leading to non-engagement in the process. It may be possible to integrate these tests into the environment's workflow, particularly in games where a contrived reason could be derived to make the player standstill for an extended period, such as riding an elevator in an enclosed cab. Alternatively, the recording of stability data could occur during loading times. However, this approach's effectiveness may be questionable as, during loading periods, the computer is under high load, and the effectiveness of the tracking systems used by the headset may be compromised. The real problem is forcing the user to stand still to make the recording. Downtime in the environment is an opportunity for subjects to stretch or adjust the headset for comfort. Finally, unless measurements are being made frequently to the point of disruption, capturing the moment a subject starts exhibiting symptoms before the point where they remove themselves from the scenario is improbable.

What is needed is a method of monitoring the user while performing their normal activities. However, capturing this data from generic motion during VR usage would likely be tricky due to the unpredictable nature of motion during VR usage. However, if a period of activity could be identified where the head position is likely stable during a VR activity, then obtaining a suitable reading may be possible. Engineering a scenario for this stable period may be possible. By identifying tasks in which the user keeps their head position constant and rotated in a particular direction.

The current experiment environment has two possible scenarios where this may be possible. Firstly, a scenario where a user walks down a corridor between tasks may provide an opportunity as the user's head position and viewing angle is likely to be stable and in a roughly constant direction. Two such scenarios exist in the testing environment, both occurring in the maze testing scenario. Two short corridors take the user back to the test's starting point, taking a few seconds to navigate. One has no distractions; the other sees the user reading a number from the test scenario while walking down the corridor. The second scenario is during the shooting task when the subject aims the gun at the target, which will provide a frequent stable head position but for a much shorter duration and with an inconsistent viewing angle. The sorting test can be utilised as a control, collecting measurements whenever the filter allows it.

The other issue relates to the duration required to achieve a reasonable measurement compared to the baseline 30-second quiet stand. The amount of data collectable will be significantly shorter than that of the quiet stand periods, and these short durations may prove inadequate. The data collection process was identical in manner and frequency to the quiet stand measurements. The primary difference was that data capture occurred during the testing period instead of the quiet stand period. As the test wanted to record natural interaction methods, these passive recording should give a clear indication as to whether this approach is feasible or not.

4.5.1.4 Nasa TLX Data

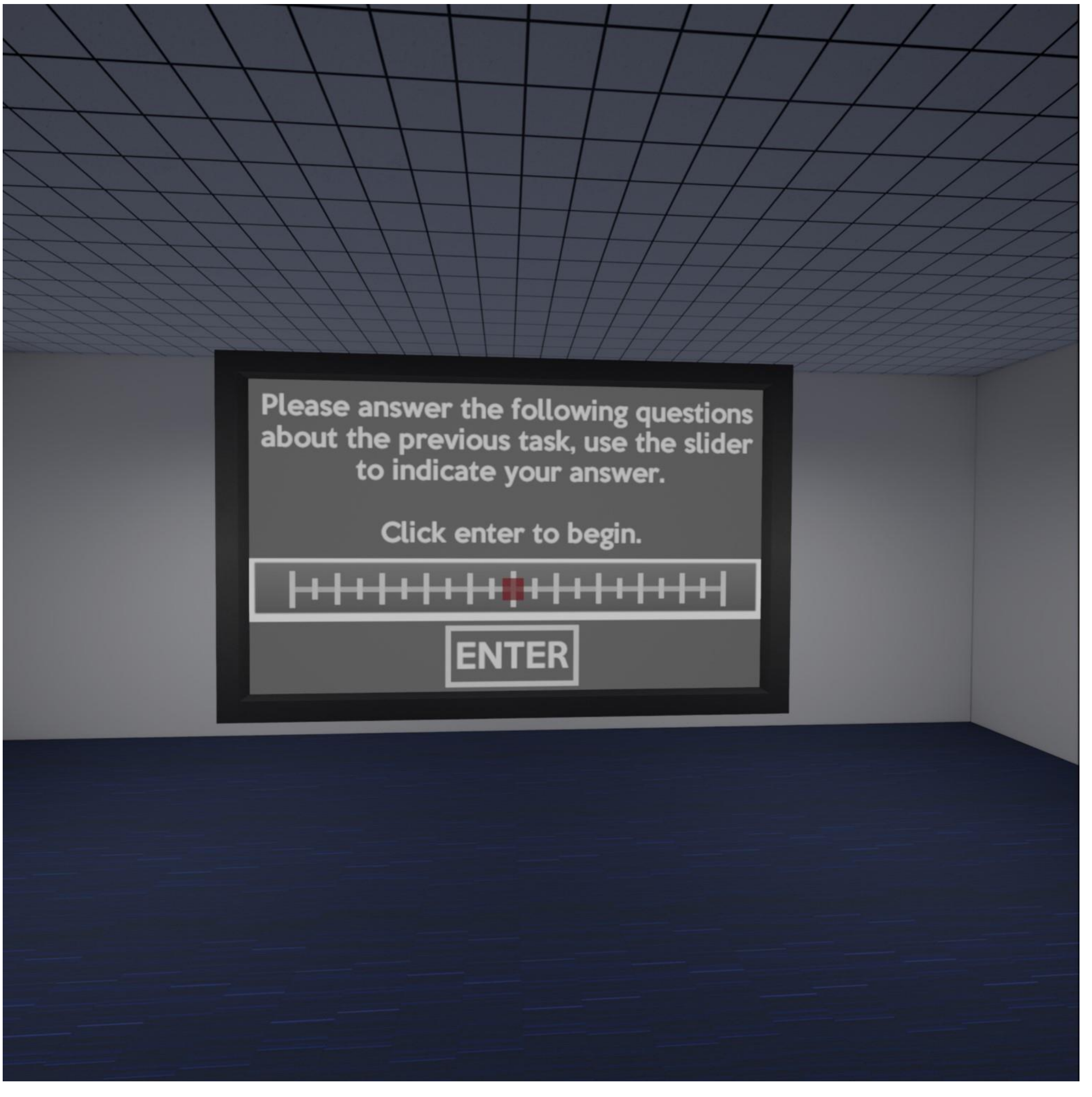


Figure 4-2:- NASA TLX Data Capture Section.

The NASA Task Load Index (NASA TLX) (Hart, 1986) is an assessment tool widely used to assess perceived workload. The NASA TLX survey has six categories (Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration). Each indicates how each of the tasks in our test affects the subjects. From this data, we can correlate TLX scores against instability and SSQ scores to identify if any correlation exists between the results identifying a link between task complexity and cybersickness and instability measured via the HMD.

The TLX test involves users providing a score for each category on a scale between 0 and 100 in steps of 5, giving the RAW TLX Scores for each task. Post-test, subjects perform a series of pairwise comparisons between the categories to identify the factors with the most significant contribution to the task's load. Capturing the TLX data within the VR environment is performed by having subjects enter their answers into a console after each test scenario. Subjects use a slider on the console to indicate their responses and a short description of each category's meaning. A paper-based test captures post-test comparison data.

4.5.1.5 Simulator Sickness Questionnaire (SSQ)

The Simulator Sickness Questionnaire (Kennedy and Lane, 1993) is an assessment tool used to assess the severity of cybersickness symptoms experienced by a subject. The SSQ test involves having the subject complete a questionnaire rating each of 16 symptoms into one of 4 categories (None, Slight, Moderate or Severe). Each of these ratings is assigned a score based upon the severity indicated in the response. By applying a weighting to each response, four scores SSQ scores can be derived (SSQ-O, SSQ-D, SSQ-N and SSQ-T), indicating the severity of each of the components of cybersickness a subject is experiencing. In this experiment, administration of the SSQ takes place post-test. Generally, SSQ results are collected pre-test for consistency and proof that subjects did not enter the environment with significant sickness symptoms.

4.6 Testing

All participants ran the test scenarios Test A, and B. Participants were exposed to different tests first based on an alternating pattern to observe the effect of test order, moving from a high load environment to a low load environment and visa versa. Each participant completed the test environment twice, with the second test condition being the alternative to their initial environment. At least 24 hours had elapsed between tests to allow for any residual sickness effects to dissipate. Experiments were conducted in the University of Derby XR development lab under the supervision of the principal investigator. Participants were screened for visual acuity based on the Snellen C test and for health conditions which may affect the test via verbal inquiry. Each participant was then given a short presentation explaining how to operate the VR environment and what to do should they wish to discontinue the experiment.

Participants performed the COP balance tests at the start and end of each testing phase to provide baseline stability measurements. For each test, the subject then completed the three activities in the test condition (maze shoot and sort). Four quiet stand measurements were taken during this period, once at the start and once again after each test condition. After each of the quiet stand measurements after each test, subjects filled in the NASA-TLX questionnaire to score the task difficulty of the task they had just completed. Each test run took approximately 20 minutes and was performed for each test conditions A and B.

Due to an oversight in the procedure of the test, no collection of pre-test SSQ data occurred. This oversight has two potential impacts on the study. First, as the SSQ is a subjective measurement, the results recorded here will not be comparable to those of studies that implemented both the pre and post-test SSQ. Young et al. (2006) identifies the likely effect on the recorded result will be a reduction in magnitude to our SSQ results. The second effect is the inability to remove the possibility any subject was already sick when they commenced the test. The likelihood of sick subjects entering the testing pool with any significant magnitude is minimal, partly due to the screening process, but more likely due to the increased need to isolate from unwell individuals due to COVID 19. However, the possibility cannot be ruled out entirely, and therefore both issues presented here should be considered during analysis.

4.7 Study Participants

The accuracy of the participant's vision was screened using the Snellen C test (Snellen, 1862). Subjects were allowed to participate with medium to correct vision problems such as contact lenses, glasses etc.). Additionally, subjects were screened using the MSSQ to identify any individuals with severe motion sickness susceptibility and to remove them from the study.

Thirty (30) Subjects agreed to participate in the experiment. Of these, six (6) subjects withdrew from the test, one due to eyesight problems, two due to potential COVID-19 Infections and three subjects withdrew due to extreme sickness responses. Of the remaining twenty-four subjects, nineteen (19) were male, and five (5) were female. The median age was twenty-two (22) years old (Min = 18, Max = 35) and primarily consisted of students from the University of Derby.

As part of the experiment, subjects were asked to indicate how much experience they had using VR technologies using one of four categories. Two categories defined subject experience with VR, subjects who described their VR experience as None or Once formed the low experience group. Subjects who described their experience as Some or Often formed the high experience group. This process placed ten subjects in the low experience group and fourteen subjects in the high experience group

4.8 Data Analysis

During the quiet stand period, the position and orientation of the headset were captured. Then post-test, the total motion was summed and divided by the recording length to give a measurement of motion per second. Analysing this data will determine if any difference exists in the number of head position corrections between sick and well individuals. Initially, this data was incredibly noisy; thus, applying a rolling average of 16 frames (0.18 seconds) to all data before analysis was necessary to obtain a reasonable view of the data without losing too much detail. Erroneous data was removed in line with the process identified in section 4.4.1.1.

Looking solely at raw output data may not be the only option for analysing the results to detect cybersickness. The process of instability involves a series of subconscious motions made by the subject, micro corrections to posture and head position, designed to maintain the desired body position within the environment. If these corrections' frequency could be detected and measured, it may make detecting the onset of cybersickness symptoms a possibility.

This report presents two possible models for detecting subconscious changes, zero crosses and velocity changes. Zero crosses refer to the point at which motion on an axis goes from positive to negative or vice-versa, changing the direction of motion on that axis. Velocity changes refer to the point at which the direction of velocity changes at the peak of motion representing the subconscious point of motion correction. Zero crosses will give a sense of the number of major directional corrections made by the subject. In contrast, the velocity changes will give a sense of certainty in those changes. The smoothing of data and removal of erroneous results were performed identically to the previous test.

Post each test activity (maze, shoot and sort) NASA TLX data was captured to quantify the task load the subject was under during the test. This data was used to validate the difference in perceived load between the two test scenarios, specifically looking at the mental load factor to validate the increase in cognitive load required between the two environments. Additionally, the TLX data was used to identify if an increase in raw load influenced stability measurements.

4.9 Statistical Analysis

The use of statistical methods to prove the reliability of data is common in evaluating cybersickness sickness symptoms. Common approaches involve the use of an ANOVA to determine if significant differences exist between groups of data. Unfortunately, the data presented in this study does not conform to a normal distribution of results, instead showing a heavy left skew, peaking at an SSQ-T value of zero. ANOVA analysis assumes a normal distribution, and as such, using it would result in misleading results. T-tests are also unsuitable for the same reason. Therefore, the bulk of statistical analysis in this report utilises the Mann-Whitney U Test to prove or disprove the even distribution of values within the data. This method is tolerant of the lack of standard distribution in the data.

There are several overlapping components to consider within these results, it is entirely plausible that a combination of the investigated factors may influence SSQ-T scores and stability measurements. Therefore, there is a need to conduct parametric tests on the results to ensure that the combined influence of factors is not directly attributed to the individual factors. This was achieved using ARTool (Wobbrock et al., 2011) to compare: test order, sick/well and VR experience factors against the SSQ-T and stability results. These parametric tests revealed no combination of any of these factors was significant in any of the tests. Results for this are included in Appendix F for completeness.

5 Results

5.1 Summarised Results/Statement

5.1.1 SSQ

Task A (Low cognitive load) returned a mean SSQ-T score of 14.65 (STD DEV = 19.79) and Task B (High cognitive load) returned a mean SSQ-T score of 19.79 (STD DEV = 27.30). This result suggests that the sickness symptoms caused by each test are mostly similar, with a slight increase observable in Task B. However, the standard deviation shows a wide distribution of values.

The frequency distribution of SSQ-T scores skews very heavily toward zero (Test A: Skew = 2.17, Kurtosis = 5.02 and Test B: Skew 2.19, Kurtosis = 4.81), suggesting most subjects did not report any symptoms beyond mild discomfort (SSQ-T <= 20). Those who did report more severe symptoms were not exclusive to Task A or Task B. Performing a 2-sided Mann-Whitney U test revealed that the means for Task A and Task B are not statistically different (U = 266.5, P= 0.66). This result suggests that neither test generates a significantly higher SSQ-T score than the other, inferring that an increase in task difficulty and by proxy (cognitive load) does not significantly increase this sample's sickness symptoms.

When comparing the effect of test order, ignoring the high and low load conditions, the results still show a heavy positive skew in both scenarios (1st runs Skew = 2.16, Kurtosis = 4.25 2nd runs Skew = 1.80, Kurtosis = 2.60). The first runs generated a mean SSQ-T score of 21.66 (STD DEV = 27.09), and the second runs generated a mean SSQ-T score of 12.78 (STD DEV = 17.87). This result shows a significant difference between the two runs, with the first test run commonly generating higher sickness levels than the second run. A Mann-Whitney test of the frequency distributions (See Figure 5-1) of SSQ-T scores (U = 373.5, P = 0.03) confirms this finding.

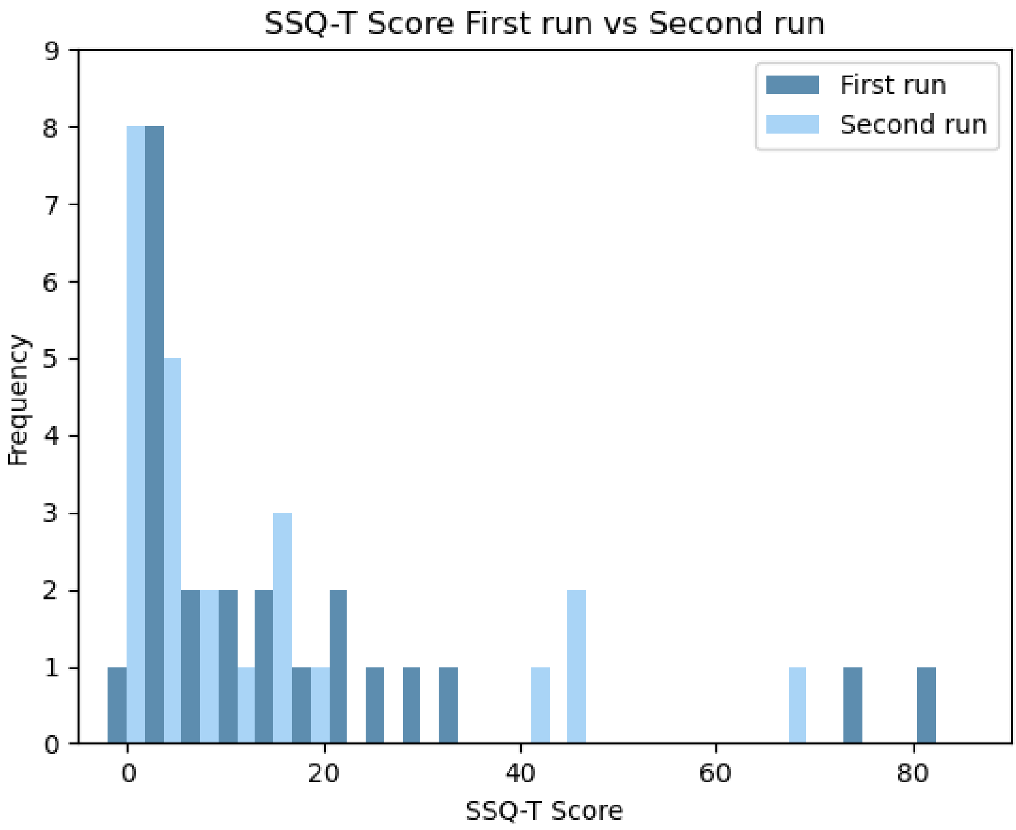


Figure 5-1:- SSQ-T Score First Run vs Second Run

Finally when comparing user experience (Low experienced users, Skew = 1.38, Kurtosis = 1.03, High Experienced Users, Skew = 1.54, Kurtosis = 2.23). Highly experienced users generated a mean SSQ-T score of 10.02 (STD DEV = 11.22), and low experienced users generated a mean SSQ-T score of 27.30 (STD DEV = 31.85). This comparison showed a major difference in average SSQ-T scores, with experienced users exhibiting lower sickness scores on average. Figure 5-2 shows that the more severe sickness scores seem to come from inexperienced users.

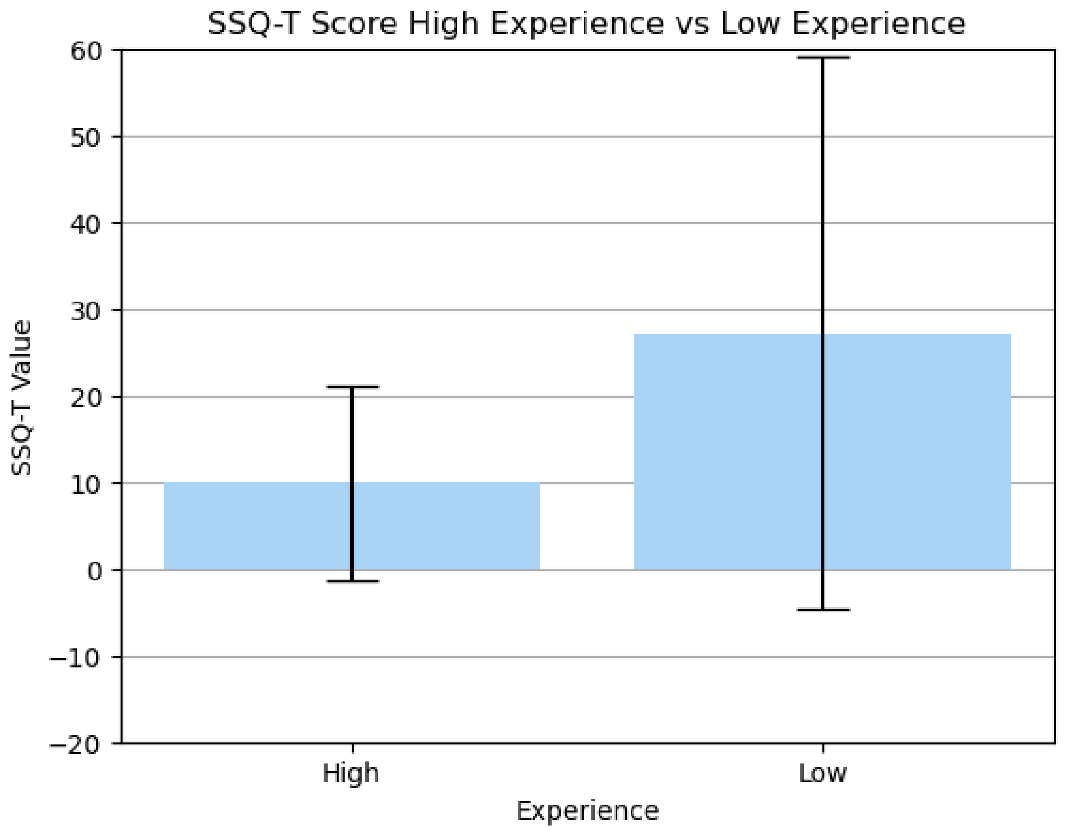


Figure 5-2:- Mean SSQ Scores, Comparing Users with High Levels of VR experience Vs Users with Low Levels of Experience

5.1.2 TLX

Multiple TLX data samples were collected throughout the experiment. The following tables document the RAW TLX data collected by the study in 6 categories (physical demand, mental demand, temporal demand, effort, performance and frustration along with the total scores. A higher score represents an increase in perceived task load for the category.

5.1.2.1 Task A Maze Task TLX Scores

Sample Number	Physical Demand	Mental Demand	Temporal Demand	Effort	Performance	Frustration	Total
P 1	0	10	0	0	0	0	10
P 2	25	40	30	35	40	15	185
P 3	5	15	25	50	0	10	105
P 4	25	35	15	15	10	10	110
P 5	5	5	50	5	10	0	75
P 6	5	5	100	5	25	0	140
P 7	10	60	50	70	20	10	220
P 9	5	5	20	20	5	0	55
P 10	20	60	60	50	20	20	230
P 12	60	55	60	35	30	50	290
P 13	30	20	50	10	40	5	155
P 14	30	35	60	40	30	30	225
P 15	0	10	75	10	0	15	110
P 16	10	30	30	65	30	15	180
P 17	10	25	70	70	25	25	225
P 18	20	15	75	25	15	0	150
P 20	0	10	65	20	0	10	105
P 21	5	20	35	20	40	0	120
P 22	40	25	20	20	20	20	145
P 25	10	15	75	65	30	15	210
P 26	10	35	60	25	20	20	170
P 27	10	30	60	65	5	15	185
P 28	25	10	70	30	35	5	175
P 29	25	25	60	55	50	10	225

5.1.2.2 Task B Maze Task TLX Scores

Sample Number	Physical Demand	Mental Demand	Temporal Demand	Effort	Performance	Frustration	Total
P 1	0	20	0	10	10	0	40
P 2	60	40	75	65	65	60	365
P 3	10	75	75	70	30	10	270
P 4	15	80	80	70	75	60	380
P 5	5	10	25	25	5	0	70
P 6	10	70	100	50	25	0	255
P 7	10	60	50	70	40	50	280
P 9	0	20	15	10	10	5	60
P 10	20	30	60	50	25	25	210
P 12	35	70	70	65	60	65	365
P 13	30	40	40	30	70	5	215
P 14	10	25	60	25	30	20	170
P 15	0	15	30	50	0	0	95
P 16	20	25	65	40	60	20	230
P 17	15	70	65	25	20	20	215
P 18	20	35	75	70	20	0	220
P 20	35	15	70	30	25	5	180
P 21	5	35	30	55	80	10	215
P 22	20	40	20	30	10	20	140
P 25	15	80	30	70	60	15	270
P 26	0	65	75	65	30	15	250
P 27	55	65	60	55	0	55	290
P 28	0	20	65	35	25	5	150
P 29	20	35	60	60	60	60	295

5.1.2.3 Task A Shoot TLX Scores

Sample Number	Physical Demand	Mental Demand	Temporal Demand	Effort	Performance	Frustration	Total
P 1	20	20	30	20	20	0	110
P 2	55	5	5	60	25	5	155
P 3	70	75	70	80	65	20	380
P 4	75	25	35	80	15	30	260
P 5	70	20	90	75	0	50	305
P 6	25	5	90	10	20	5	155
P 7	15	70	5	30	30	50	200
P 9	40	20	65	35	20	5	185
P 10	50	75	80	65	35	60	365
P 12	65	50	55	60	35	35	300
P 13	40	40	30	50	30	20	210
P 14	60	65	80	65	35	45	350
P 15	70	10	80	80	20	60	320
P 16	20	65	30	70	40	15	240
P 17	70	35	75	75	85	75	415
P 18	65	65	80	80	35	10	335
P 20	75	70	70	70	30	10	325
P 21	25	35	70	35	20	5	190
P 22	70	30	50	60	20	10	240
P 25	80	15	5	35	15	0	150
P 26	90	70	60	75	35	65	395
P 27	65	65	70	70	10	55	335
P 28	65	35	75	20	35	5	235
P 29	55	60	30	60	40	60	305

5.1.2.4 Task B Shoot TLX Scores

Sample Number	Physical Demand	Mental Demand	Temporal Demand	Effort	Performance	Frustration	Total
P 1	20	10	20	5	5	0	60
P 2	35	65	20	30	30	5	185
P 3	20	25	25	70	70	5	215
P 4	65	30	25	70	10	10	210
P 5	60	30	80	75	5	40	290
P 6	80	20	100	5	20	0	225
P 7	35	65	60	10	10	10	190
P 9	30	10	40	25	10	5	120
P 10	35	75	70	65	40	60	345
P 12	70	40	40	70	55	20	295
P 13	65	30	35	60	25	5	220
P 14	40	60	80	70	35	55	340
P 15	50	50	75	70	10	55	310
P 16	10	15	60	60	35	20	200
P 17	65	20	75	70	35	20	285
P 18	65	70	90	65	25	10	325
P 20	90	90	95	80	15	20	390
P 21	25	20	60	65	20	5	195
P 22	70	20	50	65	35	20	260
P 25	65	10	20	80	25	10	210
P 26	85	65	30	50	20	60	310
P 27	70	65	65	65	20	50	335
P 28	45	30	75	55	35	25	265
P 29	35	35	20	60	30	30	210

5.1.2.5 Task A Sorting TLX Scores

Sample Number	Physical Demand	Mental Demand	Temporal Demand	Effort	Performance	Frustration	Total
P 1	15	40	0	20	5	0	80
P 2	65	35	15	80	30	25	250
P 3	85	90	80	80	25	75	435
P 4	70	65	85	70	35	60	385
P 5	20	40	75	65	20	50	270
P 6	0	25	100	30	35	10	200
P 7	60	80	70	65	70	70	415
P 9	35	30	65	60	35	30	255
P 10	85	80	70	80	25	70	410
P 12	60	65	70	60	30	45	330
P 13	80	80	65	75	70	60	430
P 14	60	75	55	65	60	55	370
P 15	100	80	90	100	35	75	480
P 16	60	45	40	70	45	20	280
P 17	10	80	65	70	40	30	295
P 18	20	70	85	75	70	70	390
P 20	70	80	80	80	25	35	370
P 21	45	70	85	65	65	10	340
P 22	50	85	75	65	70	15	360
P 25	65	80	65	80	30	60	380
P 26	65	85	65	80	65	70	430
P 27	80	80	75	65	30	60	390
P 28	60	75	75	60	25	5	300
P 29	70	75	70	80	40	30	365

5.1.2.6 Task B Sorting TLX Scores

Sample Number	Physical Demand	Mental Demand	Temporal Demand	Effort	Performance	Frustration	Total
P 1	10	70	80	80	80	15	335
P 2	60	75	90	70	45	65	405
P 3	65	95	100	50	85	30	425
P 4	75	70	80	75	70	85	455
P 5	50	100	100	100	50	75	475
P 6	5	75	100	65	90	25	360
P 7	80	85	85	75	80	85	490
P 9	30	70	80	75	45	40	340
P 10	70	80	80	80	85	70	465
P 12	80	90	100	90	100	95	555
P 13	70	60	80	70	75	30	385
P 14	65	80	85	70	65	75	440
P 15	75	95	100	85	30	60	445
P 16	60	80	70	70	65	20	365
P 17	25	80	80	80	90	85	440
P 18	10	80	75	65	75	60	365
P 20	70	90	85	70	60	65	440
P 21	55	75	85	70	60	20	365
P 22	20	80	70	60	80	20	330
P 25	60	90	30	85	80	35	380
P 26	55	100	75	85	70	75	460
P 27	75	65	85	80	35	70	410
P 28	5	80	65	70	80	35	335
P 29	60	80	65	70	65	70	410

5.1.3 HMD Positions

Figure 5-3 shows the mean motion per second experienced by subjects for each testing scenario. Differences are evident between Test A (Low Load) and Test B (High Load), with the low load condition generating consistently lower motion per second than the high load condition. This difference is slight (3 mm/second) during the pre-test measurements, suggesting a relatively consistent baseline measurement for both groups. From this point, the gap in the mean amount of instability widens during the shooting tests and sorting tests (9 mm / second and 11mm / second, respectively). Table 5-1 shows subjects recorded higher final SSQ-T scores after Test B than Test A, subjectively stating subjects were experiencing greater sickness levels at the end of Test B than at the end of Test A (an increase of 35% in SSQ-T scores).

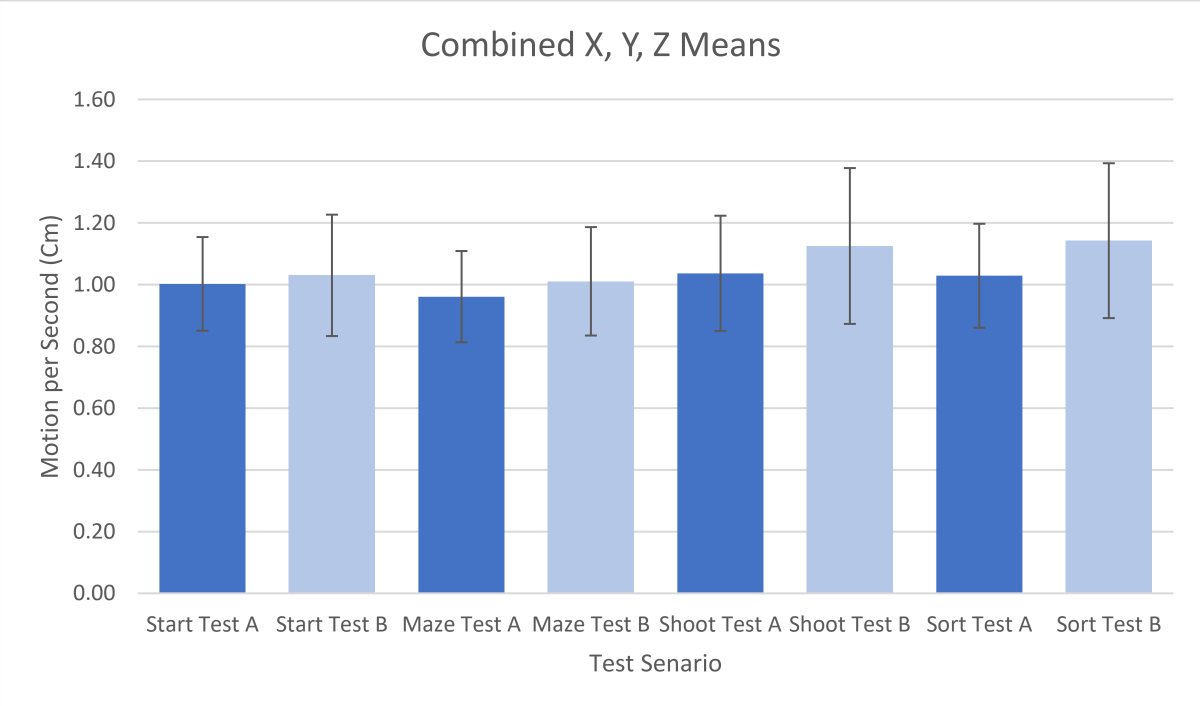


Figure 5-3:- Mean Motion of all Subjects per Test.

	Start Test A	Start Test B	Maze Test A	Maze Test B	Shoot Test A	Shoot Test B	Sort Test A	Sort Test B
Mean	1.00	1.03	0.96	1.01	1.04	1.13	1.03	1.14
STD	0.15	0.20	0.15	0.18	0.19	0.25	0.17	0.25
SSQ-T	14.65	19.79	14.65	19.79	14.65	19.79	14.65	19.79

Table 5-1:- Mean, STD Deviation and SSQ-T Values for Each Test Scenario (SSQ-T scores are post-experiment test results).

5.1.4 Balance Board

Table 5-2 shows us that there is very little difference in the mean values for the pre-test values (17.72 mm). This measurement equates to roughly equal baseline stability in each subject at the start of testing. Table 5-2 also shows an even smaller difference between the post-test scenarios (0.71 mm). This result means, on average, a slight increase in instability is present in both scenarios. The result is that both test scenarios generate the same level of instability in subjects within the test.

Scenario	Mean Value	Standard Deviation
Pre-Test, Low Load (Test A)	470.37 mm	84.35 mm
Post-Test Low Load (Test A)	492.65 mm	104.94 mm
Pre-Test, High Load (Test B)	488.09 mm	112.04 mm
Post-Test, High Load (Test B)	491.94 mm	103.89 mm

Table 5-2:- Mean Values for COP Distance Travelled in Each Experiment Scenario

5.2 Validation of Hypotheses:

5.2.1 Hypothesis 1:- Cognitive load has a significant impact on cybersickness onset and severity during the usage of virtual reality head-mounted display environments.

The results suggest that increases in cognitive load provide a significant increase in SSQ-T scores and this significantly impacts the level of cybersickness experienced by a subject. This increase in SSQ-T scores is accompanied by an increase in head motion (shown in Table 5-3), which, while not as severe as the increase in SSQ-T scores, is still significant evidence of a detectable difference observable solely from headset positional data.

Of interest, the type of cognitive load may be a factor in the overall relationship between cognitive load and cybersickness. The mean motion values for the maze task seem to indicate an increase in stability, with a reduction of path length evident in both tests (A 4% reduction in Task A and a 2% reduction in Task B). While this result is not statistically significant (U = 244, P = 0.185), it is still a very unexpected result. It seems to show subjects becoming more comfortable with the environment after a period of significant stimulation induced by user-controlled locomotion and rotation. These two factors are known to have a significant impact on cybersickness symptoms.

		Start Test A	Start Test B	Maze Test A	Maze Test B	Shoot Test A	Shoot Test B	Sort Test A	Sort Test B
X Motion / s	Mean	0.672	0.685	0.654	0.690	0.704	0.766	0.696	0.779
	STD	0.110	0.158	0.117	0.137	0.156	0.201	0.137	0.185
Y Motion / s	Mean	0.482	0.497	0.451	0.471	0.495	0.522	0.489	0.539
	STD	0.075	0.090	0.083	0.084	0.084	0.123	0.062	0.121
Z Motion / s	Mean	0.317	0.331	0.299	0.318	0.316	0.353	0.323	0.350
	STD	0.053	0.065	0.040	0.044	0.058	0.097	0.055	0.067

Table 5-3:- Axis Motion / s for Individual Axis per Test

While analysis of mean values shows a general trend, this sample contains a range of subjects with different tolerances for sickness. The small sample size means these results may not be reliable. Therefore, these results require further analysis before a model can be proposed based upon these findings with any confidence.

5.2.1.1 Effect of user experience on subject sickness

When comparing the difference between SSQ-T sickness scores between experienced users and inexperienced users, Figure 5-4 shows the frequency distribution of SSQ-T scores for both groups showing inexperienced users generating most of the higher SSQ-T scores. Moreover, a Mann-Whitney U Test (U = 369.0, P-Value = 0.03) shows the difference between the two data sets are significant. This result is not unexpected as user experience with VR is a factor in determining sickness, mainly because people tend not to use equipment that makes them sick and thus never become experienced users.

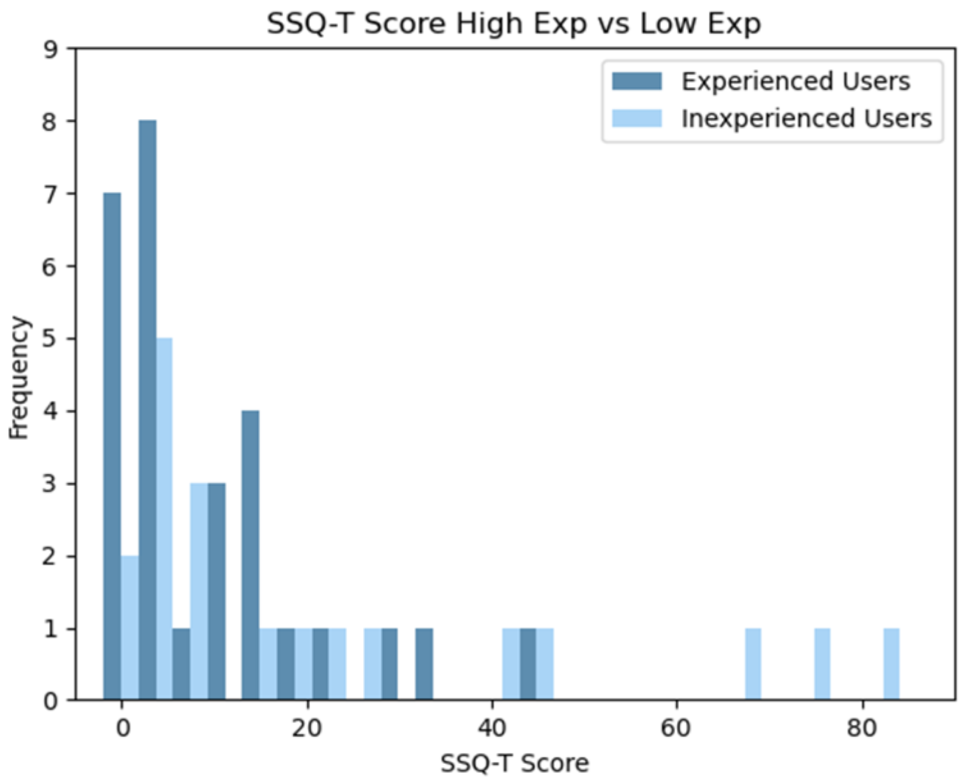


Figure 5-4:- Frequency Distribution of SSQ-T Scores for low Experience and High Experience Users.

5.2.1.2 Effect of test order on subject sickness

Figure 5-5 Shows the mean SSQ-T scores of the first test undertaken by the subject vs the second test undertaken. This comparison ignores whether the test was performed under high or low cognitive load and is just concerned with the order in which subjects completed the test. Results still show a heavy positive skew in both scenarios (1st runs Skew = 2.16, Kurtosis = 4.25 2nd runs Skew = 1.80, Kurtosis = 2.60) 1st runs generated a mean SSQ-T score of 21.66 (STD DEV = 27.09), 2nd runs generate a mean SSQ-T score of 12.78 (STD DEV = 17.87). This result shows a significant difference between the two runs, with the first test run commonly generating higher sickness levels than the second run. A Mann-Whitney test of the frequency distributions (Figure 5-6) of SSQ-T scores (U = 373.5, P = 0.03) confirms this finding.

Figure 5-5:- Mean SSQ-T Scores, First Run Vs Second Run

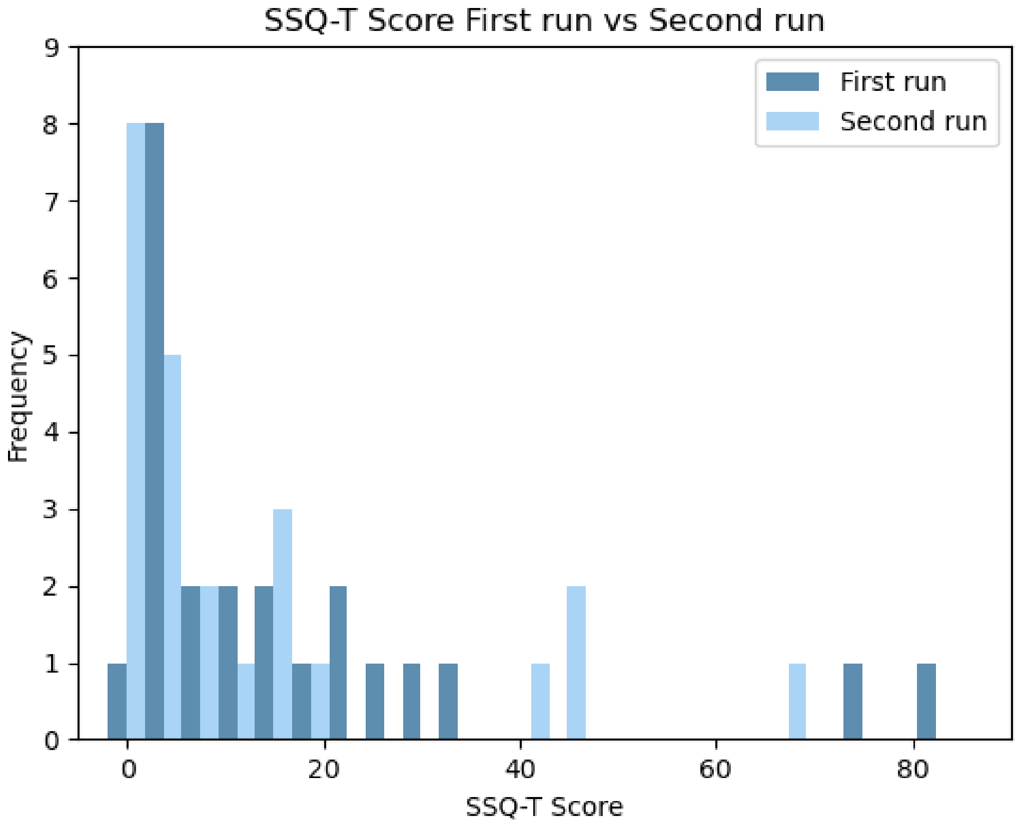


Figure 5-6:- Frequency Distribution SSQ-T Scores, First run and Second run

Two possible explanations exist for this result. The first possibility is that users are rapidly acclimatising to the VR environment and therefore getting less sick in the second run of the experiment. The effect of repeated exposure to VR environments has shown a possible decrease in sickness symptoms (Domeyer et al., 2013; Taylor et al., 2011) and could explain the reduction in scores. However, as an alternative, our experiment requires learning how to navigate and perform the test during the first experiment run. This learning effect may have a significant impact on the SSQ scores of the user. In either case, the test order significantly impacts the user's sickness levels, although its source cannot be determined.

5.2.1.3 Confounding factors

The results show that both test order and user experience significantly impact subjects reported SSQ-T scores, with a much higher significance than task complexity. This result suggests that while task complexity may have a minor impact on SSQ-T scores in general, the effect is much less significant than other more well-known factors.

Having seen the impact test order and subject experience has on SSQ-T scores, a comparison of High vs Low cognitive load under these conditions may yield a difference in SSQ-T scores exacerbated by these conditions. However, the current sample has insufficient data to analyse these factors and therefore needs to be considered during any further study. It is worth noting that each of these results has a large standard deviation. The large standard deviation of the results presented here suggests that these results may not hold for a larger sample.

5.2.1.4 Review

From the above sections, it is clear that cognitive load has a significant impact on cybersickness onset and severity during the usage of virtual reality head-mounted display environments. The exact impact on onset time and severity cannot be established due to the limited sample sizes of these tests. This result may be partially explained by user experience and test order, which gave inexperienced users significantly more experience in VR environments. As such this hypothesis can be considered to be confirmed for the purposes of this work.

5.2.2 Hypothesis 2:- A model representing actual VR usage severely impacts stability in subjects when recorded via COP pressure measurements or VR HMD positional data.

As described earlier most research scenarios are either simple limited movement environments or extremely provocative. The design of this experiment utilised tasks similar to traditional game environments involving movement, actions, and visual scanning across a range of cognitive loads.The Figure 5-7 shows the mean motion per second for each of the task types and shows clear differences in the path lengths between Test A and Test B.

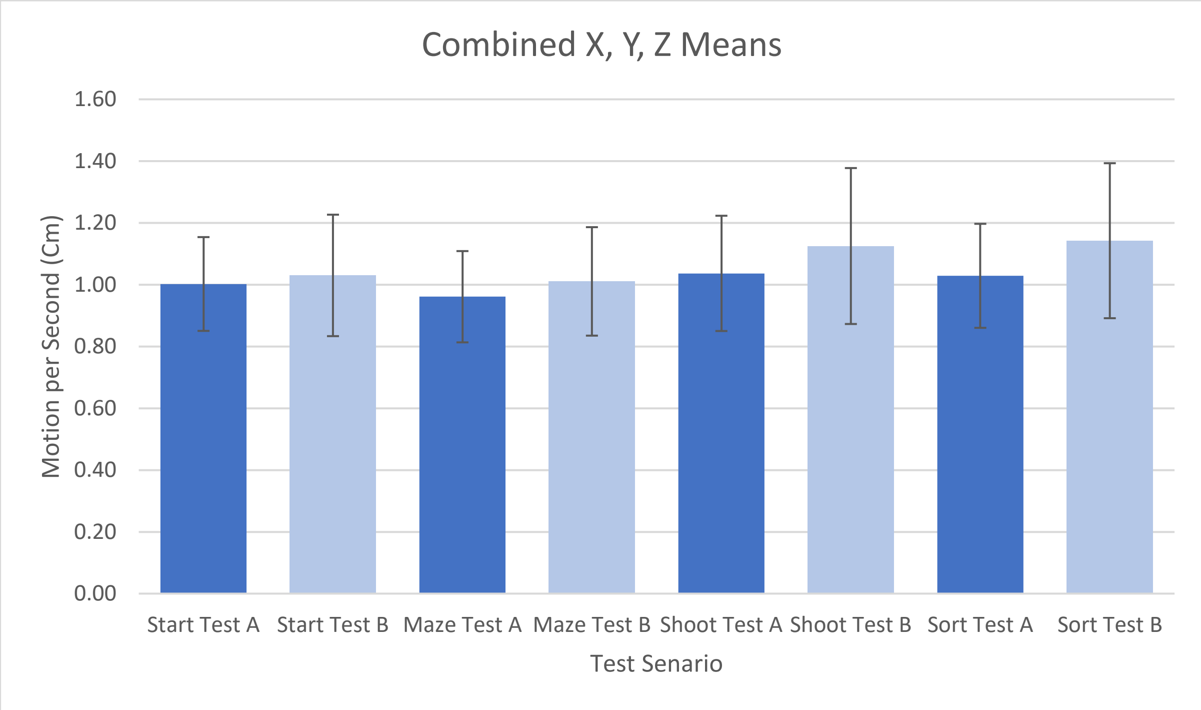


Figure 5-7:- Mean Motion of all Subjects per Test.

5.2.2.1 Maze task

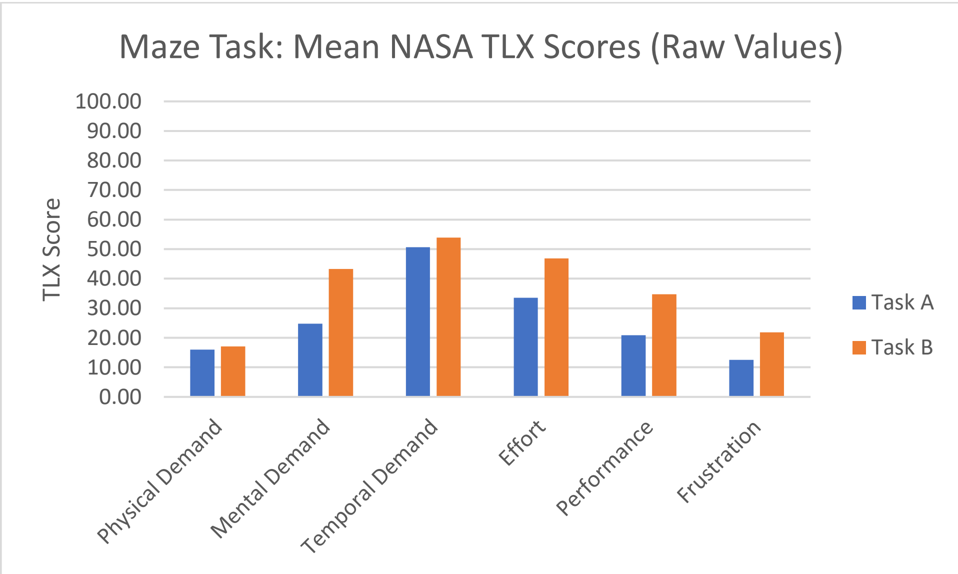


Figure 5-8:- Maze Task: Mean NASA TLX Scores (Raw Values)

The Nasa TLX results Figure 5-8 indicate that the Task A and B differences primarily affected the amount of effort and frustration encountered by the test participant. This is correlated with the very minimal differences in COP path lengths shown above.

5.2.2.2 Shooting Task

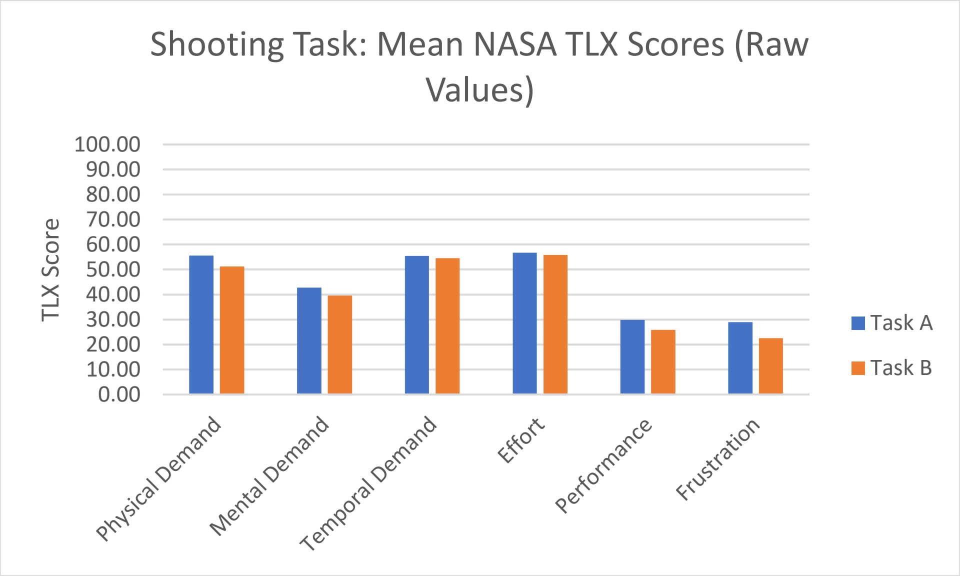


Figure 5-9:- Shooting Task: Mean NASA TLX Scores (Raw Values)

The Nasa TLX results Figure 5-9 indicate a perceived identical task load of Task A and B for the shooting test. When cross-referenced with Figure 5-7 , a significant increase exists in subjects' post-test instability after the shooting test in both conditions. Two explanations exist for this. Firstly, as the shooting task introduces additional search requirements into the test, the additional head motion added during this search task may significantly impact cybersickness. Secondly, the TLX results may not be sensitive enough to capture the difference between the two tasks, resulting in perceptually similar results.

In either case, it is impossible to separate the two scenarios without additional testing. What is required here is an additional test, adding spatial load into the scene in a manner so as not to induce additional head motion. Achieving this could mean adopting a model closer to traditional spatial load tests (Song et al., 2013), however identifying if these tests are truly reflective of assessing spatial load in VR requires additional research and appraisal.

5.2.2.3 Sort task

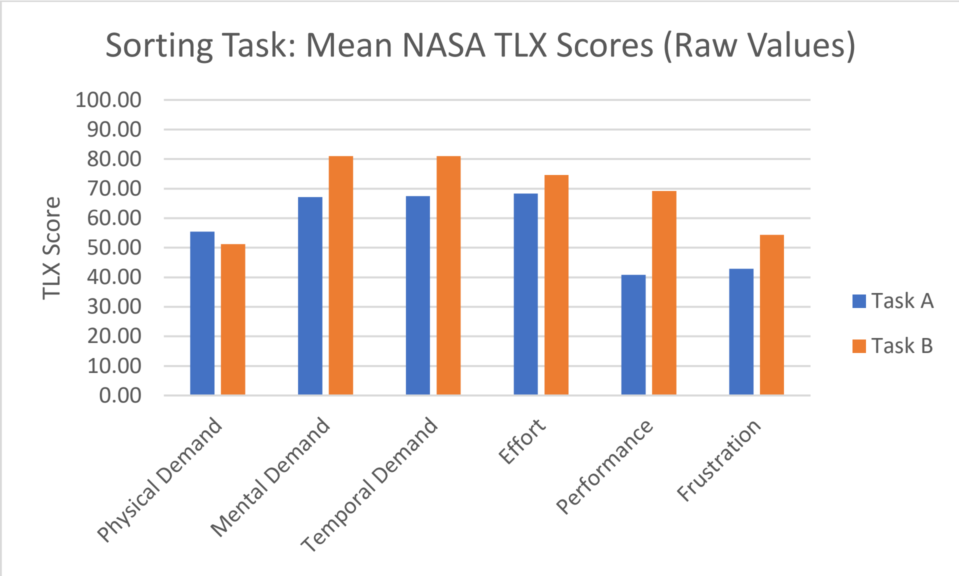


Figure 5-10:- Sorting Task: Mean NASA TLX Scores (Raw Values)

The sorting task (Figure 5-10) shows a significant jump in 5 of 6 task difficulty categories with the physical category being approximately even, showing a significant increase in perceived task difficulty accompanying a significant difference in instability. In this case, both tests' head motion is virtually identical, with no difference in the physical motions required by each task. This finding suggests that the increase in task difficulty correlates to the increase in instability.

Unfortunately, this task always came after the shooting test, and the instability measurements for each test are virtually identical. This limitation was due to the timeframe for experiments and low sample size due to COVID 19 risks. This issue means the possibility exists that the high instability measurement detected here may be a residual effect from the previous experiment. This scenario is not likely, due to the nature of the experiment design. Low activity periods separate the test periods, allowing instability to subside. However, the possibility exists that instability and sickness cannot dissipate within the environment while wearing the headset once acquired.

5.2.2.4 Location vs Orientation

The VR HMD was also capturing the rotational position of the headset. Figure 5-11 shows the mean values of the sum of all rotational measurements for each test. The pattern of increases mostly matches the pattern found in Figure 5-7, except that the increase in positional motion for Task B observed in Figure 5-7 during the shooting test occurs earlier when considering rotational measurements occurring during the Maze test and never reducing throughout the previous test. As these results are similar to those of the positional data, the reasons they occur are mostly similar. However, the increase in head rotation observed during the maze test rather than the shooting test suggests the contribution is either purely from VR usage or the specific task's effect.

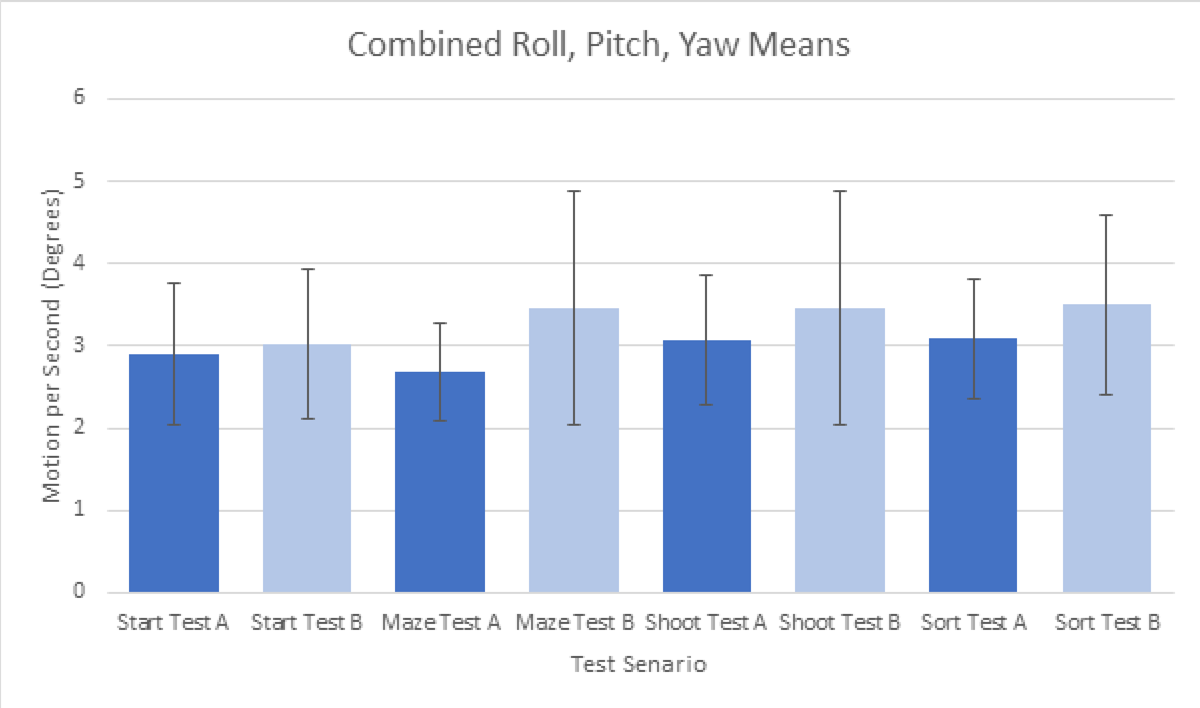


Figure 5-11:- Combined, Roll, Pitch and Yaw Means

A pure contribution from VR does not fit the data. If this were true, Task A and Task B would show an impact from the effect; yet Task A shows the same reduction during the maze task shown in the positional measurements. The same fact evidence also rules out the contribution coming entirely from the physical task in VR. Comparing the results from Figure 5-11 to the TLX data in Figure 5-8 shows a significant increase in mental complexity during Task B of the maze task. This result suggests a contribution from the memory load of the task is inducing the increase.

Combining this with the discoveries made regarding positional instability, it seems that two explanations are readily apparent. First, either mental load has an uneven effect on the different instability components, or the onset of physical motion precedes an increase in head motion. Unfortunately, the data collected during this study is insufficient to make this determination fully.

Examining Figure 5-12, we see the same trend of increased motion for the first test runs compared to the second test runs. However, the difference is much smaller in ratio than found in the positional data. This finding still suggests that the effect of test order is significant in the rotational headset measurements. However, its lower ratio may indicate a less severe impact and that the effect of environment learning may have a lesser impact on head rotation than on head movement.

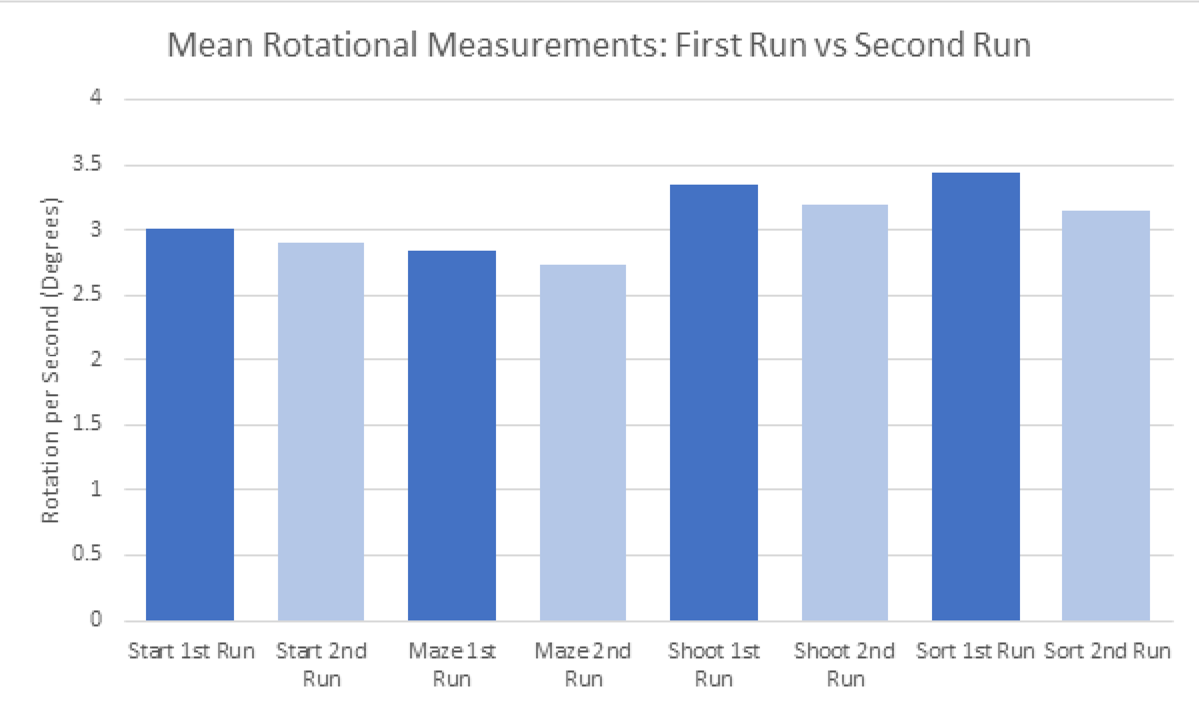


Figure 5-12:- Mean Rotational Motion Measurements: First Run vs Second Run.

Rotational motion also shows a significant increase with experience in an identical pattern to the positional motion. Additionally the distribution of rotational instability on a per user basis follows the same pattern as the positional data. The distribution of sick participants is mostly inseparable from their counterparts in the well group. The fact that all measurement patterns compared so far in terms of rotational and positional data mostly follow identical patterns is no surprise. Earlier suggestions that the rotational data may influence the positional data recorded by the headset prove apt. This result identifies the lack of accuracy initially highlighted by Niehorster et al. (2017) and identifies the issues created by not having an accurate head position representation. Instead, the measurement is a projection caused by the nature of the operation of HMD VR devices. Usage of an external head-tracking system would provide higher accuracy, like that used in Rebenitsch and Quinby, (2019). However, these are very specialised sets of hardware unlikely to be found in many VR setups.

The muddling of rotational and positional data may prove helpful for identifying subject posture changes suitable for assessing susceptibility to cybersickness onset. However, the results here definitely confirm statements by Niehorster et al. (2017) regarding the unsuitability of using headset measurements to record head position and orientation within the virtual environment accurately. In a practical usage scenario, it is clear that one set of measurements is polluted by the other.

5.2.2.5 Effect of user experience on instability

No difference in the distribution patterns of pre and post-test path length difference exists. This result suggests that novice users start with higher instability and end with higher levels of instability. One further comparison that would be useful to make would be to compare novice users' first run instability with their second run to see if an improvement is evident. However, with only eight samples matching these criteria, any results drawn here would be incredibly susceptible to outliers and be of low value. It does represent an avenue appropriate for future investigation with larger sample sizes.

5.2.2.6 Effect of test order on stability

Figure 5-13 shows the comparison of instability measurements for first runs vs second runs. First runs show significantly higher instability than second runs with the shoot test seeing a decrease in stability measurements of 7.7% during the shoot task and 5.31% in the sorting task. This result suggests that familiarity with the specific environment reduces instability within the environment. However, Figure 5-14 shows that experienced users exhibit more instability than low users across all tests.

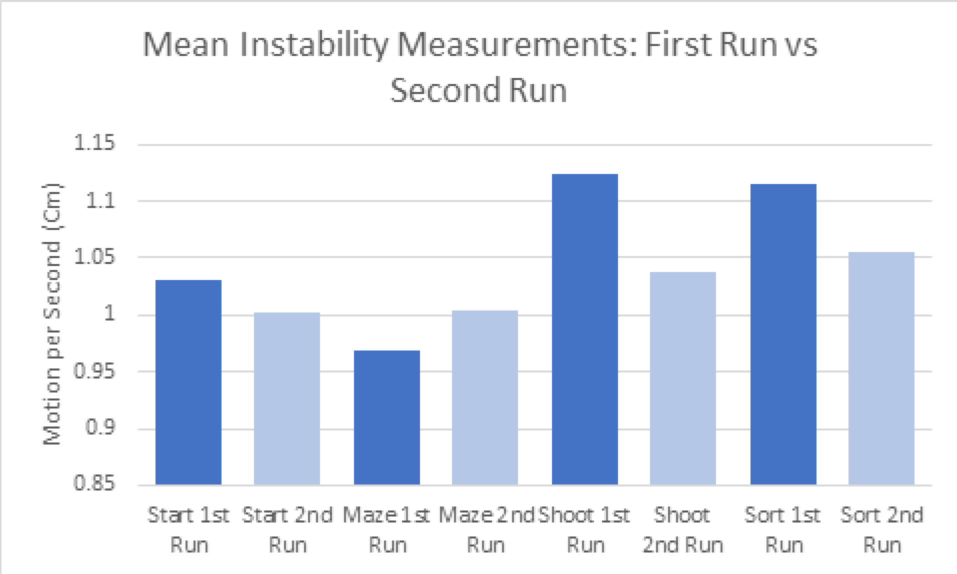


Figure 5-13:- Mean Instability Measurements: First Run vs Second Run

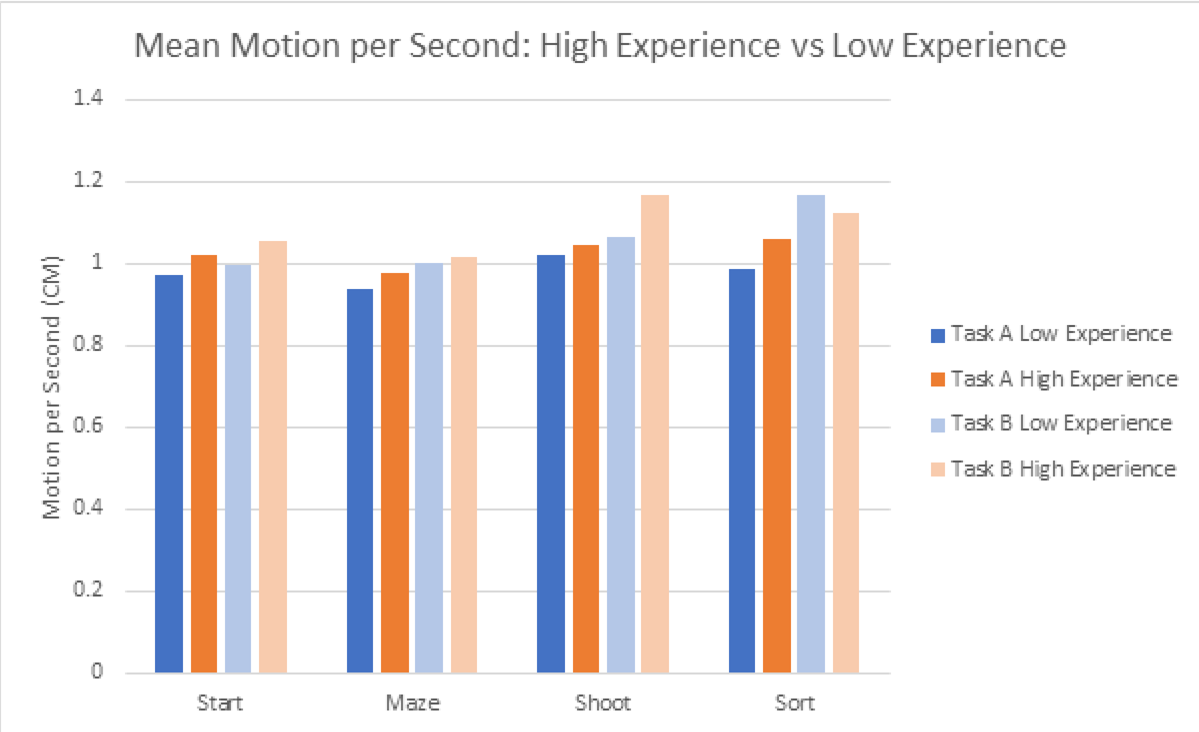


Figure 5-14:- Mean Motion per Second: High Experience vs Low Experience Users

This result suggests that reductions in instability from experience with VR environments do not necessarily transfer from different experiences with alternate virtual environments and VR hardware setups. However, tolerance to sickness does increase with experience, but associated instability does not go down, with the possibility that an increase in instability is indicative of experience with VR. It is worth noting that the sample number is very low (10 low experienced vs 14 high experienced). To make any definitive conclusions would be too unreliable from such a small data set. The results mean that the relationship between VR experience and instability is worth further study.

5.2.2.7 Contribution of individual axes to overall COP

Table 5-3 shows the variance in individual axis measurements for each testing scenario. The X-axis represents anterior/posterior motion made by the user (forwards backwards), the Y-axis motion is medial/lateral motion (side to side motion), and the Z-axis is vertical (up and down motion). Motion recorded suggests all tests saw the most significant motion in the anterior/posterior axis, followed by lateral motion with vertical motion showing the smallest contribution. The more substantial anterior/posterior measurement is consistent with the measurements reported in Widdowson et al. (2019), which shows a significant increase in anterior/posterior motion compared to medial-lateral motion. However, this study also demonstrates an increase in these motions attributed to the task's cognitive load, rather than wholly from the visual stimuli presented in Widdowson et al. (2019). A similar process has been performed for the rotational motion by looking at the individual contributions from the roll pitch and yaw axes. The Pitch axis shows the most significant contribution to overall instability, with Yaw providing the second highest.

Ideally, measuring head position Z motion should be minimal, as the user's head position should not rise or fall very much if the stance is correctly maintained. As such objective z-motion can be considered to largely be noise within the samples, however, as the data is polluted by rotation and position of the headset (see section 6.5.2) this noise can be difficult to remove. Several factors could be impacting this. Firstly, maintenance of the stance may not be perfect, rising and falling slightly as the user records the data. Secondly, the issues relating to the ground plane's incorrect alignment will take some anterior/posterior and lateral motion and translate it into represented vertical motion. Finally, noise from tracking influences will contribute to this. This finding importantly highlights this data's noisy nature and the fact that it is essential to understand these measurements' imperfect nature, especially in a scientific context. The results acquired from the respective axis are indicators of approximate motion in these directions and not absolutes.

Interestingly, no significant change in the distribution of either the individual axis of motion or any of the rotation axis from task to task. Individual tasks do not appear to affect each measurement's components. Instead, they seem to apply a uniform increase or decrease in all components.

5.2.2.8 Replication of quiet stand during gameplay

5.2.2.8.1 Maze Task, Corridor Measurements.

The maze task provided the opportunity to establish the effectiveness of recording stability measurements during known locomotion periods. Two volumes were specified within the environment to determine when subjects were performing a known task (shown in Figure 5-15), the approach corridor to the exit of the environment (defined as volume A), and the corridor connecting the exit to the start of test used to restart the test (defined as volume B). Both volumes' virtual dimensions were defined as 7 meters long by 4 meters wide, both in straight corridors with no expected changes in motion from the subject. Volume B faces the point where the subject acquires the code for the maze, and as such, the subject will be under increasing mental load during this period. As subjects do not behave in any specific way, the possibility did exist for backtracking or deviation from this model. During the observation of participants, no subjects deviated significantly from the expected pattern of interaction. However, due to the lack of restrictions on the subject, a deviation from the expected interaction plan is a significant possibility during actual usage scenarios.

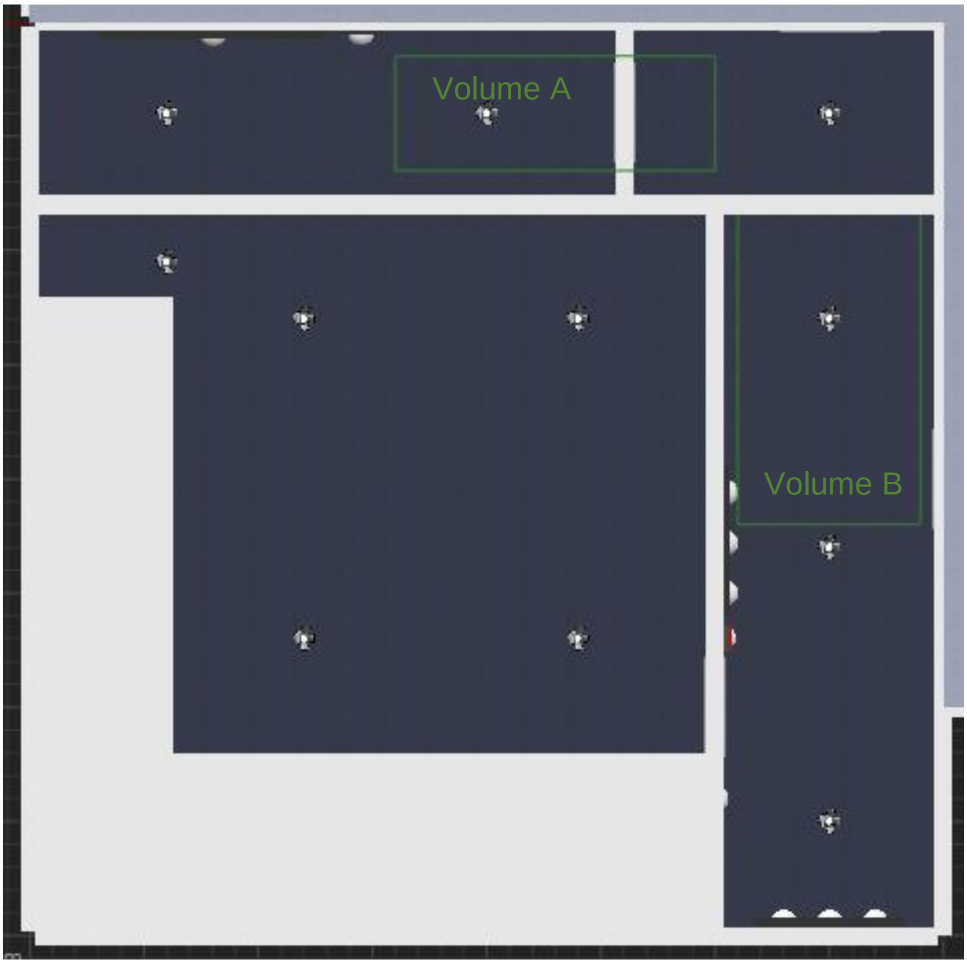
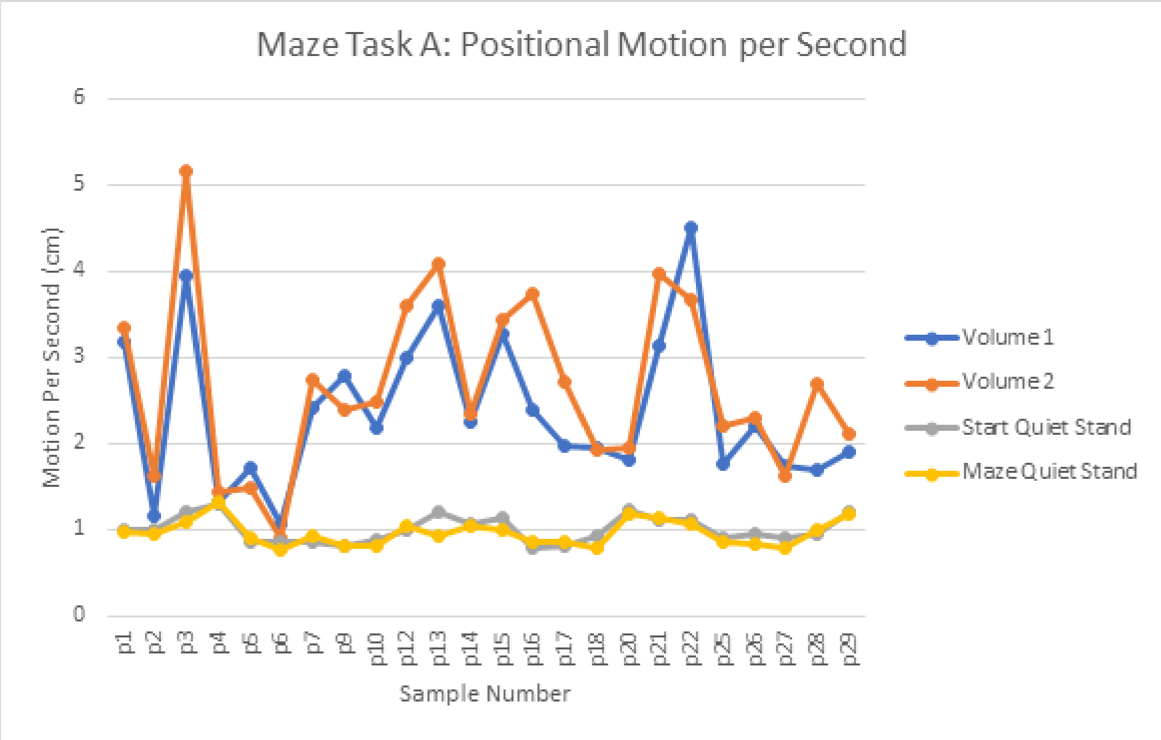


Figure 5-15:- Top-Down View of the Maze Task, Volumes A and B Highlighted in Green.

Figure 5-16 shows the physical motion and rotational motion per second for both volumes during task A. Results for each volume show significant variation from sample to sample. Some samples even showed increased stability during the maze test sample compared to the quiet stand tests. Most results, however, show a significant increase in head motion compared to the baseline. This significant increase is not unexpected due to the increase in the number of potential sources of cybersickness influencing the subject within the environment, including the significant increase in vection induced by the locomotion method. In many cases, the amount of motion recorded is significantly different between the two volumes, but no trend is evident, and many samples record no change between the two. Task B follows a similar pattern but with higher magnitudes of difference.



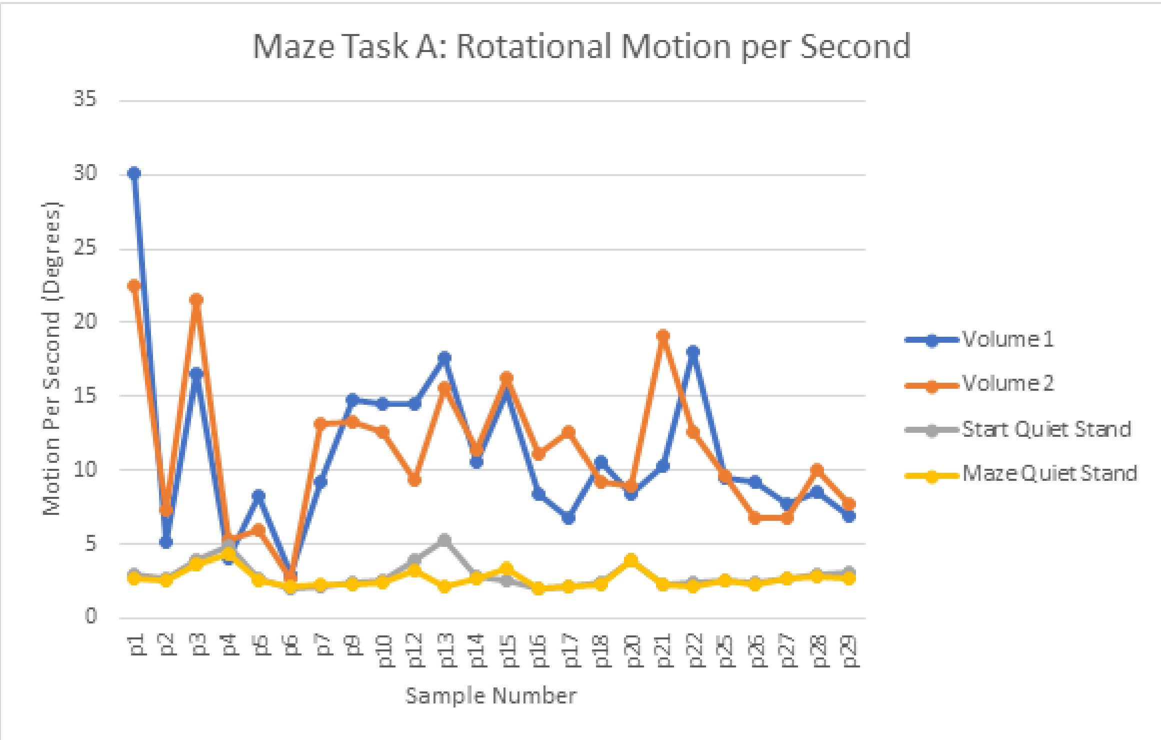


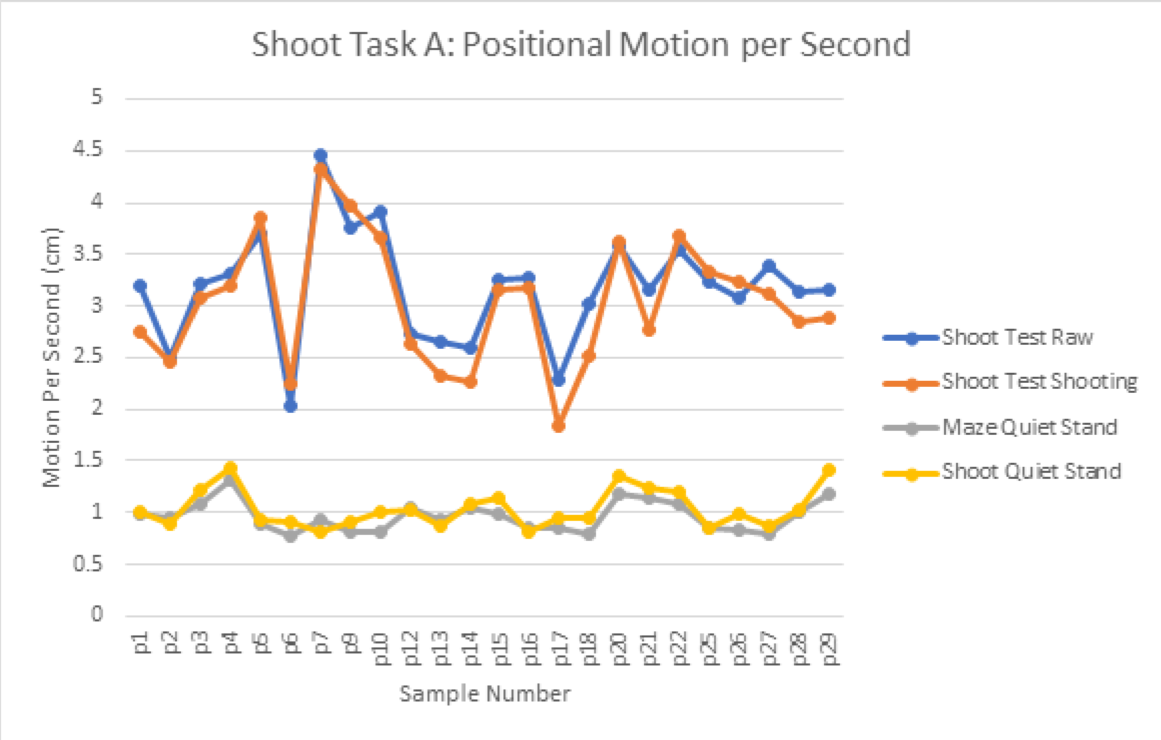
Figure 5-16:- Maze Task A: Positional and Rotational Motion per Second

Comparing these stability measurements to the SSQ-T scores recorded at the end of the test shows similar results to the comparisons made to the quiet stand measurements taken before and after the tests. These measurements' quantities do not correlate to values calculated for the final SSQ-T scores for either Task A or Task B. This result is not unexpected due to the lack of success in identifying sickness through the quiet stand measurements earlier in the report.

5.2.2.8.2 Shoot Task, Taking the Shot Measurement

The shooting task offered up another potential opportunity to take a stability sample during gameplay. When shooting in real life, a stable shooting stance is an integral part of making a good shot (Su et al., 2000). While this fact may or may not apply to shooting tasks conducted in VR environments, it is reasonable to assume users would be attempting a stable stance to improve their performance of the task. Thus, potentially providing a suitable opportunity to take a stability measurement during gameplay. It is also worth noting that a shooting task does not have much application outside of a video game environment. However, should the method yield valuable results, then the justification can be made to investigate the suitability for other tasks to be used in this manner.

The method of obtaining samples from this task was to observe the subject's actions during the test. The data sample for this process was created by selecting fifteen frames of data each time the user took a shot within the environment. The shot's accuracy was deemed unimportant due to the actions taken to make the shot being indifferent to the subject. This sampling method could generate frame overlap on shots; any occurrences were identified and corrected before calculating results.



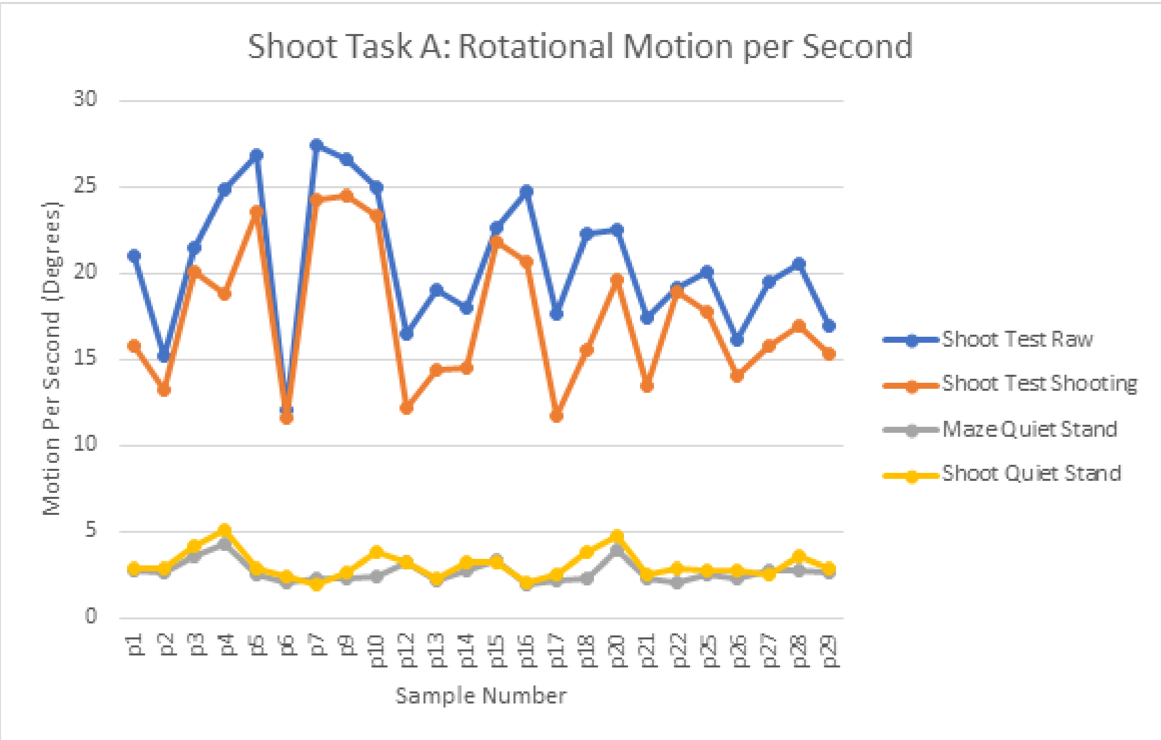


Figure 5-17:- Task A Shoot Task: Positional and Rotational Stability Measurements. Raw data vs Shooting Sampling Method

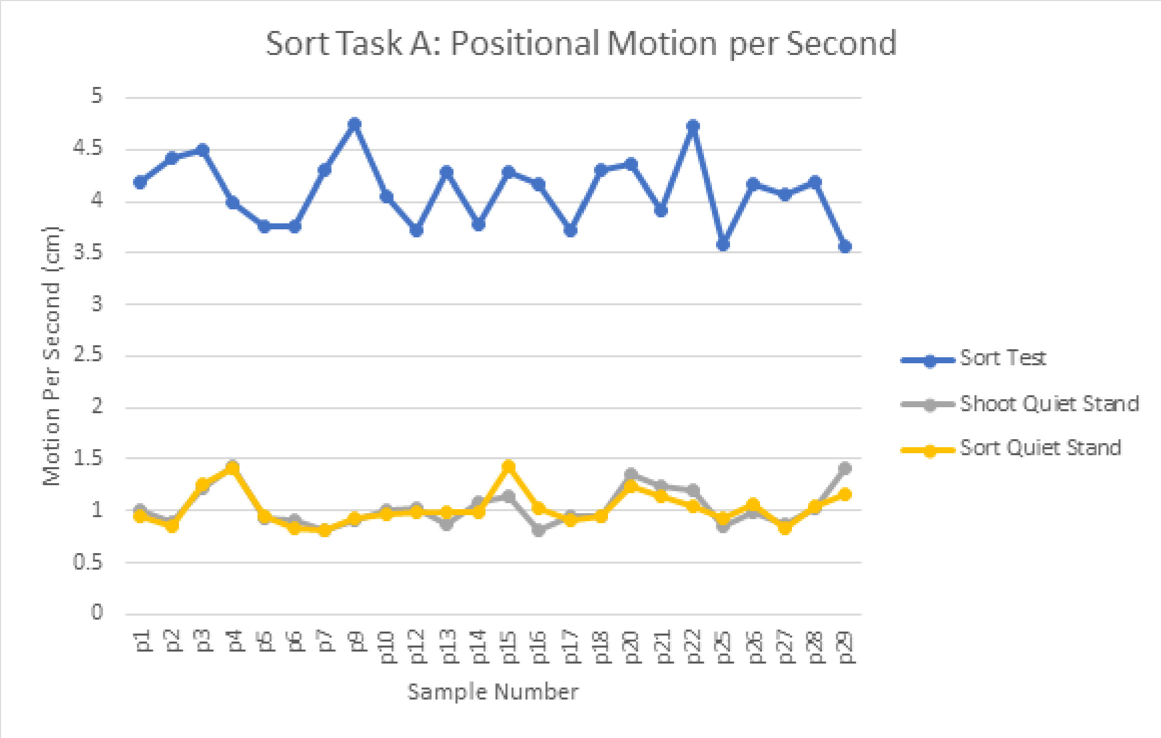
Figure 5-17 show a comparison between the shooting sampling method and just processing the raw results for Test A. Examining the positional data, the difference between the raw test analysis and the shooting results method seems minimal, suggesting no significant benefit to this data capture method for positional data. However, examining the rotational data shows a consistently lower value for most tests when comparing the raw test to the shooting data capture method. This data suggests that the magnitude of head rotation decreases while concentrating on lining up the shot. However, like the previous maze test, these results do not correlate with the results collected in the quiet stand period, with increases and decreases in the data not correlating to the pre or post-test quiet stand measurements. No differences were detectable between Task A and Task B

When comparing the magnitude of positional and rotational instability with the final SSQ-T results, like other similar comparisons made in this study, no discernible difference in results was identifiable between sick and well participants based on positional or rotational instability.

5.2.2.8.3 Sort Task Measurements.

The final method for this task analysis used the Sort task as an environment describing a situation with no predictable behaviour. While the task has the subjects performing the same task, there were various approaches to completing the task. As such, this made predicting periods of stability during the task difficult. Therefore, the proposal here was to take measurements from the whole sample whenever the filter criteria allowed. The aim was to see if the filter alone was sufficient to gather an approximation for stability during VR usage.

Figure 5-18 show the positional and rotational motion recorded during the Sort test. What is incredibly surprising is the consistency of the recorded positional data, which is significantly better than the other tests, even compared to the similar full test measurement made in the shooting task. This result is entirely unexpected as the Sort task by and far has the widest variety of different approaches. Explaining the consistency difference between this test and the results recorded in the shooting test is difficult. Task A's pace is generally slower than the shooting task; However, Task B’s pace is faster, and the same consistency improvement is evident. This issue aside, the results recorded still do not correlate with either the pre or post-test quiet stand samples. Again these results do not correlate with any of the SSQ-T scores taken shortly after the test.



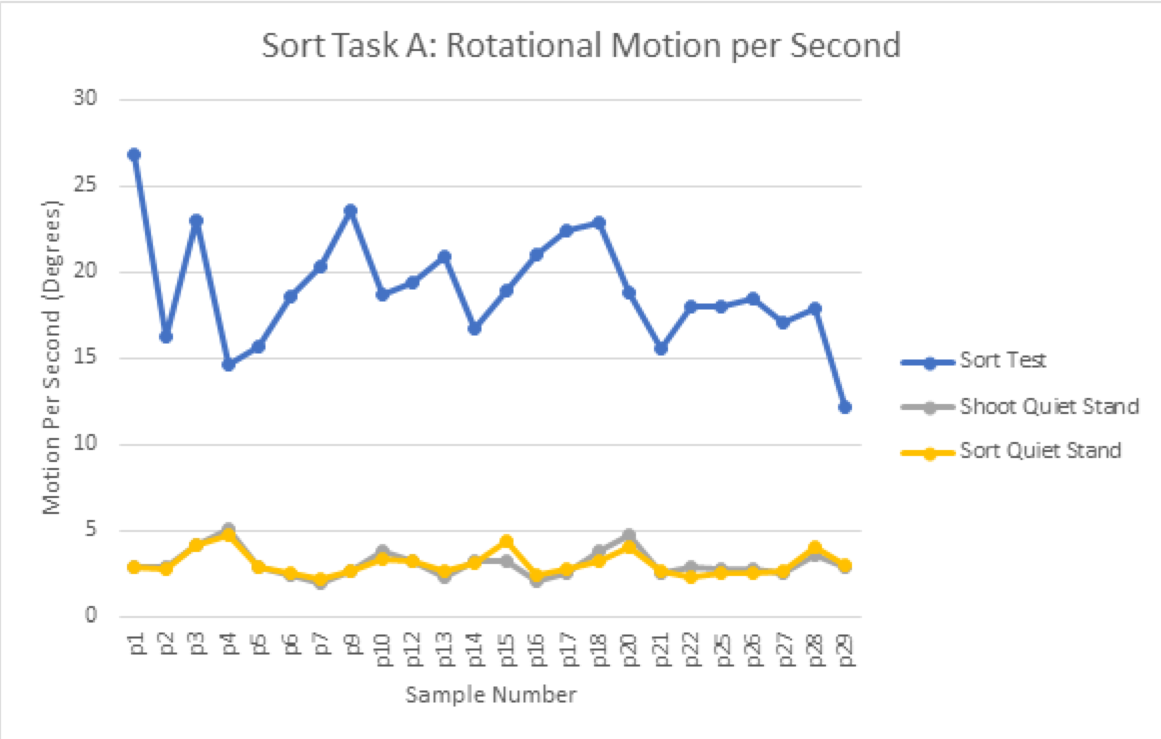


Figure 5-18:- Task A Sort Task: Positional and Rotational Stability Measurements.

5.2.2.8.4 Filter Adjustment

Calculation of the in-task results has used the filters developed during the quiet stand tests. As a result, any large motion with a magnitude greater than 0.1cm in a single frame (1/90th of a second) is considered erroneous and removed. The filter developed during the quiet stand took the standard deviation of the mean motion for the axis with the highest magnitude and multiplied it by 4. This filter may not be appropriate for this scenario, owing to the likelihood of a significant increase in baseline motion from the quiet stand. The subject is also dealing with significantly increased environmental stimulation and mental load at this time. The existing filter shows significant data losses even during periods where head motion should be minimal. Therefore, the results were reprocessed with different filter values to identify if this approach was overly aggressive and removed potentially useful results from the study. Therefore, as a final assessment, the filter value was relaxed to 0.4, four times the value's quiet stand measurements. The effect of this was to increase the magnitude of most results, but no improvements in the consistency of measurements were evident. Relaxing the filter did not resolve the issues with the maze task corridor measurements reporting instability values identical to those recorded in the quiet stand test. While this relaxing process could continue beyond this point, it is probably inappropriate. The further the filter is relaxed from the initial value, the more likely it is to incorporate deliberate motion or errors into the results accidentally. Analysis of these results shows no likelihood of improvement emerging.

5.2.2.9 Review

In this experiment, the realistic environments utilised showed that there was a significant difference in COP path length for certain aspects of cognitive load, however, the interaction between these aspects of cognitive load cannot be fully determined from these results. The use of an in-environment quiet stand equivalent was shown to be infeasible using COTS VR HMD equipment. As such this hypothesis can be considered to be confirmed for the purposes of this work.

5.2.3 Hypothesis 3:- Cognitive load has a significant impact on the stability of subjects during VR usage sufficient to make stability measurements made via tracking of VR HMD position inappropriate as an indicator of cybersickness.

The main driving force behind attempting to detect instability via VR HMD is to detect either the onset or a user's susceptibility to cybersickness. Therefore, any detectable differences in the motion of the sick and well participants could be used to propose a model capable of identifying these groups. Subjects are grouped into two categories: sick with a post-test SSQ-T score >= 20 or well based upon post-test SSQ-T scores < 20. Of the 24 samples, for test condition A (low load), five subjects were determined to be sick and nineteen well. For test condition B (high load), seven subjects were sick and seventeen well.

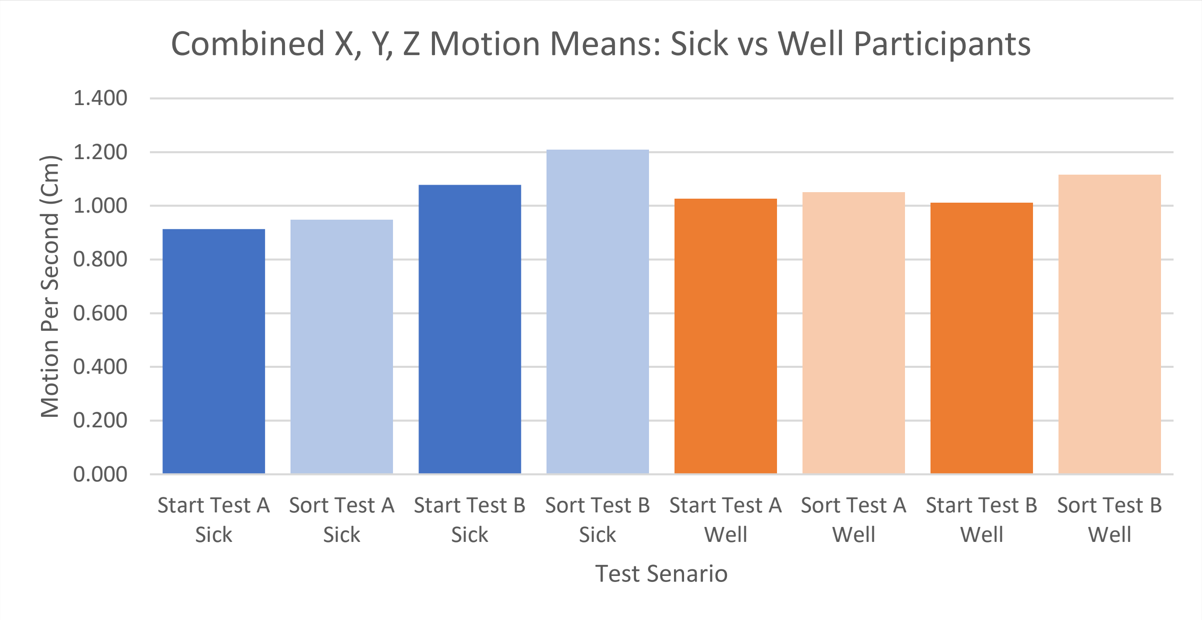


Figure 5-19:- Combined X, Y, Z Motion Means: Sick Participants vs Well Participants

This finding indicates that some effect of the load condition is impacting user stability. The high load (Task B) condition demonstrated a 10.6% increase in motion per second from the baseline measurement compared to a 4% increase for the low load condition. The difference between the two groups is reinforced by a Mann-Whitney U test showing a statistically significant difference in the distribution of the two groups' ratios (U = 194, P = 0.027).

5.2.3.1 *Stability of sick and well participants*

Figure 5-19 compares the mean instability of sick and well participants for the starting initial measurements and post-sort end of test measurements. Instability measurements rise at a greater rate for sick participants (3.7% for Test A and 12% for Test B) than well participants (2.3% for Test A and 10% for Test B). While this result demonstrates an increase in instability attributed to the sick and well conditions, it also shows the task and the mental load involved in that task have a far more significant impact (roughly four times the impact) than whether the user is suffering cybersickness. Man Whitney U tests show no significant difference in the distribution of ratios between sick and well participants in either Test A ($U = 47$, $P = 0.5$) or Test B ($U = 51.0$, $P = 0.306$). This result suggests the ratio gain observed in the mean may be obscuring variance in the individual results.

Figure 5-20 compares the measured instability of sick and well participants; the graphs illustrate the range of values recorded for each group's subjects. While the sick group contains a small number of participants, the range of motion values generated by these individuals generally falls within the range of values produced by the well group of participants. This result means that detecting sick or well participants based entirely on a single postural measurement is unfeasible.

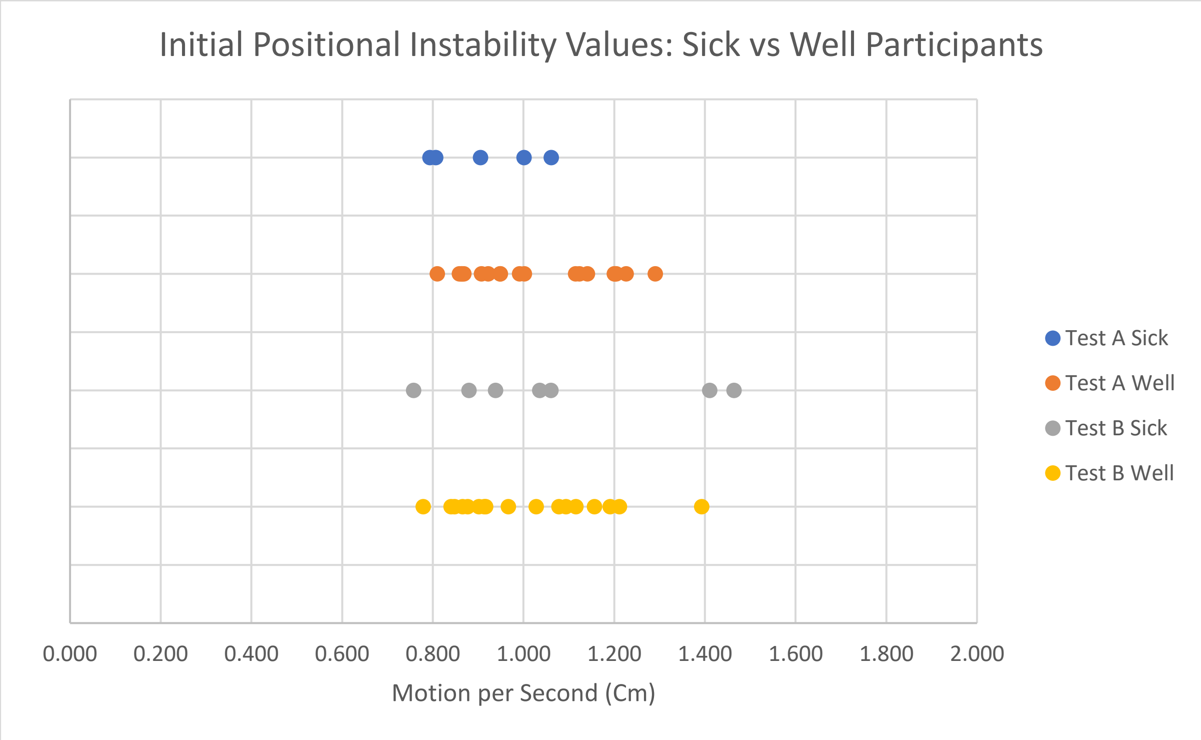


Figure 5-20:- Positional Instability Measurements, Sick vs Well Participants.

The mean instability values calculated earlier suggested an increase of approximately 10% in the total instability gained. However, these values' standard deviation is substantial, suggesting significant variability in these values. Therefore, analysing the variability on an individual level will help determine the feasibility of developing a model based on this data.

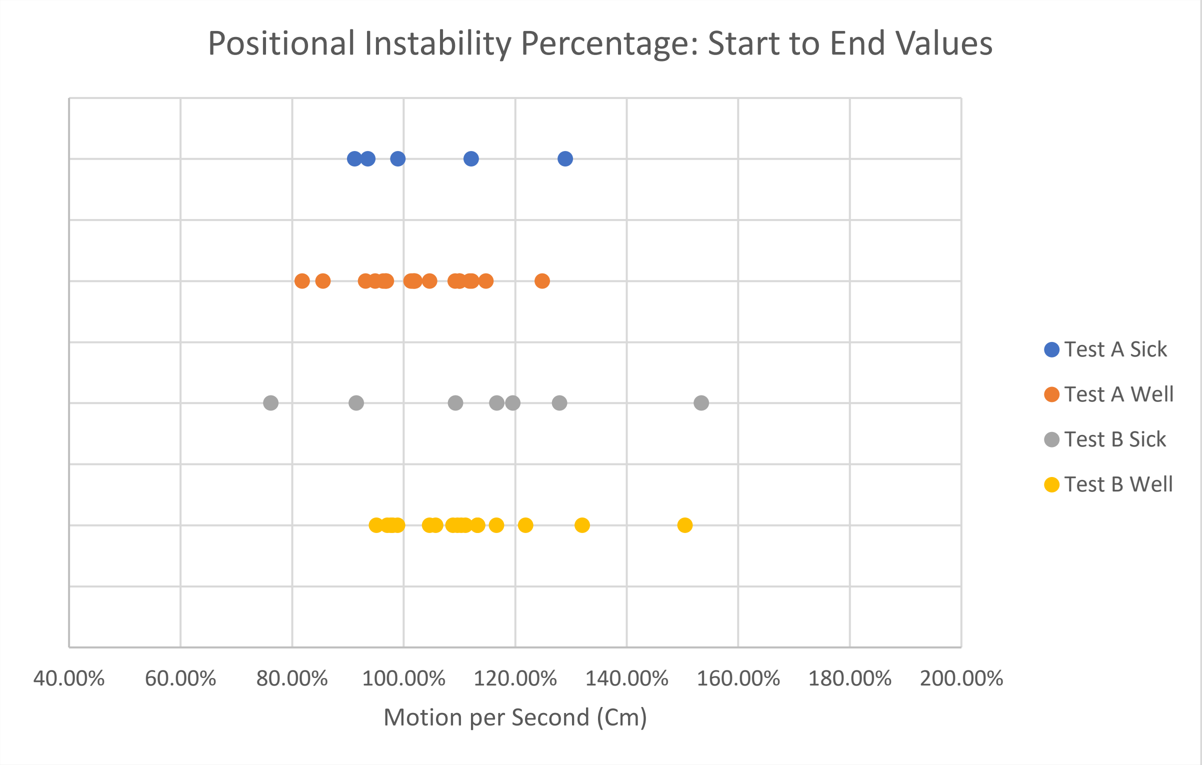


Figure 5-21:- Positional Instability Percentage: Start to End Values.

Figure 5-21 shows the difference in starting and ending measurements as a percentage. For a model based on the 10% increase found in the mean increase to be feasible, the sick participants' distribution should be around 110%. Figure 5-21 demonstrates that this is not the case. The range of sick values and well values represented in the chart cover the same range of 90 – 110% increase. The range of sick and well participants shows a similar distribution in each test condition.

The mean appraisal model seems to be an inaccurate representation of the actual picture here. Significant outliers (with +130% gains) seem to be polluting the accurate representation of what is going on. The data shows sick participants gaining and losing instability from start to end of testing in roughly equal measure. However, some cases exhibited severe instability gains. However, these severe gains were seen in both sick and well participants. They did not correlate to an increase in sickness. Examination of the rotational data in yields similar conclusions.

This discovery highlights a significant issue with most cybersickness studies; most have a low number of participants (see appendix D). Many studies have 24 – 30 participants, but a significant number have fewer than these. The study conducted here has proven 24 samples to have significant variance on the mean instability measured and SSQ-T scores recorded during testing. This finding may explain why so much disagreement exists around the question as to whether instability is an indicator of sickness, as outliers in small studies can have significant effects on the mean scores reported by these studies. Unfortunately, increasing the sample size in these studies is problematic primarily for logistical reasons. The number of participants available to recruit is limited and further hampered by the experiment's nature; generally, people do not want to be sick and thus are reluctant to participate.

Achieving a definitive answer would require the mass distribution of the test scenario would be required. Achieving this would require mass cooperation and investment from many research institutions or smaller-scale testing performed over a more extended period.

5.2.3.2 Cognitive Load Analysis

5.2.3.2.1 Maze cognitive load analysis

The maze task induced a memory load onto subjects. This test showed that the mean amount of motion in both rotational and positional measurements decreased from the starting quiet stand measurement to the post-test quiet stand measurement. While this difference was slight, it did suggest that the additional memory load of the task, even under the more significant load test B condition, did not cause an increase in instability. This result was unexpected as the task exposed the subjects to high provocation stimuli in locomotion, rotation and vection. The Maze test was also the only point at which any subjects withdrew from the test. This fact confirms that the environment was capable of inducing sickness symptoms, and as such, the results are not merely a case of subjects adjusting to the environment. As the difference between measurements is slight, the appropriate conclusion is that, on average, there is no significant difference between the results for the initial stability measurements and the post-task stability measurement. This result means that no evidence is found in this test that memory load influences instability in VR HMD environments.

5.2.3.2.2 Shooting cognitive load analysis

The second task, the shooting task, introduced a spatial load into the environment. This test showed a significant increase in both rotational and positional motion during the post-test quiet stand to the baseline. Explaining this increase requires careful consideration of all the factors for the test. The task performed here is different from the previous one and, as such, could be responsible for the increase. If this was true, and the increase was entirely down to the task rather than the task load, then the increase in instability would be uniform between the two testing conditions. However, the high load task showed a more significant increase than the low load condition. While this does not allow ruling out the task difference impacting instability, it does allow for the conclusion to be drawn that spatial load impacts VR HMD instability.

5.2.3.2.3 Sorting cognitive load analysis

The final test was the sorting test, which aimed to establish if a subject currently experiencing overstimulation would exhibit more instability. The results show a similar level of increase in rotational and positional motion to that of the shooting test, including the low load and high load condition differences, ruling out the difference in the task being wholly responsible for the increase. However, as recorded instability levels are not significantly different from those of the shooting test, and the fact that the shooting test always preceded the sorting test, the possibility that the instability gained during the previous test carried over to the sort test cannot be discounted. Therefore, no conclusion could be reached as to the impact of overstimulation on instability in VR HMD.

5.2.3.2.4 TLX cognitive load component analysis

To further reinforce this, the raw TLX values for each task can be compared to the instability measurements to see which components of the task influence the instability detected. Section 5.1.2 shows the raw TLX values obtained from the surveys conducted within the VR environment after each task. The three graphs compare the scores for each category for each test.

Physical demand produced low scores for the maze task and high scores for the shoot and sort tasks. This result is not unexpected as the trend in these scores seems to follow the amount of arm movement required to perform the task rather than the amount of movement performed in the environment, suggesting subjects can effectively separate the task's digital and virtual components.

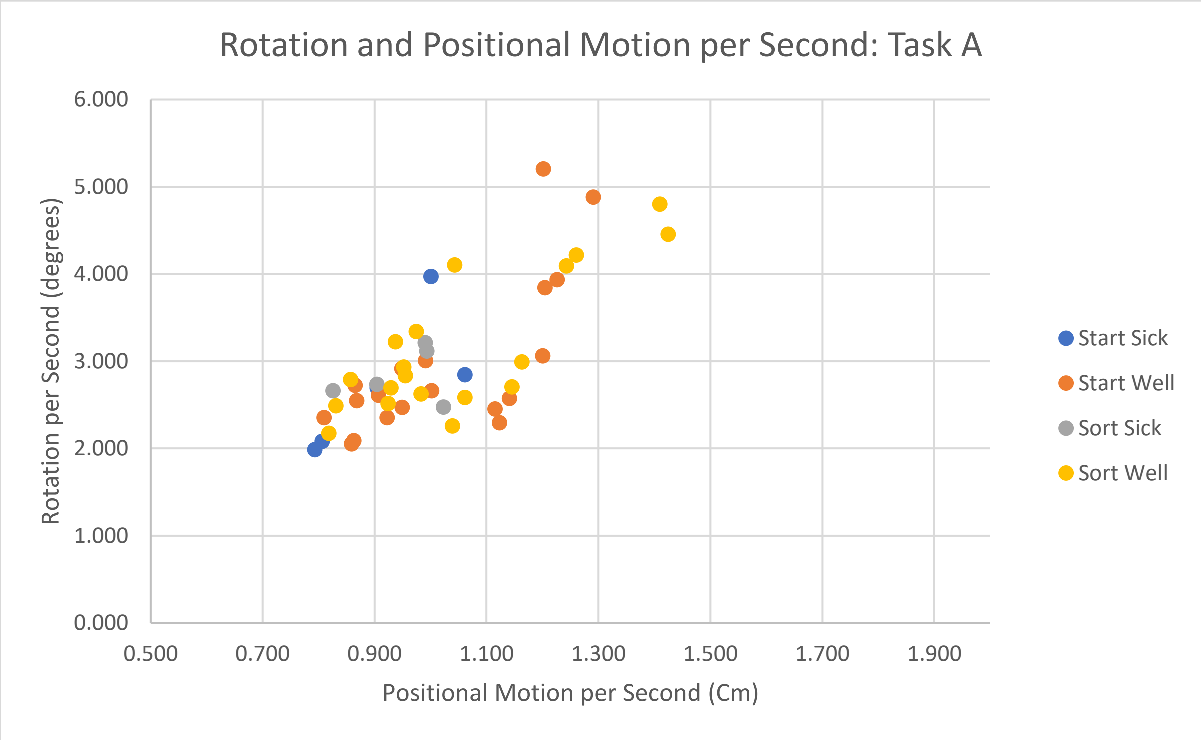
Mental demand shows a significant spike in the sorting task, suggesting a significantly higher mental task load was present during this condition. However, the shooting task scores are almost identical for each task (Task A value = 42.71, Task B = value = 39.58), suggesting no significant difference in the two tasks' mental load. Temporal load and effort and performance seem to follow a similar pattern.

Frustration values were low for the maze and sort tasks but relatively high for the sorting task. This result is not unexpected as the sorting task is the most complex in both conditions and is certainly more taxing than the other tests.

What is conclusive is that within the environment, pure memory load added has little to no impact on the instability measured via the headset as only a minimal change is evident between the two test conditions. Task A (low load) saw a 4mm reduction, and Task B (high load) saw a 2mm reduction. It was also the period during which all major sickness cases terminating the experiment occurred. Despite this, no significant change between the instability measured in the maze test condition and the test's baseline measurements was evident. This result means that the difference in memory loads of the task did not affect user instability.

5.2.3.3 Combined factors effects of instability and cybersickness

Finally, looking at both data sets in unison to see if any combination of factors presents some insight into instability as a predictor of cybersickness. Figure 5-22 presents Task A and Task B's graphs showing each subject’s positional and rotational instability at the start and end of the test. Interestingly the post-test well condition (shown as yellow dots on the graph) seem to exist in the majority of position, rotation combinations exhibited during the other conditions. It is doubtful that any way of separating them based on this data could be derived. While some evidence has suggested that as instability increases, the likelihood of cybersickness onset and severity increases. No statistical evidence can identify a way of separating the two groups using positional or rotational motion.



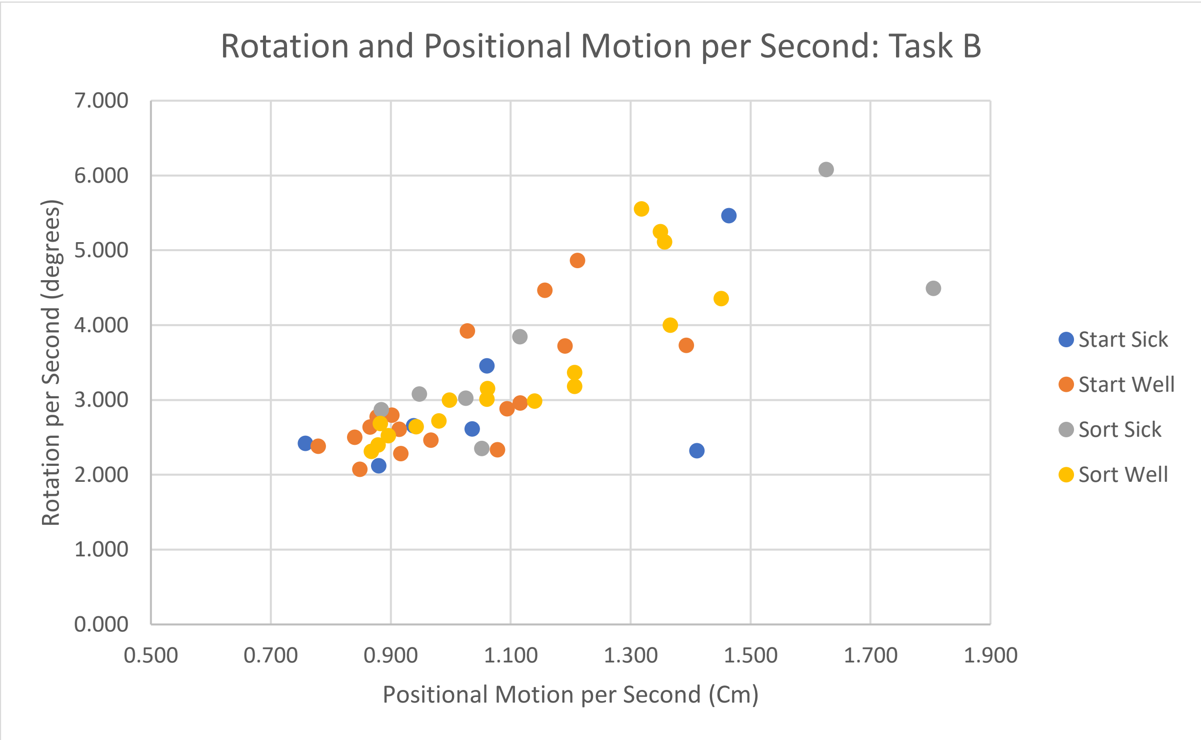


Figure 5-22:- Rotational and Positional data per Subject.

Figure 5-23 shows the difference between pre and post-tests for each subject. Surprisingly the mean gain for Test B (High Load) (Mean = 3.85mm) is significantly smaller than the mean gain for Test A (Low Load) (22.78 mm). However, the graph clearly shows a significant distribution within these results, which showed that some subjects increased their stability post-test. Neither test condition showed a prevalence for significant gains or losses in stability than the other condition, suggesting no significant contribution to stability from the cognitive load factor. Higher SSQ-T scores also do not seem to translate to higher instability measurements, although significant response results (SSQ-T > 20) tend to trend higher. This finding may mean that a positive trend in instability is only observable when sickness onset is more severe, enhancing our study's claim that rollercoaster VR rides are an inappropriate model for assessing cybersickness in VR under normal conditions. However, more results would be required to confirm this. One significant outlier result exists at around -220 mm. This outlier may be a result of significant anxiety or excitement. However, no evidence of either of these two factors was forthcoming during the observation of the subject during testing.

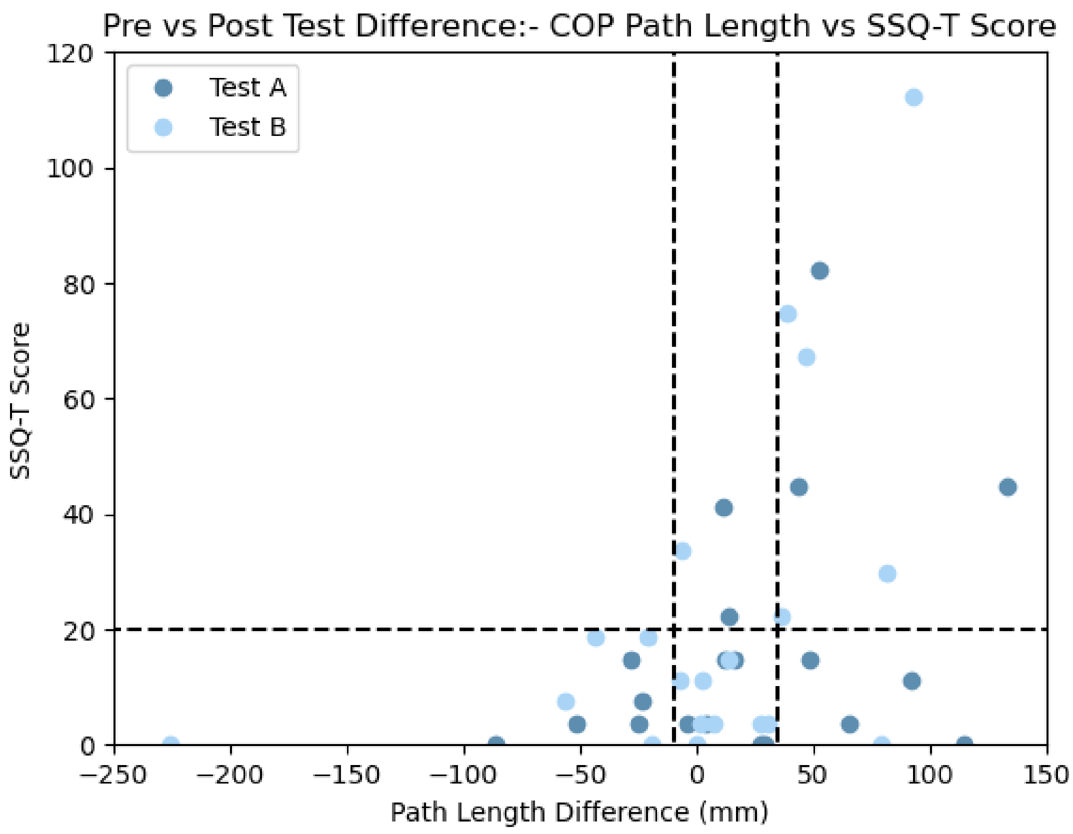


Figure 5-23:- Pre vs Post Test Difference in COP Path Lengths Compared with SSQ-T Scores, Highlighting Test Case. The dashed lines highlight SSQ-T 20 on the vertical axis and Path length difference values of -10 and 35mm on the horizontall axis.

Using the SSQ-T of greater than 20 as a threshold for determining sickness, there appears to be a distinct grouping (shown with dashed lines in Figure 5-23) that shows participants who saw at least a minor decrease in path length (> -10mm) were unlikely to be show symptoms. Those demonstrating a minimal change (-10 to +35mm) demonstrated moderate increases in SSQ-T scores, and those above this path length demonstrated a wide variety of symptoms. While this cannot show that an increased path length results in increased sickness, it can possibly be used as a detection threshold for other mechanisms and to say with some certainty that decreased path lengths indicate little concern over cybersickness symptoms.

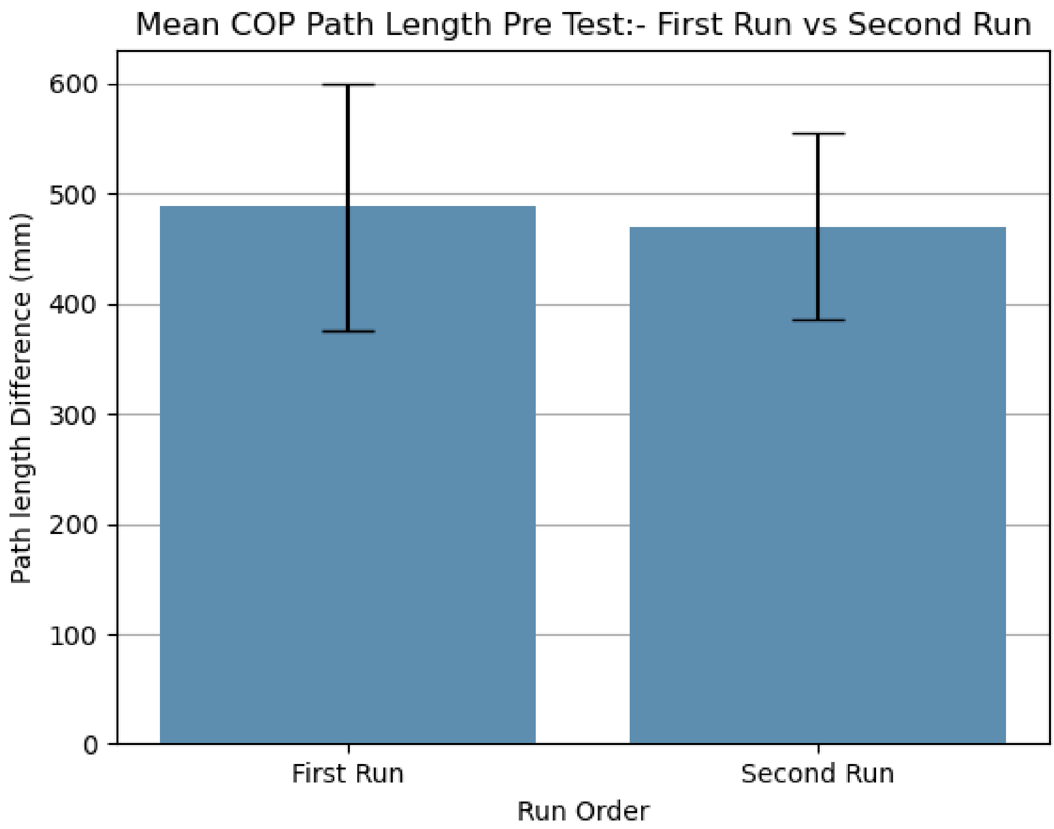


Figure 5-24:- Mean COP Path Lengths Pre-Test, First Run Vs Second Run (First Run Mean = 488.09 mm, STD = 112.04 mm, Second Run Mean = 470.37 mm, STD = 84.35 mm)

One theory as to why some subjects recorded short paths post-test may be pre-test anxiety, caused by the equipment or the experiment procedure (particularly the minor risk of COVID exposure) with the increase in stability from relief and relaxation. Figure 5-24 compares mean path lengths for pre-test first run tests vs second run tests. Mean values for each test value do seem to support this theory. A small reduction in mean path lengths for second runs is evident (Mean First Run = 488.09 mm, Mean Second Run = 470.37), representing an increase in second runs' stability. The means for each test run post-test (Figure 5-25) are practically identical (Mean First Run = 491.94 mm, Mean Second Run = 492.65) with a difference of less than 1 mm. This measurement suggests that the experiment order does not influence the difference in post-test path length recorded for each run.

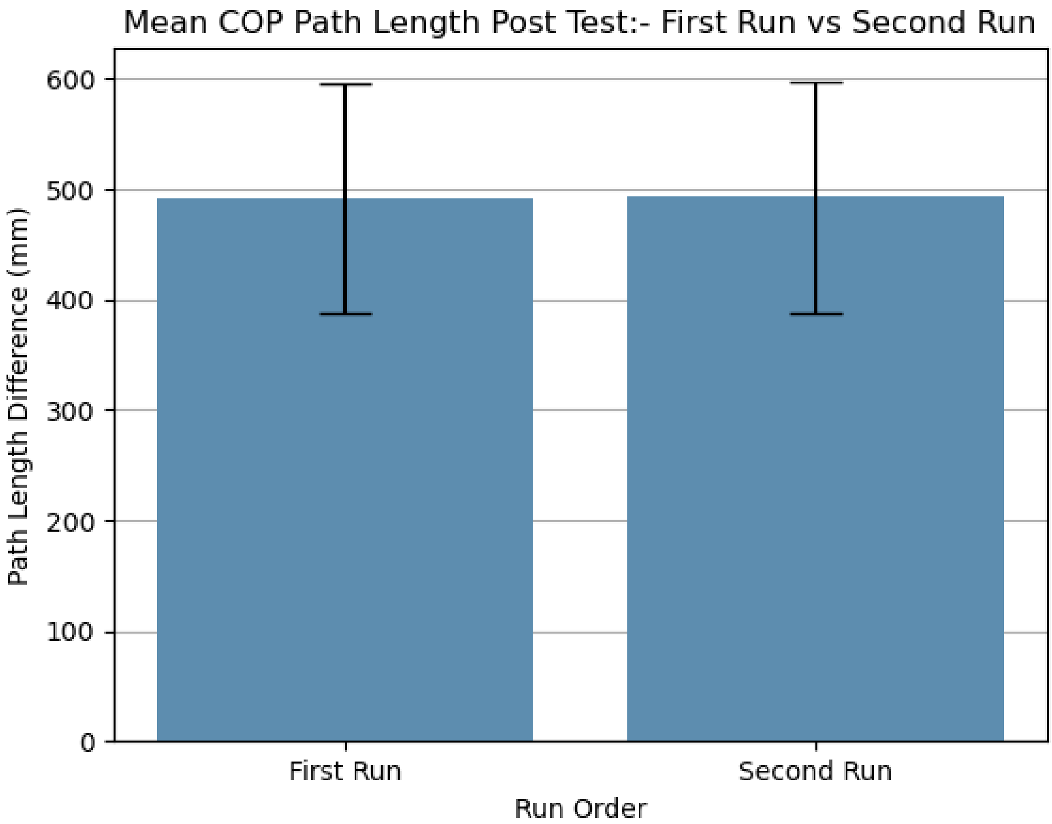


Figure 5-25:- Mean COP Path Lengths Post-Test, First Run Vs Second Run (First Run Mean = 491.94 mm, STD = 103.89 mm, Second Run Mean = 492.65 mm, STD = 104.95 mm)

Identifyin the run order of each sample shows no correlation between run order and path length difference, failing to show a predominance of negative path length differences associated with first-run tests. This result debunks the theory that the experiment order had a significant impact on this test's results.

Figure 5-26 shows the Mean COP path lengths for Low and High experienced subjects. A significant difference exists in the means of the two user groups (Low Experience Mean = 543.89 mm, STD = 99.13 mm, High Experience Mean = 457.90 mm, STD = 93.077 mm) with a difference of 85.99 mm throughout the recording (an increase of 2.87 mm / second). Figure 5-27 shows the distribution of instability, highlighting user experience visually. A trend seems to be emerging with a higher concentration of high path lengths for novice user’s post-test.

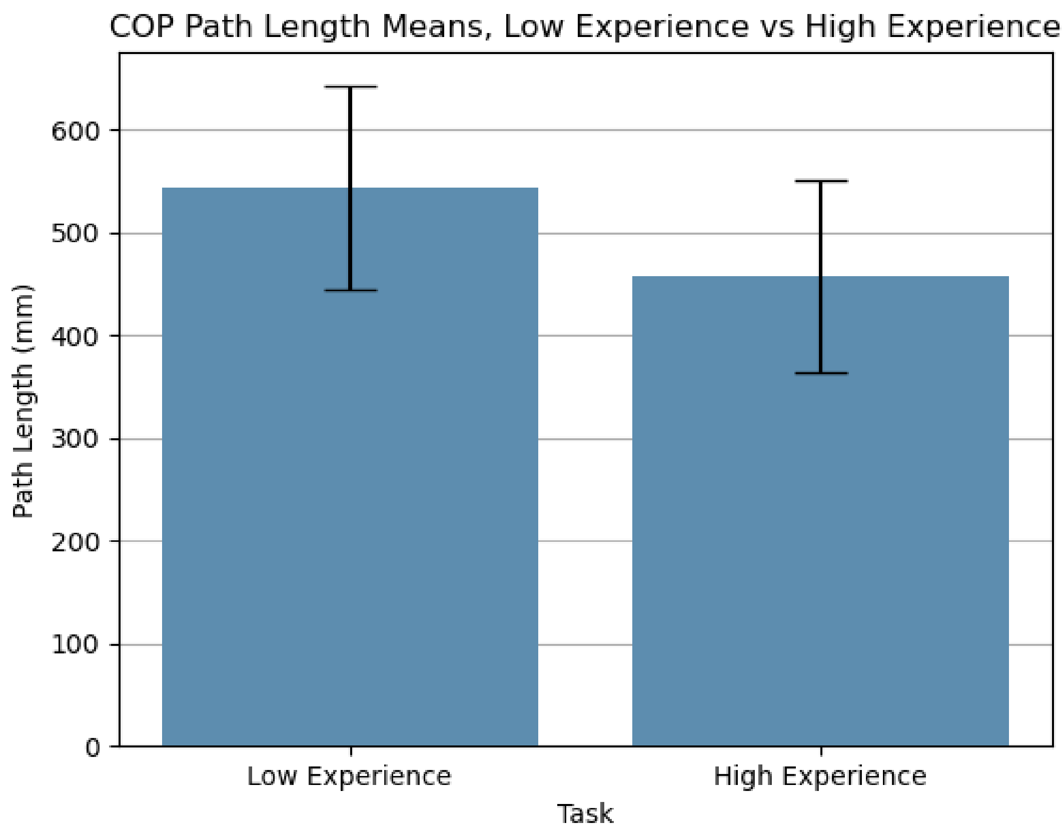


Figure 5-26:- Mean COP Path Length Difference, Low Experience Vs High Experience (Low Experience Mean = 543.89 mm, STD = 99.13 mm, High Experience Mean = 457.90 mm, STD = 93.077 mm)

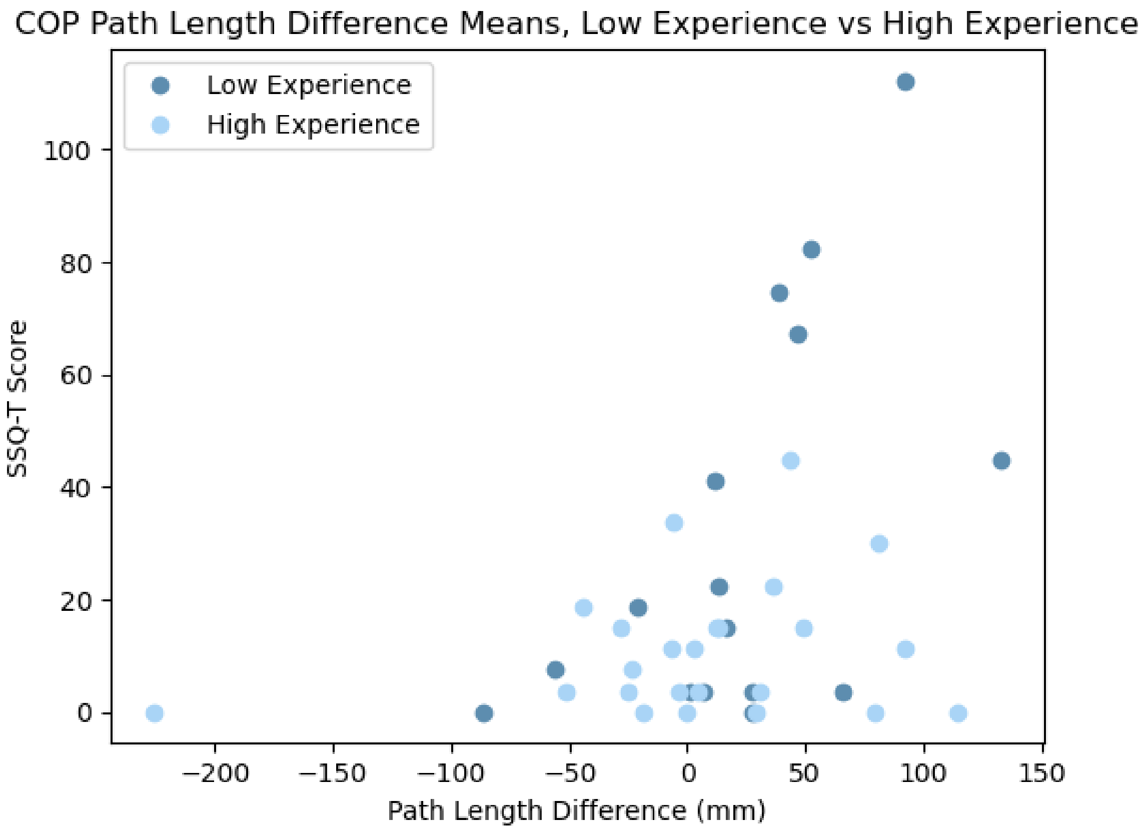


Figure 5-27:- COP Path Length Difference, Low Experience vs High Experience Subjects

5.2.3.4 COP calculation methods

5.2.3.4.1 Zero Crosses

Error! Reference source not found. shows a consistent pattern of motion in each axis between the two testing scenarios suggesting the frequency with which subjects corrected their posture during each test was largely the same.

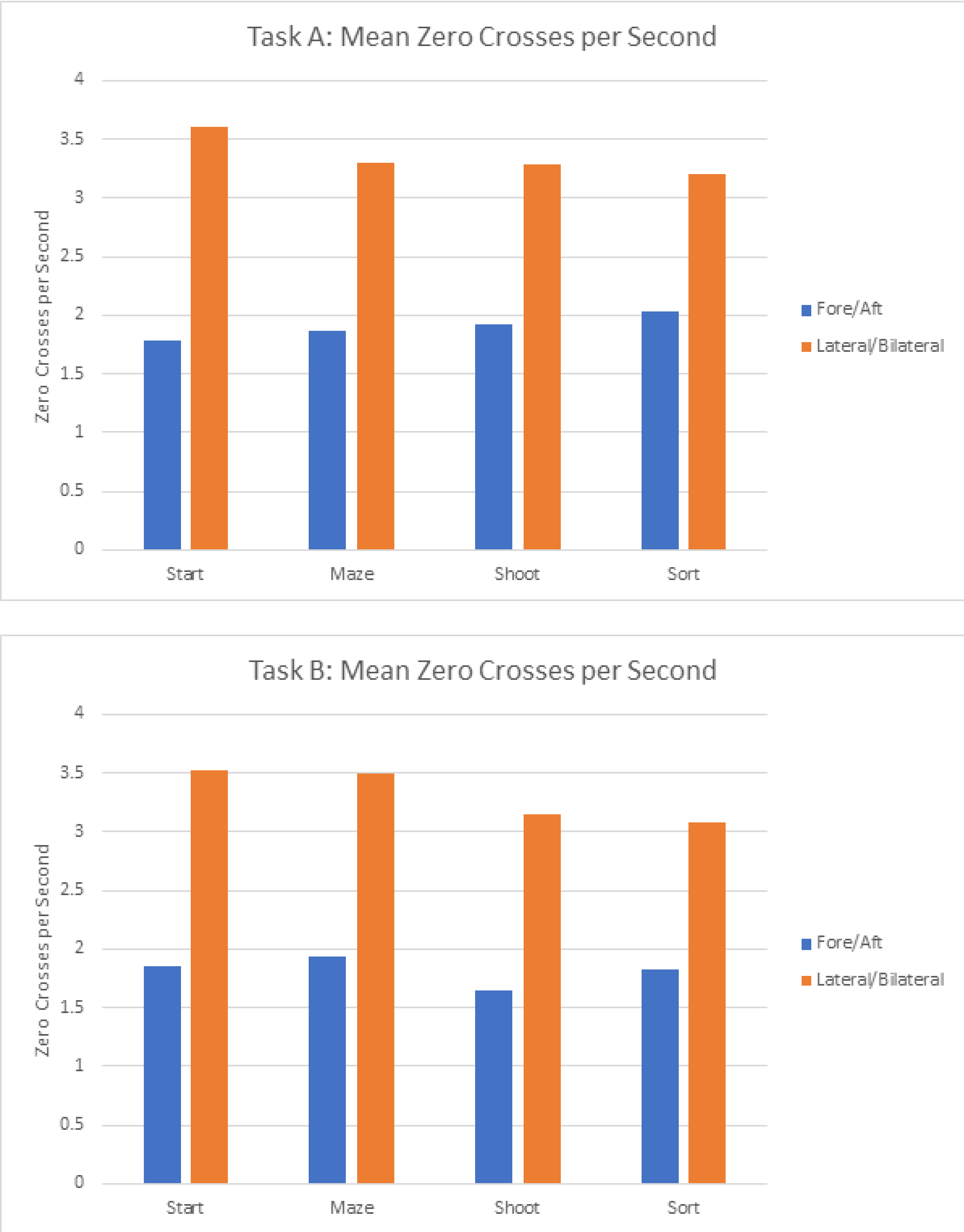


Figure 5-28:- Zero Crosses per Second for anterior/posterior and Lateral/Bilateral Motion Axis

This result means the frequency with which a subject corrects their position has no bearing on the magnitude of instability they acquire. It is more likely that the increase in the magnitude of positional and rotational motion measurements suggests instability occurs because of a series of overcorrections made by the subject attempting to manage their balance. A significant difference in anterior/posterior and Lateral/Bilateral correction exists, which is interesting as lateral bilateral data showed significantly less total motion than anterior/posterior. This result suggests more corrections are required in the medial/lateral axis, suggesting either balance is easier to maintain in this axis or lower tolerance to instability in this axis requiring additional corrective effort to maintain.

A similar pattern occurs in the rotational data with a greater number of major corrective measures in the roll axis than the pitch or yaw axis. This pattern matches the axis exhibiting the lowest motion amount. However, the significant difference in the amount of motion between the pitch and yaw axis in is not found in the zero crosses for the other two axes, with them being almost identical. This finding suggests corrections in these two axes may be connected, and further investigation is undoubtedly worthwhile.

Next, the results were split further into sick and well groups using the same categories as the previous positional instability analysis (sick participants SSQ-T >= 20, well participants < 20). Examining Figure 5-29, we see significant differences in the number of crosses between sick and well participants. While this would suggest a significant difference between the two states, this is unlikely to be the case. The number of sick samples is very small (5 for Test A and 7 for Test B), and thus the influence of outliers is likely to be high. However, the possibility exists that some difference does exist between these subjects. However, examining these graphs grants no insight into any patterns that may be emerging. Sick users *may* make more major corrections to head position, but the increase is likely minimal to the point of insignificance if they do.

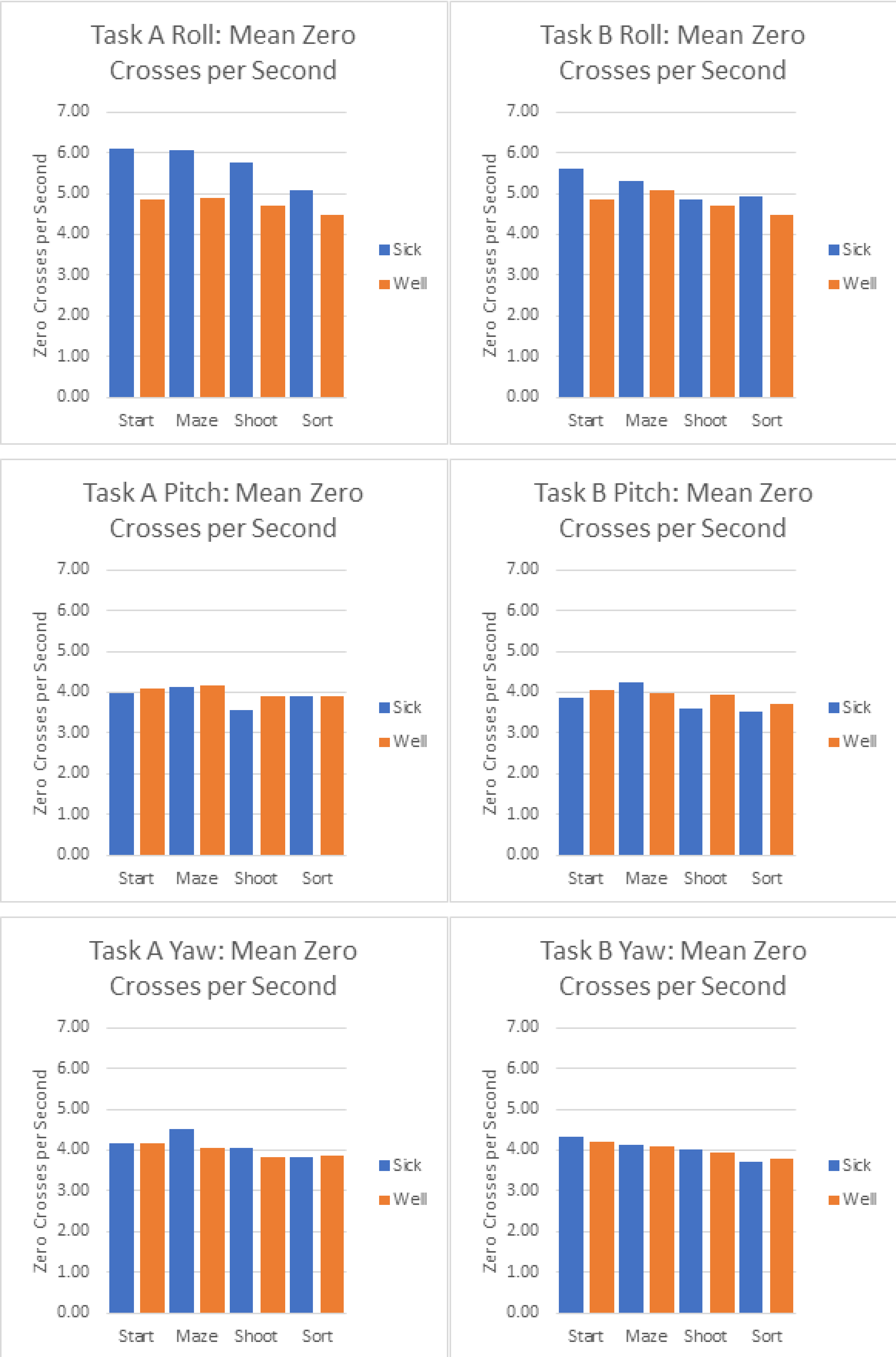


Figure 5-29:- Zero Crosses per Second: Roll, Pitch and Yaw Motion Sick vs Well Participants.

Figure 5-29 shows the number of zero crosses for each rotational axis for sick and well participants. Evidence of a general downward trend in the roll measurement is present in both Task A and Task B results. The trend is significantly more pronounced in sick participants than well participants, especially under Test A Condition. The significant difference in initial sick and well measurements suggests the number of major corrections reduces as the subject gains experience with the environment. This observation is hard to interpret. Assuming users get progressively sicker as the test goes on, the expected results would show divergence between the two sets of results rather than convergence. Something is occurring here, but interpreting its exact meaning is difficult with the current dataset. Monitoring sickness scores after each test would yield further insight here.

Looking at the impact of user experience, little difference in the values for either the anterior/posterior or lateral/bilateral axis. Low experience users present slightly more crosses per second than high experience users in all but one case. This observation matches the general trend that inexperienced users generate higher amounts of instability, which manifests as more major corrections in positional motion. A similar decrese in roll motion is observed similar to those found in Figure 5-29. This result draws similar conclusions to those that major corrections in roll decrease as the test progressed. Similarities in the results are likely due to cross-over between the two groups, with two-thirds of sick users identified as novice users. This issue suggests the findings relating to sick users (the smaller of the two populations) may be a function of user experience. Support for this conclusion comes from the observation that the difference in motion between these two groups converges towards the test's end. With less difference in motion observed between these two groups towards the end of the test, this suggests subjects learn the environment and adapt to it. However, with insufficient results in the number of experienced users getting sick, separating the two factors will prove tricky.

Finally, examining the relationship between first and second runs where the most significant difference in instability was detected earlier in the report, shows little difference in the number of zero crosses exhibited between runs, certainly not to the magnitude exhibited during earlier analysis. This result suggests that subject's make a similar number of major corrections across repeated runs within the same test environment. The rotational measurements show a similar pattern.

5.2.3.4.2 Velocity Changes

The other analysis method, velocity changes, aims to identify if small subconscious changes in the direction of velocity, speeding up and slowing down the subjects' motion indicate any changes in the subject's position or viewing direction. Figure 5-30 shows a minimal difference in Task A and Task B's direction changes and between Anterior/Posterior and Lateral/Bilateral motion. The rotational data shown in Figure 5-30 shows a similar pattern to that of the zero crosses data, with the roll axis showing more velocity changes than the pitch and yaw suggesting more minute corrections in that axis, again suggesting a lower tolerance for instability in that axis than others.

There seems to be no significant difference between the test conditions of high and low-experienced users. No observable differences between Task A and Task B testing scenarios can be found, with virtually identical trends in both scenarios. Neither sorting subjects by sick and well categories or test order yields any significant deviation from the pattern or the alternative test scenario in each case.

Analysis of these results leads to two possible conclusions. Either the number of velocity changes recorded during testing gives little indication as to the amount of sickness suffered by a subject, the level of user experience, or the test order. Alternatively, this test scenario looks at changes in direction velocity. It is highly susceptible to noise in the headset's reported position and rotation, and likely, what changes are observed here are simply noise in the recorded results. Adjusting the rolling average used to smooth out the results may yield more valuable results. However, increasing the number of samples covered by each step of the rolling average will reduce the results' sensitivity making finding any useful information unlikely. In either case, it seems that this approach to instability analysis is not effective without the usage of signal processing techniques outside the scope of this project.

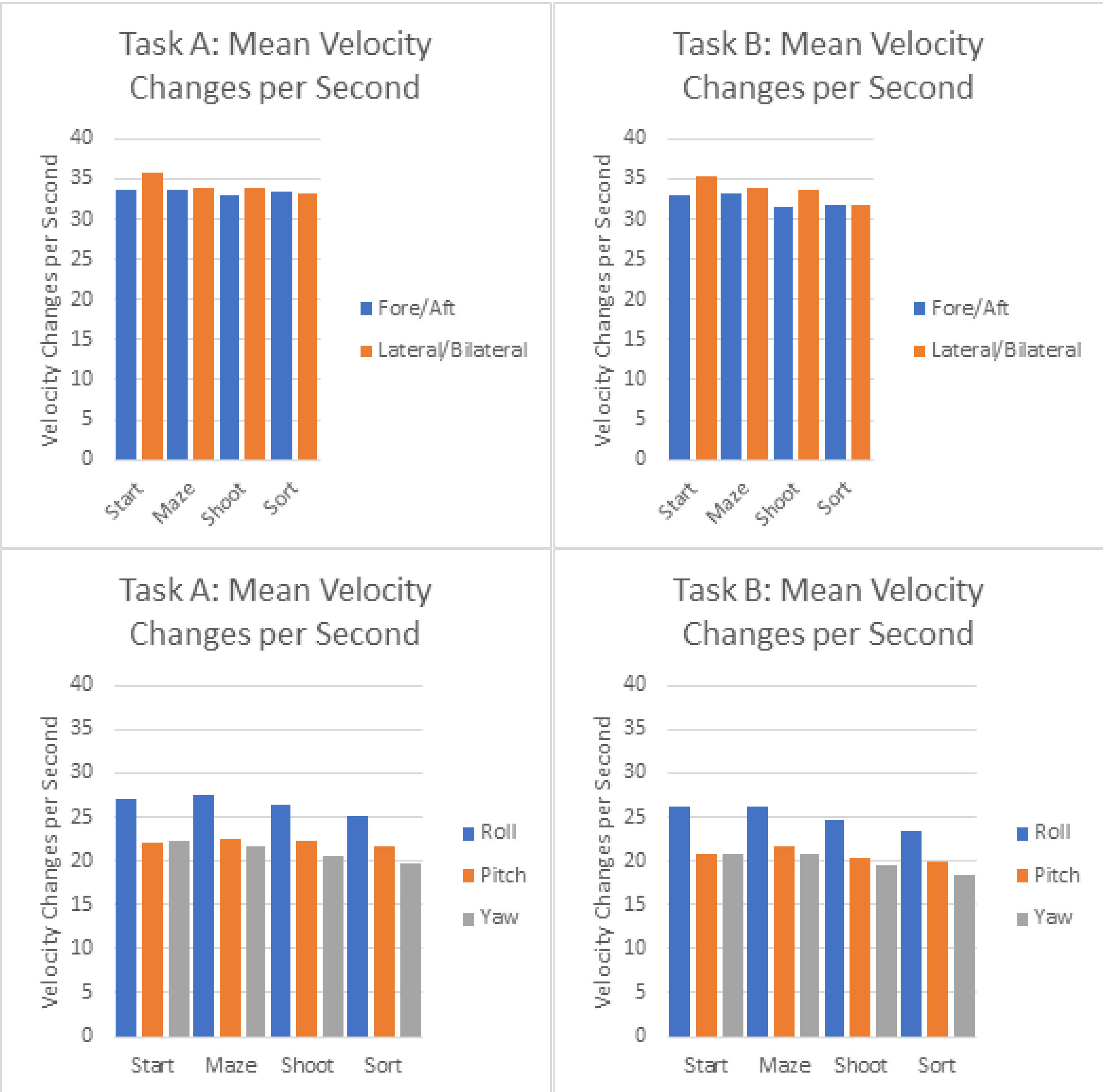


Figure 5-30:- Mean Positional Velocity Changes per Second.

5.2.3.5 Review

These results confirm the theory that task load does influence instability within a VR HMD environment. However, the task load's impact is not uniform, although a higher spatial load generally leads to higher instability. This result means that when comparing the results of any experiments utilizing instability as a metric for determining if a subject is suffering from cybersickness, it is essential to identify any differences in the task's complexity and mental load to ensure the validity of the comparison. The method used here differs from the more commonly used balance board measurement. Considering motion affects balance, performing a confirmation study will establish if these findings apply to that methodology. It is clear, though, that the instability measurement captured in this test is affected by task load. As such, this influence makes it impossible to attribute increases in instability to cybersickness wholly. The likely response would be a need to configure the threshold for instability on a per-task basis rather than having a static threshold of motion for determining sickness. As such this hypothesis can be considered to be confirmed for the purposes of this work.

6 Discussion

The study conducted here aims to identify the suitability of using measurements taken directly from a VR HMD device to detect cybersickness in subjects. The theory is that subconscious changes in VR HMD position rotation could approximate an instability measurement similar to using a balance board to record COP measurements in studies (Arcioni et al., 2019; Risi and Palmisano, 2019a; Widdowson et al., 2019). Replacing the COP method is desirable due to the unsuitability of using the pressure plate to test most use cases for VR HMD. Most VR activities allow the user to move around the environment physically; the pressure plate's physical properties restrict the play space to the plate's surface area, generally requiring the user to maintain a static foot position. The pressure plate also introduces a dangerous trip hazard into the environment, as the user entirely blind to the real world is a poor choice for safety.

The method proposed in this study would be more appropriate for monitoring the user in VR HMD environments. This data capture method poses no restriction on users' physical motion within the play space by utilizing the inbuilt measurements for headset position and rotation. As an additional benefit, the method requires no additional equipment, such as the balance plate. It could potentially be recorded externally to the VR experience, potentially allowing data capture in any VR play space.

The second study objective aims to identify the contribution to instability for VR HMD from differing cognitive loads. The reasoning for investigating this was the lack of consistency in VR HMD stimulation used in cybersickness studies. Studies generally follow one of two paths, either using an off-the-shelf VR experience, which usually is a very provocative experience, such as a rollercoaster ride. The alternative option is to create a bespoke environment, usually developed when considering the impact of a specific factor on cybersickness. In both cases, the level of engagement with the environment differs from scenario to scenario, and therefore the level of cognitive load differs.

An additional theory was that the task's additional cognitive load would increase instability. The task would occupy more of the subjects working memory and leave less mental capacity to process the environment. Identifying the impact of cognitive load on instability in VR HMDs is essential in this case to prove if increases in the movement recorded are being induced as a response to cybersickness or as a response to additional cognitive load within the environment.

6.1 Question 1: Does an increase in cognitive load within a virtual reality task influence the severity of cybersickness symptoms?

During the experimentation, subjects experienced different forms of cognitive load (memory, spatial and overload). Each load condition seems to have had a different impact on the stability measurements recorded for each participant. The maze task, which induced a memory load onto the subject, seems to have reduced instability in subjects. This finding suggests that increasing memory resources required to perform a task in VR HMD environments increases stability without affecting sickness scores. Spatial and overload tasks, however, seem to increase the instability of the subject. Additionally, all the high load scenarios (Test B) showed increased instability over the low load scenarios (Test A).

These findings suggest that our initial theory that cognitive load impacts postural stability in VR HMD environments does seem to be accurate, with the possibility that different load types have different impacts in different ways. This finding would mean the potential explanation given as to the conflicting results from studies may be down to the cognitive load of the individual tasks rather than the onset of cybersickness, or the differing loads of the tasks may be polluting the data during study comparison. The results presented here are not wholly conclusive. Attributing all of the effects to testing order or subject fatigue is possible, mainly because the load condition was randomised (Task A, Task B), and the load types were not (maze, shoot, and sort).

As shown in Table 4-1 the validation of hypothesis 1 was required to validate this research question, as this hypothesis has been confirmed and using the above rationale we can confirm this research question. As such this research question can be considered to be true within the limitations defined in this study.

6.2 Question 2: Does an increase in the cognitive load experienced by a subject within a virtual reality environment impact postural stability measurement?

This study aims to confirm or refute the plausibility of using the headset measurements from a VR HMD to detect cybersickness. Mean COP path length measurements recorded via the Wii balance board showed a slight, non-significant increase in path length during the low load task from start to finish. The mean high load path length measurements were even from start to finish with no observable change. Mean ending measurements for both tasks were even. The final ending result suggests no noticeable impact from the task load condition on stability measurements made in this way.

Further analysis attempted to identify a metric capable of separating users into sick and well participants based on their COP measurements and their responses to the SSQ. The analysis revealed no correlation between total COP path length and SSQ-T scores, nor any correlation between the difference in starting and ending COP measurements and SSQ-T results. Similar comparisons to SSQ-O, SSQ-N and SSQ-D scores showed similar results.

The results collected here do not correlate with the results collected in Widdowson et al. (2019); this is probably due to the different scenarios under which the data collection occurred. Data capture for Widdowson et al. (2019) occurred during VR exposure, whereas data collection for this study occurred before and after VR exposure. The plausibility of performing this measurement during VR immersion was deemed implausible due to safety concerns as excessive close contact would be required (forbidden due to covid-19 concerns) to introduce the board for measurement and guide the subject onto it safely. It appears the instability acquired from the VR environment dissipates very quickly and as such, capturing the data post-exposure is not applicable.

However, it is essential to acknowledge that, while the results gathered here show no significant change, the results likely would have been much more significant had the COP test been conducted under VR immersion. As such, this study does not make any claims regarding the effectiveness of this method. However, conducting this test as a pre and post-test measurement outside of the virtual environment seems to be mostly ineffective as a method of establishing sickness susceptibility.

The second part of this test was capturing VR HMD positional data during four quiet stand periods. Measurements were made once at the start of the test and immediately after completing each test activity. The initial analysis of mean values seemed to suggest a correlation between increased positional motion and SSQ-T scores, with sick participants gaining significantly more motion from the initial measurements to the final measurements made post-sorting task. While the high load task observed a more significant increase than the low load task, mean increases were significant for both task conditions.

As the sample size was small for this test (24 samples), and the standard deviation was high. Further analysis of the results was appropriate, attempting to validate the appropriateness of deriving a model based upon the ratio of increase from start to finish in instability measurements. Analysing individual results showed that few individual results correlated to this model for sick participants. Also, the final recorded magnitude of instability on an individual basis did not follow the trend suggested by the mean analysis; the instability measurements recorded for sick participants fell almost entirely within the range of non-sick participants. The conclusion drawn from these results was that neither the ratio of instability gain nor the absolute magnitude of instability recorded in this manner was suitable for identifying a participant as sick or well.

Additional analysis of other factors revealed other potential sources which contribute to instability in VR HMD motion. As discussed earlier, task load impacts HMD instability, with spatial loads adding the most instability, and the possibility exists that cognitive overload in the subject has a similar effect. Surprisingly, VR usage frequency did not significantly affect instability, with experienced VR subjects exhibiting slightly higher instability levels than novice users. The majority of subjects recorded reduced motion in the second test than the first, regardless of the test's task load.

Individual axis analysis showed the most significant increase in motion on the forward/aft axis, followed by the lateral/bilateral axis. The head rotation showed pitch to be the most mobile followed by yaw then roll. This study differs from previous studies (Arcioni et al., 2019; Dennison and D'Zmura, 2017; Rebenitsch and Quinby, 2019) due to placing no restrictions on VR usage during this test. Previous studies have greatly restricted player movement throughout the environment.

In contrast, this study captures the data during normal VR usage, a more realistic model for actual VR usage, and revealing more information about the practicality of applying this method in the real world. Studies by Dennison and D'Zmura, (2017) and Rebenitsch and Quinby, (2019) also, do not consider the reality of VR usage. The results gathered here show an influence on instability from multiple other sources such as task load and user experience, both from general familiarity with VR and multiple exposures to the same VR experience.

As shown in section Table 4-1 the validation of hypothesis 2 and 3 was required to validate this research question, as this hypothesis has been confirmed and using the above rationale we can confirm this research question. As such this research question can be considered to be true within the limitations defined in this study.

6.3 Question 3: Can the alteration of task performance be used to identify the onset and severity of cybersickness in an individual?

Data analysis yielded no meaningful evidence of a correlation between perceived task performance and SSQ-T scores. The perceived load of tasks experienced by the subject did not seem to influence the final sickness scores of the subject. However, analysis has uncovered a potential avenue for exploration. As SSQ-T was only gathered as a post-test measurement, it is impossible to identify the post-task impact of individual tasks (maze, shoot and sort) on sickness scores. The possibility exists for peaks and troughs throughout the experiment that our data did not capture. Repeating the study with multiple data capture points for SSQ-T data post-task would likely yield interesting results.

As shown in section Table 4-1 the validation of hypotheses 3 was required to validate this research question, as these hypotheses have been partially confirmed and using the above rationale this research question is false at the current time within the limitations defined in this study.

6.4 Question 4: Can the onset of cybersickness be practically identified mid-task using commercial off-the-shelf VR equipment?

This study has shown that approximate measurements for instability can be derived from VR HMD positional and orientation measurements, but several major factors limit the effectiveness of this approach. Other significant factors were found to impact stability measurements, with user experience being the largest contributor. Thus the factor of instability cannot be wholly relied on as an indicator of sickness levels. The levels of instability captured pre and post-test do follow the findings presented in other literature, but the magnitude of these results is significantly lower. Probably due to capturing this data outside the VR environment.

Additionally, the data captured from the headset is extremely noisy. While the literature identified a likely significant drop in tracking accuracy compared to the existing methods, the data captured proved much more challenging to work with than anticipated. Issues relating to tracking the headset, such as redetecting the ground plane, loss of tracking and axis flipping of the recording data, led to a significant amount of data manipulation being required. This increase in manipulation limits the plausibility of this approach in a real-time scenario, particularly in demanding applications where performance is already stretched thin. Additionally, as these tests were performed under laboratory conditions, this recording environment likely represents a best-case scenario, and real-world data sets are likely to be even more inconsistent than the ones gathered here.

Additionally, the study identified a significant issue while capturing the results. While most external tracking systems identify the head as the tracked point, the VR HMD measurements are taken as the device's position. In layman's terms, the VR HMD is not recording the subject's head position but a projection of a position in front of the user. This fact means that rotations of the user's head significantly impact the position recorder via the HMD, polluting the positional data with orientation data and visa versa.

Finally, considering the practicality of this approach, we find that the quiet stand measurement in its current form is entirely unsuitable for the task in a practical setting. Without adding complex solutions into the environment (such as lift rides during level transitions in a games environment), having a user stand still in a known pose, which is unlikely to be performed correctly, is majorly obtrusive to the workflow. Additionally, suppose a user is suffering from cybersickness. In that case, they would likely remove the headset in response to being asked to create a measurement to determine if they are sick.

Overall, these factors combined, with instability being shown to have a more significant contribution from external factors than cybersickness, the noise and pollution of the collected data, and the sheer impracticality of the data collection method. This study concludes that this method is too impractical to deploy in a real-life scenario.

As shown in section Table 4-1 the validation of hypotheses 2 and 3 were required to validate this research question, as these hypotheses have not been confirmed and using the above rationale this research question is false within the limitations defined in this study.

6.5 Potential issues:

6.5.1 Issues with positional tracking in commercial HMD units

This data capture method does have some issues, primarily, the expected loss of accuracy in the recording. The test device chosen for the experiment (HTC VIVE) has some minor issues with tracking accuracy identified by Niehorster et al. (2017). However, this study's goal was not to obtain entirely accurate head position and rotation measurements, but rather to identify if the measurements made by the headset would be sufficient to identify motion in the subject related to the onset of cybersickness. Rebenitsch and Quinby, (2019) identified an increase in headset motion during the onset of cybersickness symptoms. If the technique developed here works correctly, some correlation of results should be evident.

Results identified a significant misconception when attempting this approach; the assumption made is that the headset positional data generated represents the user's head position. Testing revealed this not to be the case. The position recorded represents the position of the headset within the real world. This small but subtle difference has a significant impact on the results, as essentially, the result recorded is a projected position in front of the user's face. As the user rotates their head, the result of this is that the headset's position also changes. Therefore, using the HMD positional data for reporting the head's exact position is inappropriate as the HMD recorded position is being polluted by the user's head's rotational movement, Potentially this issue could be corrected by projecting the positional measurement back to the centre of the head. Achieving this would require knowing the distance from the headset measurement point to the centre of the head. Insufficient data was collected to test the validity of this theory.

It is worth discussing the application of these findings to the broader range of VR HMD available on the market. The headset used during this experiment, the HTC Vive, is last-generation technology. As such, significant advances are likely to have been made in terms of these headsets' tracking accuracy. However, the only way to be confident in any headset's base accuracy is to conduct a study in line with Niehorster et al. (2017). Performing this is time-consuming and expensive, even more so considering multiple headsets are required to verify the results thoroughly. Considering the turnaround for results publication and the pace at which new headsets are coming to market, it is unlikely that detailed accuracy data for current-generation headsets to be available during the headset's primary lifespan. The most significant difference is likely to come from the headset's tracking methodology. The methods for tracking the headset's position are significantly different between headsets, and as such, results may vary on other brands and models of VR HMD

6.5.2 Headset Data capture issues

The experiment collects four sets of stability data throughout the experiment. A significant problem exists in user orientation and translation within the world environment, as during a standing interactive simulation, some movement is inevitable. Additionally, as users were free to rotate within the play space, user orientation was not static. This issue presents the question of the accuracy of recording the positional data across the play space. Niehorster et al. (2017) assessed the feasibility of utilising the HTC Vive for scientific research, providing a generally positive review, citing fast-tracking with low noise. However, the paper does cite a few issues that require consideration.

Niehorster et al. (2017) identifies that the ground plane by which the device detects the floor location, referred to as the ground plane, does not match the location and orientation of the actual floor in real space. Recalculation of the ground plane also occurs each time the headset loses and regains tracking. This issue is hugely problematic as the mismatch between the virtual ground plane and the actual physical floor plane will introduce a random offset into the measurement of roll pitch and yaw representations within the environment. This issue introduces a unique problem to this experimental setup. This experiment does not follow the common model of having the user face a specific real-world direction within the environment. Therefore, the axis of motion recorded by the headset is not consistent, nor does it line up with the cardinal directions of the tracking space. This issue means a rotation of the co-ordinate space is needed to establish the movement's direction relative to the subject. This experiment captured rotational data in the form of Euler angles (roll, pitch, yaw). Therefore, to determine the subject's axis movement, the positional co-ordinates captured need to be rotated by the inverse of the recorded angles to match them to the user’s orientation. This process would match the users' motion to the virtual ground plane, practically lining the X, Y and Z axis to the virtual world in a static direction.

This approach results in a slight mismatch with the headset's actual real-space representation based on the ground plane's offset. This solution inevitably will have a slight effect on the accuracy of the data used, establishing the direction of recorded motion. Niehorster et al. (2017) suggests the approach of calibrating for this in short term experiments by recording the ground plane at the start of the experiment and using that to identify the offset, and then operating the headset within areas of the space that are less likely to have tracking issues. While this approach is not flawless, it does minimise the risk of these issues impacting the results.

Unfortunately, this approach is limited in its applicability to real-world VR HMD usage. Very few VR HMD setups are perfect in their implementation Niehorster et al. (2017) shows tracking dead spots when the lighthouses placing the lighthouses in the manufacturers recommended positions and a difference in tracking quality based upon distance from each of the lighthouses. Most room-scale VR setups utilise space beyond the optimal tracking range, and very few environments have the user in a static position for the entire duration. Even when restricting user position to the optimal tracking space, tracking can occasionally be lost, resulting in creating a new ground plane. Niehorster et al. (2017) shows reasonably accurate tracking over a one-minute recording representing a short duration compared to actual VR usage, in which sessions can commonly last 10 minutes or more.

Recalibrating for the ground plane requires having the headset in a known position. The difference between this known position and the physical floor allows for calculating the ground plane's position. However, this process needs to repeat each time tracking is lost to accommodate the software's new ground plane. As the possibility of a loss of tracking can be minimised but not removed, and this tracking loss occurs unpredictably. Therefore, some incidents of tracking loss will occur during normal VR usage, mid-way through activities.

As the calibration process needs the headset to be in a known position, remaking these calibration measurements on the fly is impossible. While detecting these breaks in tracking is trivial as the headset reports a wildly anomalous or precisely identical position value. Recalibration without removing the headset requires getting the headset into the know position while still on a user's head. Achieving this with any accuracy or reliability, practically speaking, is just not feasible. Even attempting to hold perfectly still, a person continually makes slight motions due to natural instability as the various muscle groups counteract external forces such as gravity (Soames and Palastanga, 2018). However, taking the headset off to reperform calibration would break user immersion. While the temporary loss of tracking breaks user immersion temporarily, removing the headset removes the user from the VR experience impacting cybersickness onset by reducing presence (Risi and Palmisano, 2019a).

It is essential to question, does this matter in the context of the study's end goal? The aim is to establish whether the VR headset can capture measurements suitable for identifying a subject suffering from cybersickness during normal VR usage. The ground plane issue will have some impact on the accuracy of the results recorded by the system. The questions asked by this research does not demand resolution of the issue but more identifying it as a limitation of the data capture method. Using the best positional measurements afforded by the hardware during the actual usage of VR HMD, can we still detect an increase in instability through these measurements?

The virtual ground plane will impact the motion recorded by the headset, and an exact axis alignment to the user will not be possible with this system. Therefore, some noise and inaccuracy will exist in the anterior, posterior, lateral, and vertical motion captured. However, these values may still prove sufficient to establish instability in those directions, despite the offset in rotation added by the ground plane's inconsistent angle. Recording the headset positional data for the quiet stand tests should be largely unimpacted by differences in the ground plane. Niehorster et al. (2017) established inaccuracies in the headset's recorded position were minimal over the playfield. However, the measurements made in this study are not made from a static position and will likely have significantly more variance than those found in Niehorster et al. (2017). Moment to moment, comparisons of rotational data will have the same offset (unless directly impacted by a loss of tracking), and thus no impact should be detectable in these results.

It is also worth noting that more modern VR HMD hardware does exist than those used in this study. The HTC Vive devices used in this test are approximately five years old. Ideally, the study would utilise more modern headset technology, such as the HTC Vive Pro (https://www.vive.com/uk/product/vive-pro/). More modern hardware would likely minimise these tracking issues. Alternatively, tracking accuracy issues may be addressed by changing the tracking system to ones used by systems such as the Oculus Rift S (https://www.oculus.com/rift-s/). However, this will likely introduce their own set of problems that would require appraisal and investigation, such as in Bauer et al. (2021). While the university does have a wide range of different headsets, only the HTC Vive was available in quantities large enough to support the headset rotation policy required by the COVID-19 safety precautions.

Another vital concept to clear up is that the headset is not recording the users head position, more the headset's position. Depending on the headset, this may be either the internal displays' position or a virtual calculation of the head position based on the headset's recorded measurements. In either case, these measurements represent a projected position forward of the user’s actual head position.

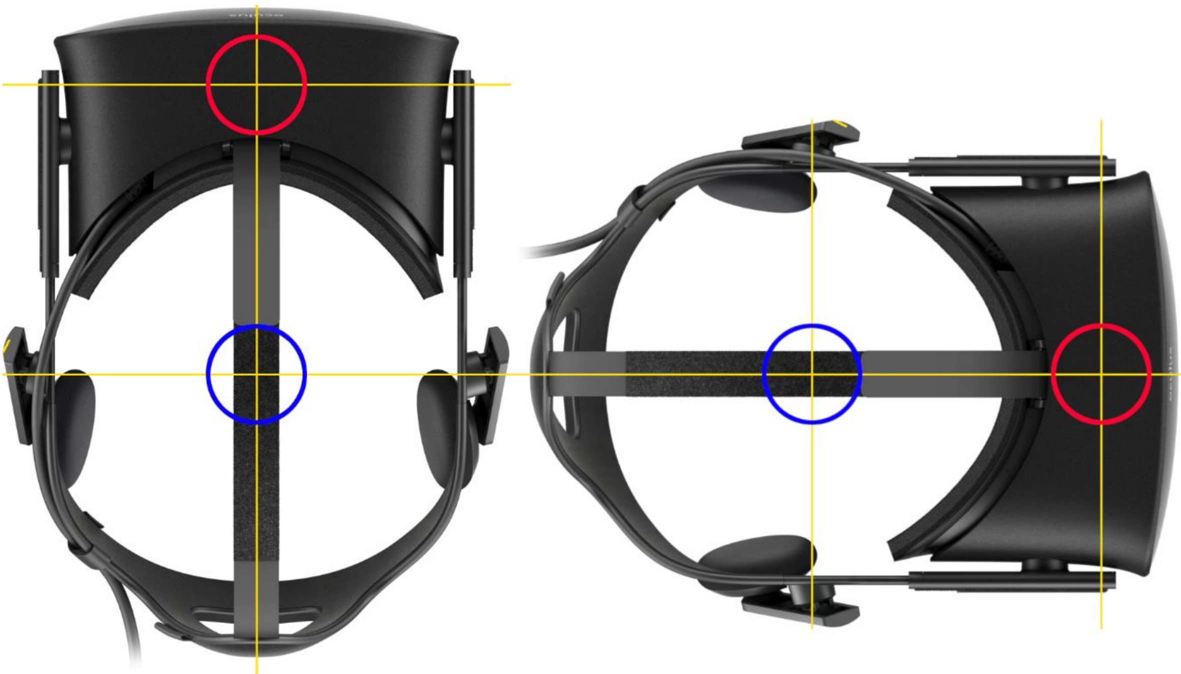


Figure 6-1:- Illustration of Headset Rotation Problem. The blue circles represent the head position, red circles represent HMD position. The diagram shows how rotating the headset can change the HMD position without a corresponding head position change.

Figure 6-1 illustrates the problem this introduces. Assuming the blue circle represents the head position, and the red circle represents the HMD position. As the user changes viewing direction, the headset rotates around the head position. Minimal changes in the user's head position will occur during this process. However, by contrast, the HMD movement is much more significant as the headset's position is projected from the point of rotation. This issue is not entirely unexpected as reporting the head's position, in this case, does not make sense. The software needs to know where the user's viewport is, not where the head's centre is, as this would require additional information about the user to establish.

It is theoretically possible to derive the centre of the head's position through a couple of methods. Firstly, if the exact point at which the headset takes its measurements is known, and the central point of rotation of the head is known, then a simple calculation could be made to establish the position of the head relative to the headset. This approach has some significant flaws in its implementation. Firstly, the system requires a biological measurement for each subject utilising it, which is tricky to obtain due to significant head size and shape variance. Measurement devices do exist for this purpose but are incredibly uncommon outside of specific scenarios requiring its usage. However, this measurement still does not give an accurate representation of the head's centre of rotation. This measurement will vary from individual to individual concerning the head's front and back measurements.

Another approach may be a possibility post-test, as we know the headset's position and its orientation. Theoretically, it should be possible to cast a vector back towards the head's centre over multiple frames, with the point of intersection of all these lines being the central pivot point of the head. Unfortunately, this solution assumes a perfect measurement of the headset position from frame to frame. Any inaccuracies will result in unpredictable variance across the vector intersection points. This approach also fails to accommodate for physical movements of the person during the capture of the data. While this could approximate a head's position over time, another question needs to answer whether this is a worthwhile practice? By calculating the head's central position, a better representation of the users' position within the environment may be achievable; however, the approaches outlined here add additional noise to the results, which reduce the reliability of the positional data and the potential to detect instability. Both approaches also add additional complexity to the acquisition of results, either requiring additional measurements during the setup phase, which would be difficult to obtain outside of a laboratory setting or require additional calculations on the measurements during data usage. This issue would not be a problem for establishing post-test instability. However, if attempting the method in real-time, this complexity would add additional load onto the system when the additional load is undesirable.

In conclusion, while acknowledging the value of knowing the position of the user’s head in this scenario would be valuable. The additional resources required to achieve this, and the potential additional variance added in this approach makes this approach unfeasible for usage in a real-world setting. Therefore, despite the noise and variance added to headset positional data, it is believed that using the raw HMD position best represents an approximation for the subject’s position within the environment, which is achievable in real-time without excessive, intrusive measurements of the subject.

6.5.3 Instability and User Experience

Comparing subjects of different experience levels is fascinating; the effects of learning in VR are well known. Taylor et al. (2011) shows that repeated exposures to the same stimuli reduce cybersickness symptoms. However, these results seem to suggest that this effect may not be transferable between different VR HMD setups or different VR environments. If instability is not an indicator of sickness, then maybe it is an indicator of acceptance of the sensory conflict. This theory would certainly explain the mismatch in instability measurements from the first run to the second run. Possibly, the problem may be in the way the brain processes the new environment. In the physical world, resistance to unusual environments, such as low gravity, can be achieved by repeated exposure. For example, the vomit comet (Nola Taylor Redd, 2017) exposes potential space travellers to a zero-gravity experience. Aptly named, one of the side effects of the experience is severe nausea. As the brain attempts to make sense of the new environment, repeated exposure to this stimulus results in tolerance to the stimuli and thus resistance to the sickness's effect.

To investigate this further, proposing a theory about the cause of this decrease in instability and why it does not cross between VR setups. That theory is that each unique combination of VR HMD environment and VR experience creates a unique scenario of mismatched sensory inputs. The brain cannot make the association between these environments as each experience is unique, and as such, the brain creates a model to deal with the environment, not VR in general. An investigation into this theory is required to determine its accuracy and represents further work for the project.

6.5.4 Alternative Data Capture Methods

The study has also noted several issues relating to the practicality of implementing the quiet stand method of detecting cybersickness. The amount of time needed to make a recording and the necessity to have the subject observe a specific posture makes the chance of getting a consistent measurement unlikely. While it is probably not necessary to take the whole 30 seconds to make a recording, and a shorter duration would probably suffice, the method still requires removing the user from their task to perform. Finally, the purpose of this method is to attempt to detect cybersickness before symptoms emerge. Achieving this would require frequent measurements during the task, likely to disrupt immersion and users' workflow to an unacceptable degree.

Therefore, an alternative method is necessary. This study has theorised that it may be possible to obtain an approximate measurement for stability during activities where the user’s head will be stable. These, in theory, would be much easier to incorporate passively into a VR experience. Two possible scenarios where this could occur exist within the experimental environment. The first was a corridor walk within the Maze task when the subjects returned to the test's starting position. The second was during the shooting task when subjects shot the target.

Analysis of the amount of motion occurring during the corridor walk showed the prediction of a stable viewpoint to be wildly inaccurate. Huge variations existed in the recorded individual motions, some of which recorded lower instability than during the quiet stand. It is doubtful that subjects became more stable during these measurements as the stimulation from motion within the environment and the increase in activity should contribute to instability. Instead, it is far more likely that the motion recorded by the subject exceeded the amount of motion needed to have the result removed as deliberate motion. The instability recordings made during the shooting task seem to be more accurate, showing significantly less variation than those made during the corridor walk recording. However, comparing them to the motion recorded during the entire test without any filtering showed only a small difference in the magnitude. The consistency increase observed in this result compared to the corridor method attempted earlier is likely attributed to the volume of results recorded during this method being significantly higher and the lack of locomotion during recording. More results to obtain the measurement would result in a significantly smaller contribution from results with motion exceeding the filter amount before being declared as deliberate movement or an erroneous recording.

While this confirms the theory that taking a shot represented a point of head stability within the test, the increase in stability measurements is minimal; therefore, this approach offers seemingly little benefit over just analysing all of the test results. This analysis, however, is likely misleading. Most of the data recorded during this period included the user taking a shot or searching for a target, therefore likely showing a result representative of just shooting targets. In a more realistic scenario, the user would likely be shooting intermittently within the experience; thus, measuring the entire experience would likely yield significantly different results. Further research is required to confirm or refute this. While shooting does represent a possible case where some stability may be successfully measured, it represents a very niche area of VR experiences, limited to a subset of games. Therefore, profiling a wide range of VR experiences would allow further research to identify more suitable candidates for this method.

This approach's possible flaw is attempting to predetermine areas of low motion and limiting measurements to that period. Analysis of motion during the Sort task sought to take the opposite approach. By sampling the whole data set and using the filter to remove results with large amounts of motion, what is left over should indicate instability. This approach yielded the most consistent results, likely due to the increased sample size and more results to average over. Again, however, increases in instability do not correlate to either SSQ scores or variations in quiet stand measurements. Adjustments to the filter, relaxing it to accommodate more base motion induced by the in-task measurements, failed to improve the results' consistency.

The method's overall goal was to identify a user’s susceptibility to cybersickness via a passive VR HMD position measurement. In all cases measured, the investigation failed to identify a link between the magnitude of motion recorded and SSQ-T scores. In all cases, Comparing the in-task measurements to the quiet stand measurements shows no correlation in the magnitudes; a high quiet stand measurement does not translate to a high in task measurement and vice versa. Also, the magnitude of measurements was significantly higher than the quiet stand measurements suggesting these measurements are not an appropriate approximation of the quiet stand measurements. However, as neither data set correlates to cybersickness measurements either, it is not easy to identify the more appropriate method to pursue further study. As the COVID precautions equally impacted both methods, both methods require further study to confirm their effectiveness.

6.5.5 Experiment Observations

During the experiment, subjects, for the most part, completed the tests as expected. However, some users did present unexpected behaviour. One subject moved significantly further back from the target range to see more of the game at once in the shooting task. This act effectively reduced the amount of head movement required to complete the task. Head movement does induce vection into the environment. A known cause of cybersickness (Bonato et al., 2008) and a reduction in head movement may have contributed to a lower SSQ score for the individual. Many of the subjects attempted to pick up guns in both hands to go “akimbo” and shoot with both hands. As this occurred before the experiment started, subjects were advised not to do this. All subjects complied with this request. Some subjects changed their firing hand midway through the test, primarily due to the discomfort of keeping the arm extended for the test's duration. Some subjects carried the gun outside of the test environment, but no evidence exists to suggest this impacted the post-test postural measurement or TLX data entry points. Safeguards placed to prevent objects from other experiments from entering other tests prevented any objects from entering a different test environment.

The Sort test saw some users attempt to mitigate the task's workload; a couple of users attempted to mark which bin related to which shape by placing a shape next to the bin to mark it. This method proved ineffective during the high load task due to the task's pace of refreshing objects’ One subject attempted this strategy during the low load task but quickly dismissed it as unnecessary. In both cases, subjects returned to the expected activity pattern quickly after attempting these strategies.

A more typical strategy, particularly in task B as the task's pace became excessive, was to sort all instances of a single shape into the correct bin before moving on to the next. This approach seems to represent the user adapting to the task attempting to make it more manageable and reduce load, sacrificing overall performance for limited accuracy on a subset of the task. This observation suggests that the subjects were overloaded, and, in response to this, subjects employed strategies to reduce task load. This behaviour generally was only exhibited in task B and only during the closing moments of the test, where the pace and load of the task were at their most intense.

These factors clearly show that there is another factor to be considered which highlights the difference between the realistic model used and the more controlled studies restricting user motions during testing. The ways users interact with VR environments are almost infinite and can be unpredictable. As such not considering this in a realistic model will result in an inaccurate representation, therefore relying wholly upon controlled studies may not reveal the whole picture regarding cybersickness within VR environments.

Finally, three subjects withdrew from the experiment due to extreme nausea. All these withdrawals came during the maze test. Withdrawal is likely due to an intolerance of the locomotion method used in the testing process as sickness symptoms quickly progressed shortly after entering the maze. Of the three withdrawals, two subjects exhibited instability to the point that their safety became a significant concern, and the supervisor stopped the experiment. One subject completed the maze task and expressed a desire to stop due to extreme nausea. No other tests caused instability to this level or saw any subjects withdraw. This observation would seem to suggest that locomotion within the environment was a significant contributor to sickness symptoms.

6.5.6 Possible Issues/Limitations

Some potential issues do remain relating to this study. The lack of recording pre-test SSQ data for subjects means that the possibility that subjects entered the test with sickness symptoms exists. This issue means that the level of sickness recorded for the subject may not be induced entirely by the VR environment. However, the likelihood of any individual with significant symptoms starting the study is improbable. The collection of the study’s results occurred during a period when covid restrictions were relaxed in England but not removed entirely. As such, part of the screening process was to ask subjects if they had any symptoms of COVID-19 or if they felt unwell in general. Subjects were withdrawn from the study if they responded that they were unwell or had any contact with a possibly unwell person due to safety concerns. Therefore it would be reasonable to conclude that any subject entering the test would at least be reasonably well, and as a result, not score highly on the pre-test SSQ, limiting the potential impact on the results. Young et al. (2006) shows the likely effect of not performing the pre-test SSQ is to reduce the magnitude of SSQ-T score, but this effect is uniform and should not influence results distribution.

It is also apparent that the COVID-19 pandemic has had a less beneficial impact on this study. Starting with the obvious, forcing participants to wear masks during testing will impact subjects breathing and respiratory rate. As research suggests, changes in respiration rate may indicate cybersickness onset (Denise et al., 2009). It stands to reason that the two factors may be linked, and as such, affecting the rate of breathing for a subject may impact cybersickness. Another source of uncertainty may be the anxiety introduced by the pandemic. Anxiety will impact the subject's mental state, possibly introducing additional instability into the environment. It also may explain the drop instability between the first run and second runs of the experiment. As subjects get more comfortable with the experiment, they relax and get more stable; however, a significant difference would be observed in the starting values measurements if this were true.

Finally, the need to use multiple headsets for testing introduces the possibility for variation in results All the headsets used in this study have seen significant service in the VR lab within the university. All headsets used in the experiment showed no apparent defects during appraisal for usage; this said the appraisal process did uncover two headsets with significant flaws. One headset had scratches on the lenses; the second had a malfunctioning IPD sensor. Both issues would have significantly affected visual quality. While this process has detected prominent areas of inconsistency, passive, invisible inconsistency may exist. This issue does not represent a problem; this represents an issue the method must tolerate to be successfully applied. In a real-world scenario, every headset is different, and as such, the method must be robust enough to accommodate this. This issue is likely to be exasperated further when considering using different headset models, especially considering different tracking methods.

To conclude, the COVID-19 pandemic has had a significant impact on the study; the addition of safety equipment may have impacted the subjects' physical and mental states. This equipment was necessary to protect participants from exposure to covid, and as such, may have potentially impacted the validity of the study. Therefore, the results presented here require further scrutiny via a confirmation study to ensure their validity. However, as every subject was impacted evenly by the situation, the impact should be relatively uniform across all subjects. Therefore, the analysis of the results in this report is valid.

6.6 Comparison to Existing Solutions

The impact of overuse of Visual displays and computerised systems is well known (Health and Safety Executive, 2019), and while often temporary, over usage can lead to uncomfortable conditions such as headaches, dry eye syndrome, diplopia and blurred vision (Parihar et al., 2016; Rosenfield, 2011). Guidance is in place for acceptable usage limits (Health and Safety Executive, 2019) to prevent these conditions from occurring. However, enforcement of these guidelines is rare, nor do they account for an individual’s tolerances. The best help software seems to give is the occasional reminder to take a break (Figure 6-2). In VR, these protective systems are more critical than ever. While no substantiated evidence exists for VR HMD devices causing permanent eye damage, unsubstantiated reports exist (British Broadcasting Corporation, 2020). The discomfort caused by improper and extended usage can have significant debilitating effects manifesting as cybersickness. While these symptoms are only temporary, they often have a much more significant impact than traditional VDU (Visual Display Unit) usage fatigue. Therefore while a monitoring system, is a superb idea. Detecting when a user is reaching an acceptable limit of their tolerance is currently not possible. However, with further research, these predictions could be made from within the virtual environment itself; the information could allow real-time adjustments to the environment, such as the speed the user navigates the environment, to reduce the chances of cybersickness onset tailored to the individuals' responses.

Content removed due to copyright restrictions

Figure 6-2:- Reminder to have the player take a break in Guild Wars. (https://steamcommunity.com/sharedfiles/filedetails/?id=1390233024)

7 Conclusions

This report has demonstrated that a proxy measurement for subject stability can be captured via a VR HMD using the data produced during the headset's regular operation. However, these readings are likely to be significantly less reliable than traditional postural stability recording methods such as a CoP measurement. Data capture highlighted several significant issues with the process of data recording during the quiet stand period, such as the deliberate movement of the subject and recalculation of position, resulting in the necessary removal of significant portions of the data set to ensure reliability. This study's evidence suggests that head rotation of the subject could influence the headset's reported position due to tracking the HMD position instead of the user’s head position. The study also found the raw feed from Steam VR regarding the VR HMD position to be difficult to use, flipping the X and Y coordinates of the tracked space unpredictably.

During this report's testing, the instability measurements captured seem to accurately record a decent approximation for subject stability during VR HMD usage. Analysis of individual motion patterns recorded seems to follow the existing literature regarding increases in motion on the positional and rotational axis. Analysis of the mean instability experienced by sick and well groups seemed to confirm postural stability theory, with an increase in SSQ scores also showing an increase in measured instability. However, analysis of individual results shows a wide range of instability values recorded for both sick and well groups, with the spread of instability values being relatively even for both sick and well participants. Therefore, proposing a mathematical model to separate the two groups became impossible.

Investigating other factors, familiarity with VR environments seemed to have the opposite effect to that which was expected, with experienced users showing a slightly higher instability than their inexperienced counterparts. SSQ-T scores recorded for experienced users were generally lower than inexperienced users in line with expectations. A significant reduction in mean instability was recorded between subjects' first runs vs their second run, regardless of the task order or user experience. These findings provided a possible insight into how humans deal with the conflict between the vestibule and visual systems induced in VR environments. Suggesting that while an individual may develop a tolerance to VR HMD environment usage in general, the decrease in instability recorded over multiple exposures to the test environment suggests that adaptation must occur during each experience. With further research, this finding may result in a new model for how humans learn stability in VR environments, possibly identifying that each environment and its different sensory mismatch rates require individual learning to develop tolerance.

Also investigated in this report was the concept that task difficulty could influence instability in a VR environment. The expected effect was an increase in instability measurements could be associated with an increase in task load, caused by an increase in the demand on the subject's working memory, resulting in less space for processing environmental mismatches. Interestingly the maze task seemed to increase stability despite subjectively being the most provocative stimuli of the three used during testing. Testing with spatial load yielded a significant increase in instability measurements from users, and the overload condition yielded similar levels of instability to those in the spatial load tests. All three load conditions showed higher instability levels in their high load conditions than in the low load condition. However, the primary point of instability gain seemed to be the spatial test. As the test order was static, the increase in instability recorded here could be evidence of the effect of test order or user fatigue. However, the results here warrant further investigation into the matter.

In appraising the suitability of this methodology's real-world application for detecting cybersickness, we find it completely inappropriate in its current form. The evidence provided in this report suggested no way of detecting the difference between sick and well participants via instability measurements made via a VR HMD. However, as the sample size is small, and the standard deviation of recorded results is high; The possibility that this method may be valid still exists. The data capture method was also assessed for feasibility in deployment into VR applications; the idea of having users of the VR applications perform the quiet stand test regularly to determine if they are sick or not is impractical. Firstly a 30-second stand with any frequency sufficient to be useful will quickly become irritating to users and likely be dismissed if possible. However, in some scenarios (particularly game environments), sufficient excuses to perform this task could be engineered, such as elevator rides or sci-fi scanning.

Testing showed that getting users to adopt the quiet stand properly was difficult, with most requiring direct instruction to perform the task correctly. This likely means the user will not adopt the stance correctly in a home or work environment and is unlikely to engage with the task if perceived as not useful or tedious. The final nail in this approach is the simple fact that if a user feels sick, it is doubtful they will stand still long enough to make a measurement, more likely removing the headset and removing themselves from the source of the sickness.

Therefore, additional analysis was performed on the measurements obtained during testing outside of the quiet stand periods. While subjects navigated a straight corridor, attempting to take measurements provided wildly inconsistent results, with some subjects reporting a drop in instability from their quiet stand results. Attempting to use the moment before a subject took a shot in the shooting test yielded results with no significant improvement over just analysing the whole sample. No results acquired during this phase correlated with SSQ-T scores or pre and post-test quiet stand patterns, suggesting that instability measurements acquired here were not good indicators for detecting cybersickness in subjects.

It is also essential to consider the influence of the COVID-19 pandemic on this study. While testing, the primary concern was participant safety, and the addition of safety equipment, such as gloves, masks, and disinfectant spray, was necessary. The addition of these precautions may have had an impact on the quality of results gathered here. While all participation was voluntary, the potential risk may have altered participants' mindsets. Thus, as our results for determining sickness are purely subjective, they may have influenced the severity of sickness symptoms reported in the SSQ. The addition of masks will also have impacted breathing during the experiment and may have had an unquantified impact on the breathing of subjects affecting results. While these factors are outside the control of this experiment, repeating the experiment without the safety equipment should be considered necessary to validate the results' accuracy and reliability when it is safe to do so.

In conclusion, it seems that while possible, tracking user head position in VR HMD environments is an ineffective means of identifying sickness. This conclusion means that the primary goal of this study remains unrealized. However, several potential avenues for investigation were uncovered during the process that advance knowledge within the field. Firstly, the idea of utilizing the headset to capture positional data is plausible but likely requires investigation and implementation on a per headset basis to determine the reliability and accuracy of the results. Second, Instability measurements captured using the VR HMD measurements are derived primarily from unfamiliarity with the specific environment rather than VR experience in general, identifying a significant potential impact on the validity of pairwise trials in a VR environment. Third, the task's perceived difficulty has a significant impact on instability recorded from the VR HMD device; while other factors also have an influence, it does confirm the need to consider the type of task performed when comparing instability measurements. Finally, implementing this test into a game environment in a passive manner may be possible. However, measurements are unlikely to be comparable to quiet stand measurements and probably unlikely to identify cybersickness with any accuracy.

Referring back to the original research questions, these can now be stated to be:

1. An increase in cognitive load within a virtual reality task influences the severity of cybersickness symptoms, however, there are significant confounding factors with this influence.
2. An increase in the cognitive load experienced by a subject within a virtual reality environment impacts postural stability measurements.
3. The alteration of task performance can be used to identify the onset and severity of cybersickness in an individual, however, there are significant confounding factors to this identification.
4. The onset of cybersickness cannot be practically identified mid-task using commercial off-the-shelf VR equipment.

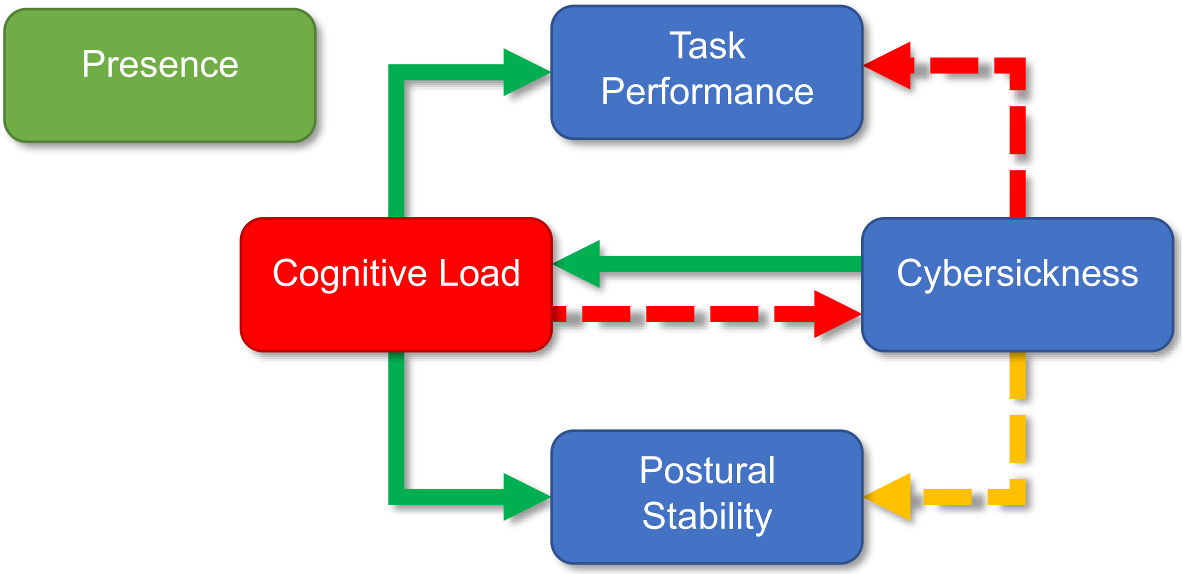


Figure 7-1:- Diagram from section 1-1 showing the structure of the thesis, now to include the new knowledge obtained through the thesis.

Referring back to our original diagram (Figure 1-1) we can now address what we know about the 3 links in question. The link between cybersickness and its relationship on task performance is confirmed, perceived task performance does drop as cybersickness symptoms increase. Additionally the study has demonstrated a link between cognitive load and cybersickness showing that higher load conditions do indeed result in high sickness scores, however the evidence is not wholly conclusive and as such further work is required. Finally the potential link between cyber sickness and postural stability has been proved to be influenced by multiple external factors, including user and familiarity with the environment, but most impornatly that the task load does influence the stability measurement, validating the projects concerns regarding comparing different studies with different task loads. This highlights a significant issue with the approach and as such further study should be completed. Figure 7-1 shows the updated knowledge graph to reflect the new knowledge gained.

Further work should focus on investigating the impact of task load on headset instability; the evidence presented in this report shows a significant contribution to headset motion from the task's spatial load. However, a possible contribution from test order does exist, masking a possible contribution from the overload condition. Therefore, a repeat experiment randomising the task order would allow for an investigation into whether a contribution to instability from overload exists. It cannot be determined if the headset instability added from task load translates into an increase in core postural instability. Investigating this with the traditional pressure plate would potentially yield significant ramifications for detecting cybersickness via postural stability.

Finally, this report's findings suggested that increases in headset motion may not correspond to cybersickness; instead, it may be the body's response to trying to figure out the environment as evidenced by the significant drop in head motion from the first run to the second. Confirming this would go a significant way toward establishing the viability of using VR HMD positional measurements to confirm the presence of cybersickness in headset users.

8 References

A. B. C. News, 2019. Why you feel motion sickness during virtual reality [WWW Document]. ABC News. URL https://abcnews.go.com/Technology/feel-motion-sickness-virtual-reality/story?id=65153805 (accessed 3.11.21).

Al Zayer, M., MacNeilage, P., Folmer, E., 2020. Virtual Locomotion: A Survey. IEEE Trans. Visual. Comput. Graphics 26, 2315–2334. https://doi.org/10.1109/TVCG.2018.2887379

Andersen, S.A.W., Konge, L., Sørensen, M.S., 2018. The effect of distributed virtual reality simulation training on cognitive load during subsequent dissection training. Medical Teacher 40, 684–689. https://doi.org/10.1080/0142159X.2018.1465182

Andersson, G., Hagman, J., Talianzadeh, R., Svedberg, A., Larsen, H.C., 2002. Effect of cognitive load on postural control. Brain Research Bulletin 58, 135–139. https://doi.org/10.1016/S0361-9230(02)00770-0

Ansorge, P., 2020. Sixty percent want to stay at home after COVID-19 | theHRD [WWW Document]. URL https://www.thehrdirector.com/business-news/work_life_balance/data-reveals-60-of-people-want-to-stay-at-home-after-covid-19/ (accessed 7.31.20).

Arcioni, B., Palmisano, S., Apthorp, D., Kim, J., 2019. Postural stability predicts the likelihood of cybersickness in active HMD-based virtual reality. Displays 58, 3–11. https://doi.org/10.1016/j.displa.2018.07.001

Armougum, A., Orriols, E., Gaston-Bellegarde, A., Marle, C.J.-L., Piolino, P., 2019. Virtual reality: A new method to investigate cognitive load during navigation. Journal of Environmental Psychology 65, 101338. https://doi.org/10.1016/j.jenvp.2019.101338

Arns, L.L., Cerney, M.M., 2005. The relationship between age and incidence of cybersickness among immersive environment users, in: IEEE Proceedings. VR 2005. Virtual Reality, 2005. IEEE, Bonn, Germany, pp. 267–268. https://doi.org/10.1109/VR.2005.1492788

August, K.G., Adamovich, Sergei, August, K., Bleichenbacher, D., D, P., Adamovich, S., D, P., 2005. Virtual Reality Physical Therapy: A Telerehabilitation Tool for Hand and Finger Movement Exercise Monitoring and Motor Skills Analysis.

Aukstakalnis, S., 2017. Practical augmented reality: a guide to the technologies, applications, and human factors for AR and VR.

Avatar, 2009. . Twentieth Century Fox, Dune Entertainment, Lightstorm Entertainment.

Barten, P.G., 1999. Contrast Sensitivity of the Human Eye and Its Effects on Image Quality. SPIE. https://doi.org/10.1117/3.353254

Bartlett, H.L., Ting, L.H., Bingham, J.T., 2014. Accuracy of force and center of pressure measures of the Wii Balance Board. Gait Posture 39. https://doi.org/10.1016/j.gaitpost.2013.07.010

Bauer, P., Lienhart, W., Jost, S., 2021. Accuracy Investigation of the Pose Determination of a VR System. Sensors 21, 1622. https://doi.org/10.3390/s21051622

Baur, D., Pfeifle, C., Heyde, C.E., 2021. Cervical spine injury after virtual reality gaming: a case report. J Med Case Reports 15, 312. https://doi.org/10.1186/s13256-021-02880-9

Boletsis, C., Cedergren, J.E., 2019. VR Locomotion in the New Era of Virtual Reality: An Empirical Comparison of Prevalent Techniques. Advances in Human-Computer Interaction 2019, 1–15. https://doi.org/10.1155/2019/7420781

Bonato, F., Bubka, A., Palmisano, S., 2009. Combined Pitch and Roll and Cybersickness in a Virtual Environment. aviat space environ med 80, 941–945. https://doi.org/10.3357/ASEM.2394.2009

Bonato, F., Bubka, A., Palmisano, S., Phillip, D., Moreno, G., 2008. Vection Change Exacerbates Simulator Sickness in Virtual Environments. Presence: Teleoperators and Virtual Environments 17, 283–292. https://doi.org/10.1162/pres.17.3.283

Bos, J.E., 2015. Less sickness with more motion and/or mental distraction. Journal of Vestibular Research 25, 23–33. https://doi.org/10.3233/VES-150541

Bos, J.E., Bles, W., Groen, E.L., 2008. A theory on visually induced motion sickness. Displays, Health and Safety Aspects of Visual Displays 29, 47–57. https://doi.org/10.1016/j.displa.2007.09.002

British Broadcasting Corporation, 2020. Developer warns VR headset damaged eyesight. BBC News, Published on 10th June 2020.

Brooks, X., 2009. James Cameron's 3D Avatar leaves reviewer queasy. The Guardian, Published on 19th August 2009.

Bruck, S., Watters, P.A., 2011. The factor structure of cybersickness. Displays 32, 153–158. https://doi.org/10.1016/j.displa.2011.07.002

Buettner, R., 2013. Cognitive Workload of Humans Using Artificial Intelligence Systems: Towards Objective Measurement Applying Eye-Tracking Technology, in: Timm, I.J., Thimm, M. (Eds.), KI 2013: Advances in Artificial Intelligence, Lecture Notes in Computer Science. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 37–48. https://doi.org/10.1007/978-3-642-40942-4_4

Caddy, B., 2016. Vomit Reality: Why VR makes some of us feel sick and how to make it stop [WWW Document]. Wareable. URL https://www.wareable.com/vr/vr-headset-motion-sickness-solution-777 (accessed 3.11.21).

Carrougher, G.J., Hoffman, H.G., Nakamura, D., Lezotte, D., Soltani, M., Leahy, L., Engrav, L.H., Patterson, D.R., 2009. The Effect of Virtual Reality on Pain and Range of Motion in Adults With Burn Injuries: Journal of Burn Care & Research 30, 785–791. https://doi.org/10.1097/BCR.0b013e3181b485d3

Chandler, P., Sweller, J., 1991. Cognitive Load Theory and the Format of Instruction. Cognition and Instruction 8, 293–332. https://doi.org/10.1207/s1532690xci0804_2

Chen, W., Chen, J.Z., So, R.H.Y., 2011. Chen-et-al-2011-EHF2011.pdf, in: Contemporary Ergonomics and Human Factors 2011: CRC Press, pp. 281–287.

Cheung, B., Vaitkus, P., 1998. Perspectives of electrogastrography and motion sickness. Brain Research Bulletin 47, 421–431. https://doi.org/10.1016/S0361-9230(98)00095-1

Clifton, J., Palmisano, S., 2020. Effects of steering locomotion and teleporting on cybersickness and presence in HMD-based virtual reality. Virtual Reality 24, 453–468. https://doi.org/10.1007/s10055-019-00407-8

CNN, S.L., 2017. The very real health dangers of virtual reality [WWW Document]. CNN. URL https://www.cnn.com/2017/12/13/health/virtual-reality-vr-dangers-safety/index.html (accessed 7.29.22).

Cobb, S.V.G., Nichols, S., Ramsey, A., Wilson, J.R., 1999. Virtual reality-induced symptoms and effects (VRISE). Presence-Teleoper. Virtual Env. 8, 169–186. https://doi.org/10.1162/105474699566152
Colgan, A., 2015. Weightless Creator on Designing a Zero-G Sci-Fi Playground [WWW Document]. Leap Motion Blog. URL http://blog.leapmotion.com/weightless-creator-martin-schubert-designing-zero-gravity-sci-fi-playground/ (accessed 3.13.21).
Collins, K., Kapralos, B., Tessler, H., Paterson, N., Conway, F., 2014. Engagement, Immersion and Presence, in: Collins, K., Kapralos, B., Tessler, H. (Eds.), The Oxford Handbook of Interactive Audio. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199797226.013.016
Cowan, N., 2008. What are the differences between long-term, short-term, and working memory? Prog Brain Res 169, 323–338. https://doi.org/10.1016/S0079-6123(07)00020-9
Dakers, P., 2020. Virtual reality – what is the screen door effect and how do we eliminate it? [WWW Document]. udl. URL https://www.udl.co.uk/insights/virtual-reality-what-is-the-screen-door-effect-and-how-do-we-eliminate-it (accessed 3.11.21).
Davis, A., Kühnlenz, F., 2007. Optical Design using Fresnel Lenses: Basic Principles and some Practical Examples. Optik & Photonik 2, 52–55. https://doi.org/10.1002/opph.201190287
Davis, S., Nesbitt, K., Nalivaiko, E., 2015. Comparing the onset of cybersickness using the Oculus Rift and two virtual roller coasters. In the proceedings of the 11th Australasian Conference on Interactive Entertainment (IE 2015) (Vol 27, p30).
Dealessandri, M., 2020. The best practices and design principles of VR development [WWW Document]. GamesIndustry.biz. URL https://www.gamesindustry.biz/articles/2020-04-01-the-best-practices-and-design-principles-of-vr-development (accessed 3.13.21).
Denise, P., Vouriot, A., Normand, H., Golding, J.F., Gresty, M.A., 2009. Effect of temporal relationship between respiration and body motion on motion sickness. Autonomic Neuroscience 151, 142–146. https://doi.org/10.1016/j.autneu.2009.06.007
Dennison, M., D'Zmura, M., 2018. Effects of unexpected visual motion on postural sway and motion sickness. Applied Ergonomics 71, 9–16. https://doi.org/10.1016/j.apergo.2018.03.015
Dennison, M.S., D'Zmura, M., 2017. Cybersickness without the wobble: Experimental results speak against postural instability theory. Applied Ergonomics 58, 215–223. https://doi.org/10.1016/j.apergo.2016.06.014
Dennison, M.S., Wisti, A.Z., D'Zmura, M., 2016. Use of physiological signals to predict cybersickness. Displays 44, 42–52. https://doi.org/10.1016/j.displa.2016.07.002
Diels, C., Howarth, P.A., 2011. Visually induced motion sickness: Single- versus dual-axis motion. Displays 32, 175–180. https://doi.org/10.1016/j.displa.2011.02.005
DiMaggio, P., 1997. Culture and Cognition. Annual Review of Sociology 23, 25. https://doi.org/10.1146/annurev.soc.23.1.263

Domeyer, J.E., Cassavaugh, N.D., Backs, R.W., 2013. The use of adaptation to reduce simulator sickness in driving assessment and research. Accident Analysis & Prevention 53, 127–132. https://doi.org/10.1016/j.aap.2012.12.039
Drexler, J., 2006. Identification Of System Design Features That Affect Sickness In Virtual Environments. University of Central Florida.
Ebenholtz, S., Cohen, M., Linder, B., 1994. The possible role of nystagmus in motion sickness: A hypothesis. Aviation, space, and environmental medicine 65, 1032–5.
Edwards, B., Edwards, B., 2015. Unraveling The Enigma Of Nintendo's Virtual Boy, 20 Years Later [WWW Document]. Fast Company. URL https://www.fastcompany.com/3050016/unraveling-the-enigma-of-nintendos-virtual-boy-20-years-later (accessed 5.14.20).
Epic Games, 2018. Blueprints Visual Scripting [WWW Document]. URL https://docs.unrealengine.com/en-us/Engine/Blueprints (accessed 4.9.18).
Faber, J., Fonseca, L.M., 2014. How sample size influences research outcomes. Dental Press J Orthod 19, 27–29. https://doi.org/10.1590/2176-9451.19.4.027-029.ebo
Facebook Technologies, 2021a. Testing and Performance Analysis | Oculus Developers [WWW Document]. URL https://developer.oculus.com/documentation/unity/unity-perf/?locale=en_GB (accessed 3.11.21).
Facebook Technologies, 2021b. Oculus Debug Tool | Oculus Developers [WWW Document]. URL https://developer.oculus.com/documentation/native/pc/dg-debug-tool/?locale=en_GB (accessed 3.13.21).
Facebook Technologies, 2020. Oculus Guardian System [WWW Document]. URL https://developer.oculus.com/documentation/native/pc/dg-guardian-system/ (accessed 6.11.20).
Facebook technologies, LLC, 2020. Rift Store: VR Games, Apps and More [WWW Document]. Oculus. URL https://www.oculus.com/experiences/rift/ (accessed 5.27.20).
Farmani, Y., Teather, R.J., 2018. Viewpoint Snapping to Reduce Cybersickness in Virtual Reality 8.
Ferdig, R.E. (Ed.), 2009. Handbook of Research on Effective Electronic Gaming in Education: IGI Global. https://doi.org/10.4018/978-1-59904-808-6
Fernandes, A.S., Feiner, S.K., 2016. Combating VR sickness through subtle dynamic field-of-view modification, in: 2016 IEEE Symposium on 3D User Interfaces (3DUI). IEEE, Greenville, SC, USA, pp. 201–210. https://doi.org/10.1109/3DUI.2016.7460053
Fowle, K., 2015. A look back at the doomed virtual reality boom of the 90s [WWW Document]. Kill Screen. URL https://killscreen.com/previously/articles/failure-launch/ (accessed 3.11.21).
Freeman, D., Reeve, S., Robinson, A., Ehlers, A., Clark, D., Spanlang, B., Slater, M., 2017. Virtual reality in the assessment, understanding, and treatment of mental health disorders. Psychol. Med. 47, 2393–2400. https://doi.org/10.1017/S003329171700040X
Garcia-Agundez, A., Reuter, C., Caserman, P., Konrad, R., Göbel, S., 2019. Identifying Cybersickness through Heart Rate Variability alterations. IJVR 19, 1–10. https://doi.org/10.20870/IJVR.2019.19.1.2907

Gavgani, A.M., Nesbitt, K.V., Blackmore, K.L., Nalivaiko, E., 2017. Profiling subjective symptoms and autonomic changes associated with cybersickness. Autonomic Neuroscience 203, 41–50. https://doi.org/10.1016/j.autneu.2016.12.004
Gilson, S.J., Fitzgibbon, A.W., Glennerster, A., 2006. Quantitative analysis of accuracy of an inertial/acoustic 6DOF tracking system in motion. J Neurosci Methods 154, 175–182. https://doi.org/10.1016/j.jneumeth.2005.12.013
Golding, J.F., 1998. Motion sickness susceptibility questionnaire revised and its relationship to other forms of sickness. Brain Research Bulletin 47, 507–516. https://doi.org/10.1016/S0361-9230(98)00091-4
Graybiel, A., Lackner, J.R., 1983. Motion sickness: acquisition and retention of adaptation effects compared in three motion environments. Aviat Space Environ Med 54, 307–311.
Haklay, M., 2001. Virtual reality and GIS: Applications, trends and directions, in: Virtual Reality in Geography. CRC Press, pp. 47–57. https://doi.org/10.1201/9780203305850.ch5
Halarewich, D., 2016. Reducing Cognitive Overload For A Better User Experience [WWW Document]. Smashing Magazine. URL https://www.smashingmagazine.com/2016/09/reducing-cognitive-overload-for-a-better-user-experience/ (accessed 3.15.21).
Handford, M., 2010. The Language of Business Meetings. Cambridge University Press.
Harper, J., 2020. Flexible working will be a new normal after virus. BBC News, Published on 22nd of May 2020.
Harris, D., Wilson, M., Vine, S., 2020. Development and validation of a simulation workload measure: the simulation task load index (SIM-TLX). Virtual Reality 24, 557–566. https://doi.org/10.1007/s10055-019-00422-9
Harris, M., 2019. Virtual & Augmented Reality Field Trips [WWW Document]. URL https://www.learnersedge.com/blog/virtual-and-augmented-reality-field-trips (accessed 3.31.21).
Hart, S., 1986. Task Load Index (NASA TLX) [WWW Document]. URL https://ntrs.nasa.gov/citations/20000021488
Hart, S.G., Field, M., 2006. NASA-TASK LOAD INDEX (NASA-TLX); 20 YEARS LATER. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Volume 50, Issue 9.
Hawthorne, M., 2021. VR Oculus Quest 2: What Is Passthrough? [WWW Document]. Technipages. URL https://www.technipages.com/vr-oculus-quest-2-what-is-passthrough (accessed 7.30.22).
Health and Safety Executive, 2020. Manual handling at work. A brief guide [WWW Document]. URL https://www.hse.gov.uk/pubns/indg143.PDF
Health and Safety Executive, 2019. Working safely with display screen equipment - HSE [WWW Document]. URL http://www.hse.gov.uk/msd/dse/ (accessed 5.1.19).
Health and Safety Executive, United Kingdom, 1974. Health and Safety at Work Act [WWW Document]. URL https://www.legislation.gov.uk/ukpga/1974/37 (accessed 2.17.21).
Heaney, D., 2019. How VR Positional Tracking Systems Work. URL https://uploadvr.com/how-vr-tracking-works/ (accessed 3.13.21).

Heaven, D., 2013. Virtual reality rises again. New Scientist 218, 20. https://doi.org/10.1016/S0262-4079(13)61540-7
Heilig, M.L., 1962. Sensorama simulator, US patent. US3050870A.
Hettinger, L.J., Haas, M., 2003. A unified approach to presence and motion sickness, in: Virtual and Adaptive Environments: Applications, Implications, and Human Performance Issues. pp. 47–66.
Heutink, J., Broekman, M., Brookhuis, K.A., Melis-Dankers, B.J.M., Cordes, C., 2019. The effects of habituation and adding a rest-frame on experienced simulator sickness in an advanced mobility scooter driving simulator. Ergonomics 62, 65–75. https://doi.org/10.1080/00140139.2018.1518543
Hill, K.J., Howarth, P.A., 2000. Habituation to the side effects of immersion in a virtual environment. Displays 21, 25–30. https://doi.org/10.1016/S0141-9382(00)00029-9
Hoffman, D.M., Girshick, A.R., Akeley, K., Banks, M.S., 2008. Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. Journal of Vision 8, 33–33. https://doi.org/10.1167/8.3.33
HTC Corporation, 2018. VIVE™ | Discover Virtual Reality Beyond Imagination [WWW Document]. URL https://www.vive.com/uk/ (accessed 4.9.18).
Ilieva, S., 2016. Peripheral Vision Drives User Attention [WWW Document]. Progress Blogs. URL https://www.progress.com/blogs/the-surprising-potential-of-peripheral-vision-in-driving-user-attention (accessed 3.16.21).
Indeed, 2020. 4 Types of Communication (With Examples) [WWW Document]. Indeed Career Guide. URL https://www.indeed.com/career-advice/career-development/types-of-communication (accessed 7.9.20).
Israel, K., Zerres, C., Tscheulin, D.K., Buchweitz, L., Korn, O., 2019. Presenting Your Products in Virtual Reality: Do not Underestimate Cybersickness, in: Nah, F.F.-H., Siau, K. (Eds.), HCI in Business, Government and Organizations. ECommerce and Consumer Behavior, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 206–224. https://doi.org/10.1007/978-3-030-22335-9_14
Jenkins, A., 2019. The fall and rise of VR: The struggle to make virtual reality get real [WWW Document]. Fortune. URL https://fortune.com/longform/virtual-reality-struggle-hope-vr/ (accessed 2.9.21).
Jeong, D., Yoo, S., Yun, J., 2019. Cybersickness Analysis with EEG Using Deep Learning Algorithms, in: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). pp. 827–835. https://doi.org/10.1109/VR.2019.8798334
JiaxinWen, 2017. Non-diegetic UI Design for VR Games. Jiaxin Wen. URL https://jiaxinwen.wordpress.com/2017/03/30/non-diegetic-ui-design-for-vr-games/ (accessed 3.16.21).
Jimenez, Y.A., Cumming, S., Wang, W., Stuart, K., Thwaites, D.I., Lewis, S.J., 2018. Patient education using virtual reality increases knowledge and positive experience for breast cancer patients undergoing radiation therapy. Support Care Cancer 26, 2879–2888. https://doi.org/10.1007/s00520-018-4114-4
Jin, W., Fan, J., Gromala, D., Pasquier, P., 2018. Automatic Prediction of Cybersickness for Virtual Reality Games, in: 2018 IEEE Games, Entertainment, Media Conference (GEM). IEEE, Galway, pp. 1–9. https://doi.org/10.1109/GEM.2018.8516469

Joseph, J.A., Griffin, M.J., 2008. Motion Sickness: Effect of the Magnitude of Roll and Pitch Oscillation. asem 79, 390–396. https://doi.org/10.3357/ASEM.2196.2008
Kallie, C.S., Schrater, P.R., Legge, G.E., 2007. Variability in stepping direction explains the veering behavior of blind walkers. Journal of Experimental Psychology: Human Perception and Performance 33, 183–200. https://doi.org/10.1037/0096-1523.33.1.183
Kellogg, R.S., Kennedy, R.S., Graybiel, A., 1965. Motion Sickness Symptomatology of Labyrinthine Defective and Normal Subjects During Zero Gravity Maneuvers. Aerosp Med 36, 315–318.
Kennedy, R.S., Lane, N.E., 1993. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. The International Journal of Aviation Psychology 203–220.
Kennedy, R.S., Lane, N.E., Lilienthal, M.G., Berbaum, K.S., Hettinger, L.J., 1992. Profile Analysis of Simulator Sickness Symptoms: Application to Virtual Environment Systems. Presence: Teleoperators and Virtual Environments 1, 295–301. https://doi.org/10.1162/pres.1992.1.3.295
Keo, M., 2017. Graphical Styles in Video Games. Häme University of Applied Sciences 75.
Keshavarz, B., Hecht, H., 2012. Visually induced motion sickness and presence in videogames: The role of sound. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 56, 1763–1767. https://doi.org/10.1177/1071181312561354
Keshavarz, B., Hecht, H., 2011. Validating an Efficient Method to Quantify Motion Sickness. Hum Factors 53, 415–426. https://doi.org/10.1177/0018720811403736
Kim, H.K., Park, J., Choi, Y., Choe, M., 2018. Virtual reality sickness questionnaire (VRSQ): Motion sickness measurement index in a virtual reality environment. Applied Ergonomics 69, 66–73. https://doi.org/10.1016/j.apergo.2017.12.016
Kim, Y.Y., Kim, H.J., Kim, E.N., Ko, H.D., Kim, H.T., 2005. Characteristic changes in the physiological components of cybersickness. Psychophysiology 0, 050826083901001-??? https://doi.org/10.1111/j.1469-8986.2005.00349.x
Klosterhalfen, S., Kellermann, S., Pan, F., Stockhorst, U., Hall, G., Enck, P., 2005. Effects of ethnicity and gender on motion sickness susceptibility. Aviat Space Environ Med 76, 1051–1057.
Koenig, S.D., 2017. Remembering What We’ve Seen: The Hippocampus and Relational Memory. University of Washington.
Kolasinski, E.M., 1995. Simulator Sickness in Virtual Environments. U.S. Army Research Institute for the Behavioral and Social Sciences.
Kramida, G., 2016. Resolving the Vergence-Accommodation Conflict in Head-Mounted Displays. IEEE Trans. Visual. Comput. Graphics 22, 1912–1931. https://doi.org/10.1109/TVCG.2015.2473855
Krug, S., 2021. Cognitive load | Design Defined | InVision [WWW Document]. Cognitive load | Design Defined | InVision. URL https://www.invisionapp.com/design-defined/cognitive-load/ (accessed 3.15.21).
Lang, B., 2020. Hands-on: CREAL is Shrinking Its Light-field Display for AR & VR Headsets. Road to VR. URL https://www.roadtovr.com/ces-2020-creal-light-field-shrinking-for-ar-vr-headsets/ (accessed 3.18.21).

Lang, B., 2018. “Borderlands 2 VR” Review – An Ok Version of a Good Game. Road to VR. URL https://www.roadtovr.com/borderlands-2-vr-review-psvr-an-ok-version-of-a-good-game/ (accessed 3.16.21).
LaValle, S.M., Yershova, A., Katsev, M., Antonov, M., 2014. Head tracking for the Oculus Rift, in: 2014 IEEE International Conference on Robotics and Automation (ICRA). IEEE, Hong Kong, China, pp. 187–194. https://doi.org/10.1109/ICRA.2014.6906608
LaViola, J.J., 2000. A discussion of cybersickness in virtual environments. ACM SIGCHI Bulletin 32, 47–56. https://doi.org/10.1145/333329.333344
Leung, A.K., Hon, K.L., 2019. Motion sickness: an overview. Drugs Context 8. https://doi.org/10.7573/dic.2019-9-4
Levine, M.E., Stern, R.M., 2002. Spatial Task Performance, Sex Differences, and Motion Sickness Susceptibility. Percept Mot Skills 95, 425–431. https://doi.org/10.2466/pms.2002.95.2.425
Lin, C.-T., Chuang, S.-W., Chen, Y.-C., Ko, L.-W., Liang, S.-F., Jung, T.-P., 2007. EEG Effects of Motion Sickness Induced in a Dynamic Virtual Reality Environment, in: 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, Lyon, France, pp. 3872–3875. https://doi.org/10.1109/IEMBS.2007.4353178
Lin, J.J.-W., Duh, H.B.L., Parker, D.E., Abi-Rached, H., Furness, T.A., 2002. Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment, in: Proceedings IEEE Virtual Reality 2002. IEEE Comput. Soc, Orlando, FL, USA, pp. 164–171. https://doi.org/10.1109/VR.2002.996519
Lin, Y.-T., Chien, Y.-Y., Wang, H.-H., Lin, F.-C., Huang, Y.-P., 2018. 65-3: The Quantization of Cybersickness Level Using EEG and ECG for Virtual Reality Head-Mounted Display. SID Symposium Digest of Technical Papers 49, 862–865. https://doi.org/10.1002/sdtp.12267
Lin, Y.-X., Venkatakrishnan, Rohith, Venkatakrishnan, Roshan, Ebrahimi, E., Lin, W.-C., Babu, S.V., 2020. How the Presence and Size of Static Peripheral Blur Affects Cybersickness in Virtual Reality. ACM Trans. Appl. Percept. 17, 1–18. https://doi.org/10.1145/3419984
Lo, W.T., So, R.H.Y., 2001. Cybersickness in the presence of scene rotational movements along different axes. Applied Ergonomics 14.
Longstaffe, K.A., Hood, B.M., Gilchrist, I.D., 2014. The influence of cognitive load on spatial search performance. Atten Percept Psychophys 76, 49–63. https://doi.org/10.3758/s13414-013-0575-1
Lubin, G., 2016. Here’s how close AI is to beating humans in different games [WWW Document]. Business Insider. URL https://www.businessinsider.com/ai-vs-humans-in-different-games-2016-11 (accessed 3.17.21).
Ludlow, B., 2015. Virtual Reality: Emerging Applications and Future Directions. Rural Special Education Quarterly 34, 3–10. https://doi.org/10.1177/875687051503400302
Magaki, T., Vallance, M., 2019. Real-time Monitoring Method for Cybersickness using Physiological Signals, in: 25th ACM Symposium on Virtual Reality Software and Technology. ACM, Parramatta NSW Australia, pp. 1–2. https://doi.org/10.1145/3359996.3364700
Magnenat-Thalmann, N., Bonanni, U., 2006. Haptics in Virtual Reality and Multimedia. IEEE Multimedia 13, 6–11. https://doi.org/10.1109/MMUL.2006.56

Marcus Andrews, 2010. Game UI Discoveries: What Players Want [WWW Document]. Game Developer. URL https://www.gamedeveloper.com/design/game-ui-discoveries-what-players-want (accessed 7.30.22).
Martindale, J., 2017. VR locomotion solved? The CAOTS system | KitGuru. URL https://www.kitguru.net/components/vr/jon-martindale/vr-locomotion-solved-the-caots-system/ (accessed 3.13.21).
Mazloumi Gavgani, A., Walker, F.R., Hodgson, D.M., Nalivaiko, E., 2018. A comparative study of cybersickness during exposure to virtual reality and “classic” motion sickness: are they different? Journal of Applied Physiology 125, 1670–1680. https://doi.org/10.1152/japplphysiol.00338.2018
Mazuryk, T., Gervautz, M., 1999. History, Applications, Technology and Future. Vienna: Institute of Computer Graphics Vienna University of Technology.
Meehan, M., Razzaque, S., Whitton, M.C., Brooks, F.P., 2003. Effect of latency on presence in stressful virtual environments, in: IEEE Virtual Reality, 2003. Proceedings. IEEE Comput. Soc, Los Angeles, CA, USA, pp. 141–148. https://doi.org/10.1109/VR.2003.1191132
Melnick, K., 2020. Virtuix Reveals Omni One At-Home VR Treadmill - VRScout [WWW Document]. URL https://vrscout.com/news/virtuix-at-home-omni-one-vr-treadmill/ (accessed 3.13.21).
Miller, G.A., 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychological Review 63, 81–97. https://doi.org/10.1037/h0043158
Miller (II), E.F., Graybiel, A., 1969. A Standardized Laboratory Means of Determining Susceptibility to Coriolis (motion) Sickness. Naval Aerospace Medical Institute, Naval Aerospace Medical Center.
Mittelstaedt, J.M., Wacker, J., Stelling, D., 2019. VR aftereffect and the relation of cybersickness and cognitive performance. Virtual Reality 23, 143–154. https://doi.org/10.1007/s10055-018-0370-3
Monad Rock, 2018. Summer Funland on Steam [WWW Document]. URL https://store.steampowered.com/app/780280/Summer_Funland/ (accessed 3.22.21).
Murata, A., 2004. Effects of Duration of Immersion in a Virtual Reality Environment on Postural Stability. International Journal of Human-Computer Interaction 17, 463–477. https://doi.org/10.1207/s15327590ijhc1704_2
Mythbusters Seasickness: Kill or Cure, 2005. . Beyond Entertainment.
Nalivaiko, E., Davis, S.L., Blackmore, K.L., Nesbitt, K.V., 2015. Cybersickness provoked by head-mounted display affects cutaneous vascular tone, heart rate and reaction time. Autonomic Neuroscience 192, 63. https://doi.org/10.1016/j.autneu.2015.07.032
Naqvi, S.A.A., Badruddin, N., Malik, A.S., Hazabbah, W., 2013. Does 3D Produce More Symptoms of Visually Induced Motion Sickness? 35th Annual International Conference of the IEEE EMBS 4.
Nelson, B.W., Allen, N.B., 2018. Accuracy of Consumer Wearable Heart Rate Measurement During an Ecologically Valid 24-Hour Period: Intraindividual Validation Study (Preprint) (preprint). JMIR mHealth and uHealth. https://doi.org/10.2196/preprints.10828

Nesbitt, K., Davis, S., Blackmore, K., Nalivaiko, E., 2017. Correlating reaction time and nausea measures with traditional measures of cybersickness. Displays 48, 1–8. https://doi.org/10.1016/j.displa.2017.01.002

Ng, A.K.T., Chan, L.K.Y., Lau, H.Y.K., 2020. A study of cybersickness and sensory conflict theory using a motion-coupled virtual reality system. Displays 61, 101922. https://doi.org/10.1016/j.displa.2019.08.004

Nichols, S., Haldane, C., Wilson, J.R., 2000. Measurement of presence and its consequences in virtual environments. International Journal of Human-Computer Studies 52, 471–491. https://doi.org/10.1006/ijhc.1999.0343

Niehorster, D.C., Li, L., Lappe, M., 2017. The Accuracy and Precision of Position and Orientation Tracking in the HTC Vive Virtual Reality System for Scientific Research. Iperception 8. https://doi.org/10.1177/2041669517708205

Nintendo, 2019. Accessories | Wii | Nintendo UK | Wii | Nintendo [WWW Document]. URL https://www.nintendo.co.uk/Wii/Accessories/Accessories-Wii-Nintendo-UK-626430.html (accessed 5.12.19).

OVERBUILDED, 2020. RollerCoaster VR Universe on Steam [WWW Document]. URL https://store.steampowered.com/app/1070690/RollerCoaster_VR_Universe/ (accessed 3.22.21).

Paas, F., van Merriënboer, J.J.G., 2020. Cognitive-Load Theory: Methods to Manage Working Memory Load in the Learning of Complex Tasks. Curr Dir Psychol Sci 096372142092218. https://doi.org/10.1177/0963721420922183

Palmisano, S., Allison, R.S., Schira, M.M., Barry, R.J., 2015. Future challenges for vection research: definitions, functional significance, measures, and neural bases. Front. Psychol. 6. https://doi.org/10.3389/fpsyg.2015.00193

Palmisano, S., Szalla, L., Kim, J., 2019. Monocular Viewing Protects Against Cybersickness Produced by Head Movements in the Oculus Rift, in: 25th ACM Symposium on Virtual Reality Software and Technology. ACM, Parramatta NSW Australia, pp. 1–2. https://doi.org/10.1145/3359996.3364699

Pane, E.S., Zuhriya Khoirunnisaa, A., Wibawa, A.D., Purnomo, M.H., 2018. Identifying Severity Level of Cybersickness from EEG signals using CN2 Rule Induction Algorithm, in: 2018 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS). IEEE, Bangkok, pp. 170–176. https://doi.org/10.1109/ICIIBMS.2018.8549968

Pao, C., 2020. Improving Head Tracking in XR Environments. CEVA's Experts blog. URL https://www.ceva-dsp.com/ourblog/improving-head-tracking-in-xr-environments/ (accessed 3.13.21).

Parihar, J.K.S., Jain, V.K., Chaturvedi, P., Kaushik, J., Jain, G., Parihar, A.K.S., 2016. Computer and visual display terminals (VDT) vision syndrome (CVDTS). Medical Journal Armed Forces India 72, 270–276. https://doi.org/10.1016/j.mjafi.2016.03.016

Paroz, A., Potter, L.E., 2018. Impact of air flow and a hybrid locomotion system on cybersickness, in: Proceedings of the 30th Australian Conference on Computer-Human Interaction. ACM, Melbourne Australia, pp. 582–586. https://doi.org/10.1145/3292147.3292229

Patney, A., Kim, J., Salvi, M., Kaplanyan, A., Wyman, C., Benty, N., Lefohn, A., Luebke, D., 2016. Perceptually-based foveated virtual reality, in: ACM SIGGRAPH 2016 Emerging Technologies. ACM, Anaheim California, pp. 1–2. https://doi.org/10.1145/2929464.2929472

Peleckis, K., Peleckienė, V., Peleckis, K., Polajeva, T., 2016. Towards sustainable entrepreneurship: role of nonverbal communication in business negotiations. JESI 4, 228–239. https://doi.org/10.9770/jesi.2016.4.2(10)
Pereira Junior, A., 2021. Explaining Dizziness in a Virtual Reality Roller Coaster Ride. Journal of Consciousness Studies 28, 183–198.
Perform Better, 2022. Pasco Force Platforms - Perform Better UK Range [WWW Document]. Perform Better. URL https://performbetter.co.uk/product/pasco-portable-force-platform/ (accessed 7.30.22).
Plouzeau, J., Paillot, D., Chardonnet, J.-R., Merienne, F., 2015. Effect of proprioceptive vibrations on simulator sickness during navigation task in virtual environment.
Pollak, C., Thorpy, M.J., Yager, J., 2010. The Encyclopedia of Sleep and Sleep Disorders. Infobase Publishing.
Porcino, T., Rodrigues, E.O., Bernardini, F., Trevisan, D., Clua, E., 2022. Identifying cybersickness causes in virtual reality games using symbolic machine learning algorithms. Entertainment Computing 41, 100473. https://doi.org/10.1016/j.entcom.2021.100473
Porter, J., Robb, A., 2019. An Analysis of Longitudinal Trends in Consumer Thoughts on Presence and Simulator Sickness in VR Games, in: Proceedings of the Annual Symposium on Computer-Human Interaction in Play. ACM, Barcelona Spain, pp. 277–285. https://doi.org/10.1145/3311350.3347159
Pouke, M., Tiiro, A., LaValle, S.M., Ojala, T., 2018. Effects of Visual Realism and Moving Detail on Cybersickness, in: 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). IEEE, Reutlingen, pp. 665–666. https://doi.org/10.1109/VR.2018.8446078
Purwar, S., 2019. Designing User Experience for Virtual Reality (VR) applications [WWW Document]. Medium. URL https://uxplanet.org/designing-user-experience-for-virtual-reality-vr-applications-fc8e4faadd96 (accessed 3.16.21).
Razzaque, S., Kohn, Z., Whitton, M.C., 2001. Redirected Walking. Eurographics Association.
Reason, J.T., 1978. Motion Sickness Adaptation: A Neural Mismatch Model. J R Soc Med 71, 819–829. https://doi.org/10.1177/014107687807101109
Reason, J.T., Brand, J.J., 1975. Motion Sickness. Academic Press.
Rebenitsch, L., Owen, C., 2016. Review on cybersickness in applications and visual displays. Virtual Reality 20, 101–125. https://doi.org/10.1007/s10055-016-0285-9
Rebenitsch, L., Quinby, B., 2019. Cybersickness and Postural Sway Using HMD Orientation, in: Chen, J.Y.C., Fragomeni, G. (Eds.), Virtual, Augmented and Mixed Reality. Multimodal Interaction, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 500–509. https://doi.org/10.1007/978-3-030-21607-8_39
Regenbrecht, H., Schubert, T., 2002. Real and Illusory Interactions Enhance Presence in Virtual Environments. Presence: Teleoperators and Virtual Environments 11, 425–434. https://doi.org/10.1162/105474602760204318
Reichelt, S., Häussler, R., Fütterer, G., Leister, N., 2010. Depth cues in human visual perception and their realization in 3D displays, in: Javidi, B., Son, J.-Y., Thomas, J.T., Desjardins, D.D. (Eds.), . Orlando, Florida, p. 76900B. https://doi.org/10.1117/12.850094

Remington, L.A., 2012. Clinical Anatomy and Physiology of the Visual System. Elsevier/Butterworth-Heinemann.
Riccio, G.E., Stoffregen, T.A., 1991. An ecological Theory of Motion Sickness and Postural Instability. Ecological Psychology 3, 195–240. https://doi.org/10.1207/s15326969eco0303_2
Riecke, B.E., 2010. Compelling Self-Motion Through Virtual Environments Without Actual Self-Motion – Using Self-Motion Illusions ('Vection') to Improve VR User Experience, in: Kim, J.-J. (Ed.), Virtual Reality. InTech. https://doi.org/10.5772/13150
Riener, R., Harders, M., 2012. VR for Medical Training, in: Riener, R., Harders, M. (Eds.), Virtual Reality in Medicine. Springer, London, pp. 181–210. https://doi.org/10.1007/978-1-4471-4011-5_8
Risi, D., Palmisano, S., 2019a. Effects of postural stability, active control, exposure duration and repeated exposures on HMD induced cybersickness. Displays 60, 9–17. https://doi.org/10.1016/j.displa.2019.08.003
Risi, D., Palmisano, S., 2019b. Can We Predict Susceptibility to Cybersickness?, in: 25th ACM Symposium on Virtual Reality Software and Technology. ACM, Parramatta NSW Australia, pp. 1–2. https://doi.org/10.1145/3359996.3364705
Riva, G., 2005. Virtual Reality in Psychotherapy: Review. CyberPsychology & Behavior 8, 220–230. https://doi.org/10.1089/cpb.2005.8.220
Rosenfield, M., 2011. Computer vision syndrome: a review of ocular causes and potential treatments: Computer vision syndrome. Ophthalmic and Physiological Optics 31, 502–515. https://doi.org/10.1111/j.1475-1313.2011.00834.x
Salomoni, P., Prandi, C., Roccetti, M., Casanova, L., Marchetti, L., Marfia, G., 2017. Diegetic user interfaces for virtual environments with HMDs: a user experience study with oculus rift. Journal on Multimodal User Interfaces 11. https://doi.org/10.1007/s12193-016-0236-5
Schwind, V., Knierim, P., Haas, N., Henze, N., 2019. Using Presence Questionnaires in Virtual Reality, in: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19. ACM Press, Glasgow, Scotland Uk, pp. 1–12. https://doi.org/10.1145/3290605.3300590
Seay, A.F., Krum, D.M., Hodges, L., Ribarsky, W., 2001. Simulator sickness and presence in a high FOV virtual environment, in: Proceedings IEEE Virtual Reality 2001. IEEE Comput. Soc, Yokohama, Japan, pp. 299–300. https://doi.org/10.1109/VR.2001.913806
Sevinc, V., Berkman, M.I., 2020. Psychometric evaluation of Simulator Sickness Questionnaire and its variants as a measure of cybersickness in consumer virtual environments. Applied Ergonomics 82, 102958. https://doi.org/10.1016/j.apergo.2019.102958
Shafer, D.M., Carbonara, C.P., Korpi, M.F., 2019. Factors Affecting Enjoyment of Virtual Reality Games: A Comparison Involving Consumer-Grade Virtual Reality Technology. Games for Health Journal 8, 15–23. https://doi.org/10.1089/g4h.2017.0190
Sharples, S., Cobb, S., Moody, A., Wilson, J.R., 2008. Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems. Displays 29, 58–69. https://doi.org/10.1016/j.displa.2007.09.005

Silva, B.M., Fernando, P., 2018. Early Prediction of Cybersickness in Virtual, Augmented & Mixed Reality Applications: A Review.
Slater, M., Steed, A., 2000. A Virtual Presence Counter. Presence: Teleoperators and Virtual Environments 9, 413–434. https://doi.org/10.1162/105474600566925
Smith, D., Granados, J., Suss, J., 2019. Evaluating the Comprehensiveness of VR PLAY Guidelines Using Elder Scrolls: Skyrim VR. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 63, 2287–2291. https://doi.org/10.1177/1071181319631164
Snellen, H., 1862. Probebuchstaben zur Bestimmung der Sehschärfe. Van de Weijer, Utrecht.
Soames, R.W., Palastanga, N., 2018. Anatomy and Human Movement: Structure and function, 7e, 7th edition. ed. Elsevier, Edinburgh ; New York.
Solimini, A.G., 2013. Are There Side Effects to Watching 3D Movies? A Prospective Crossover Observational Study on Visually Induced Motion Sickness. PLoS ONE 8, e56160. https://doi.org/10.1371/journal.pone.0056160
Song, W., Zhang, K., Sun, J., Ma, L., Jesse, F.F., Teng, X., Zhou, Y., Bao, H., Chen, Shiqing, Wang, S., Yang, B., Chu, X., Ding, W., Du, Y., Cheng, Z., Wu, B., Chen, Shanguang, He, G., He, L., Chen, X., Li, W., 2013. A Simple Spatial Working Memory and Attention Test on Paired Symbols Shows Developmental Deficits in Schizophrenia Patients. Neural Plasticity 2013, 1–7. https://doi.org/10.1155/2013/130642
Stanney, K., Fidopiastis, C., Foster, L., 2020. Virtual Reality Is Sexist: But It Does Not Have to Be. Front. Robot. AI 7, 4. https://doi.org/10.3389/frobt.2020.00004
Stanney, K., Salvendy, G., 1998. Aftereffects and Sense of Presence in Virtual Environments: Formulation of a Research and Development Agenda. International Journal of Human-Computer Interaction 10, 135–187. https://doi.org/10.1207/s15327590ijhc1002_3
Stanney, K.M., Kennedy, R.S., 1997. The psychometrics of cybersickness. Commun. ACM 40, 66–68. https://doi.org/10.1145/257874.257889
Statt, N., 2019. Epic vs. Steam: the console war reimagined on the PC [WWW Document]. The Verge. URL https://www.theverge.com/2019/4/16/18334865/epic-games-store-versus-steam-valve-pc-gaming-console-war-reimagined (accessed 3.11.21).
Stauffert, J.-P., Niebling, F., Latoschik, M.E., 2018. Effects of Latency Jitter on Simulator Sickness in a Search Task, in: 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). IEEE, Reutlingen, pp. 121–127. https://doi.org/10.1109/VR.2018.8446195
steantycip, 2020. The importance of frame rates in VR [WWW Document]. URL https://steantycip.com/blog/the-importance-of-framerates-in-vr/ (accessed 7.30.22).
Stein, S., 2016. 2016 VR, meet 1993 VR [WWW Document]. CNET. URL https://www.cnet.com/news/2016-vr-meet-1993-vr/ (accessed 3.11.21).
Steinicke, F., Visell, Y., Campos, J., Lécuyer, A. (Eds.), 2013. Human Walking in Virtual Environments: Perception, Technology, and Applications. Springer New York, New York, NY. https://doi.org/10.1007/978-1-4419-8432-6

Step, S., 2017. The Science of Locomotion in VR – Usable or not. – Medium [WWW Document]. URL https://medium.com/usable-or-not/the-science-of-locomotion-in-vr-e2314be3f1ca (accessed 4.11.19).
Stoffregen, T.A., Faugloire, E., Yoshida, K., Flanagan, M.B., Merhi, O., 2008. Motion Sickness and Postural Sway in Console Video Games. Hum Factors 50, 322–331. https://doi.org/10.1518/001872008X250755
Stone Iii, W.B., 2017. Psychometric evaluation of the Simulator Sickness Questionnaire as a measure of cybersickness (Doctor of Philosophy). Iowa State University, Digital Repository, Ames. https://doi.org/10.31274/etd-180810-5050
Su, F.-C., Wu, W.-L., Lee, W.-D., 2000. Stance Stability in Shooters. Journal of Medical and Biological Engineering 20, 187–192.
Sutherland, I.E., 1965. The Ultimate Display. Proceedings of the IFIP Congress 506–508.
Sutori, 2020. 1961 - Headsight Developed by Comeau and Bryan, the... | Sutori [WWW Document]. URL https://www.sutori.com/item/1961-headsight-developed-by-comeau-and-bryan-the-headsight-projected-screen-f (accessed 5.14.20).
Sweller, J., 1988. Cognitive Load During Problem Solving: Effects on Learning. Cognitive Science 12, 257–285. https://doi.org/10.1207/s15516709cog1202_4
Swider, M., Pino, N., Leger, H.S., Hood, V., Adam, Vjestica, 2021. PS4 vs Xbox One: which gaming console is better? [WWW Document]. TechRadar. URL https://www.techradar.com/uk/news/gaming/consoles/ps4-vs-xbox-720-which-is-better-1127315 (accessed 3.25.21).
Sylaiou, S., Karoulis, A., Stavropoulos, Y., Patias, P., 2008. Presence-Centered Assessment of Virtual Museums' Technologies. DJLIT 28, 55–62. https://doi.org/10.14429/djlit.28.4.197
Taylor, G.S., Schmidt, C., 2012. Empirical Evaluation of the Emotiv EPOC BCI Headset for the Detection of Mental Actions. Proceedings of the Human Factors and Ergonomics Society Annual Meeting.
Taylor, L.C., Harm, D.L., Kennedy, R.S., Reschke, M.F., Loftin, R.B., 2011. Cybersickness following repeated exposures to dome and HMD virtual environments.
Treisman, M., 1977. Motion sickness: an evolutionary hypothesis. Science 197, 493–495. https://doi.org/10.1126/science.301659
Ujike, H., Ukai, K., Nihei, K., 2008. Survey on motion sickness-like symptoms provoked by viewing a video movie during junior high school class. Displays 29, 81–89. https://doi.org/10.1016/j.displa.2007.09.003
Unity Technologies, 2021. Unity engine visual scripting | Game development software without coding | Unity Bolt | Unity [WWW Document]. URL https://unity.com/products/unity-visual-scripting (accessed 3.18.21).
Valve Corporation, 2020. Steam Hardware & Software Survey [WWW Document]. URL https://store.steampowered.com/hwsurvey/ (accessed 7.20.20).
van der Veen, S.M., Bordeleau, M., Pidcoe, P.E., France, C.R., Thomas, J.S., 2019. Agreement Analysis between Vive and Vicon Systems to Monitor Lumbar Postural Changes. Sensors 19, 3632. https://doi.org/10.3390/s19173632
Vergauwe, E., Dewaele, N., Langerock, N., Barrouillet, P., 2012. Evidence for a central pool of general resources in working memory. Journal of Cognitive Psychology 24, 359–366. https://doi.org/10.1080/20445911.2011.640625

Villard, S.J., Flanagan, M.B., Albanese, G.M., Stoffregen, T.A., 2008. Postural Instability and Motion Sickness in a Virtual Moving Room. Hum Factors 50, 332–345. https://doi.org/10.1518/001872008X250728
Virtual Reality Society, 2017. What is Virtual Reality? - Virtual Reality Society [WWW Document]. URL https://www.vrs.org.uk/virtual-reality/what-is-virtual-reality.html (accessed 4.11.19).
VR comfort rating, 2020. Steam Curator: VR Comfort Rating [WWW Document]. URL https://store.steampowered.com/curator/33961168-VR-Comfort-Rating/ (accessed 2.16.20).
Weaver, T.B., Ma, C., Laing, A.C., 2017. Use of the Nintendo Wii Balance Board for Studying Standing Static Balance Control: Technical Considerations, Force-Plate Congruency, and the Effect of Battery Life. J Appl Biomech 33, 48–55. https://doi.org/10.1123/jab.2015-0295
Weech, S., Moon, J., Troje, N.F., 2018. Influence of bone-conducted vibration on simulator sickness in virtual reality. PLoS ONE 13, e0194137. https://doi.org/10.1371/journal.pone.0194137
Wei, S.-E., Saragih, J., Simon, T., Harley, A.W., Lombardi, S., Perdoch, M., Hypes, A., Wang, D., Badino, H., Sheikh, Y., 2019. VR Facial Animation via Multiview Image Translation. Facebook Research. URL https://research.fb.com/publications/vr-facial-animation-via-multiview-image-translation/ (accessed 3.30.21).
Wertheim, A.H., Bos, J.E., Krul, A.J., TNO Human Factors, 2001. Predicting motion induced vomiting from subjective misery (MISC) ratings obtained in 12 experimental studies. TNO Human Factors, S.l.
Whelan, T.J., 2008. Social Presence in Multi-User Virtual Environments: [WWW Document]. URL https://www.frontiersin.org/articles/10.3389/frobt.2018.00114/full
Wickens, C.D., Carswell, C.M., 2021. Information Processing, in: Handbook of Human Factors and Ergonmics. John Wiley & Sons, Ltd, pp. 114–158. https://doi.org/10.1002/9781119636113.ch5
Widdowson, C., Becerra, I., Merrill, C., Wang, R.F., LaValle, S., 2019. Assessing Postural Instability and Cybersickness Through Linear and Angular Displacement. Hum Factors 001872081988125. https://doi.org/10.1177/0018720819881254
Wienrich, C., Weidner, C.K., Schatto, C., Obremski, D., Israel, J.H., 2018. A Virtual Nose as a Rest-Frame - The Impact on Simulator Sickness and Game Experience, in: 2018 10th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games). IEEE, Wurzburg, pp. 1–8. https://doi.org/10.1109/VS-Games.2018.8493408
Wiltz, C., 2019. The Story of Sega VR: Sega's Failed Virtual Reality Headset [WWW Document]. Design News. URL https://www.designnews.com/electronics-test/story-sega-vr-segas-failed-virtual-reality-headset/74451237860349 (accessed 5.14.20).
Winestock, N., 2018. VR & Diegetic Interfaces - Don't Break the Experience! [WWW Document]. Medium. URL https://uxdesign.cc/vr-diegetic-interfaces-dont-break-the-experience-554f210b6e46 (accessed 3.16.21).
Witmer, B.G., Jerome, C.J., Singer, M.J., 2005. The Factor Structure of the Presence Questionnaire. Presence: Teleoperators and Virtual Environments 14, 298–312. https://doi.org/10.1162/105474605323384654

Witmer, B.G., Singer, M.J., 1998. Measuring Presence in Virtual Environments: A Presence Questionnaire. Presence: Teleoperators and Virtual Environments 7, 225–240. https://doi.org/10.1162/105474698565686
Wobbrock, J.O., Findlater, L., Gergle, D., Higgins, J.J., 2011. The aligned rank transform for nonparametric factorial analyses using only anova procedures, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, Vancouver BC Canada, pp. 143–146. https://doi.org/10.1145/1978942.1978963
Wood, R.W., 1895. The “Haunted Swing” illusion. Psychological Review 2, 277–278. https://doi.org/10.1037/h0073333
Yablonski, J., 2019. Design Principles for Reducing Cognitive Load [WWW Document]. Medium. URL https://blog.prototypr.io/design-principles-for-reducing-cognitive-load-84e82ca61abd (accessed 3.16.21).
Yang, S., Schlieski, T., Selmins, B., Cooper, S.C., Doherty, R.A., Corriveau, P.J., Sheedy, J.E., 2012. Stereoscopic Viewing and Reported Perceived Immersion and Symptoms: Optometry and Vision Science 89, 1068–1080. https://doi.org/10.1097/OPX.0b013e31825da430
Yeung, A.S., Jin, P., Sweller, J., 1998. Cognitive Load and Learner Expertise: Split-Attention and Redundancy Effects in Reading with Explanatory Notes. Contemporary Educational Psychology 23, 1–21. https://doi.org/10.1006/ceps.1997.0951
Yildirim, C., 2020. Don’t make me sick: investigating the incidence of cybersickness in commercial virtual reality headsets. Virtual Reality 24, 231–239. https://doi.org/10.1007/s10055-019-00401-0
Yildirim, C., 2019. Cybersickness during VR gaming undermines game enjoyment: A mediation model. Displays 59, 35–43. https://doi.org/10.1016/j.displa.2019.07.002
Young, S.D., Adelstein, B.D., Ellis, S.R., 2006. Demand Characteristics of a Questionnaire Used to Assess Motion Sickness in a Virtual Environment, in: IEEE Virtual Reality Conference (VR 2006). IEEE, Alexandria, VA, USA, pp. 97–102. https://doi.org/10.1109/VR.2006.44
Zhang, L.-L., Wang, J.-Q., Qi, R.-R., Pan, L.-L., Li, M., Cai, Y.-L., 2016. Motion Sickness: Current Knowledge and Recent Advance. CNS Neurosci Ther 22, 15–24. https://doi.org/10.1111/cns.12468

9 Glossary of Terms

Acronym	Definition	Description
VIMS	Visually Induced Motion Sickness	Another term for cybersickness.
VR	Virtual Reality	A system designed to replace a user's actual environment with a digital one.
AR	Augmented Reality	A system designed to insert virtual elements into a user's real-world environment.
XR	Cross Reality	An umbrella term for all systems aiming to present digital elements into a user's environment.
VE	Virtual Environments	An umbrella term for AR and VR environments.
FOV	Field of View	The amount of viewing angle presented by a display. This measurement is expressed either as a horizontal, vertical pair or as a corner to comer diagonal typically given in degrees.
CAVE	Cave Automatic Virtual Environment	A VR system designed to utilise multiple displays surrounding a user to create a virtual environment.
BOOM	Binocular Omni-Oriented Monitor	A VR system designed to utilise a viewing box suspended on a multilink arm to provide a VR viewing and head tracking experience.
HMD	Head-mounted display	A display that is worn on the head by a user to display XR environments.
SSQ	Simulator Sickness Questionnaire	A method developed by (Kennedy and Lane, 1993) to quantify the severity of MS (Motion Sickness) symptoms, explicitly relating to simulator environments.

SSQ-O	Simulator Sickness Questionnaire – Oculomotor	Abbreviation for the oculomotor components of the SSQ.
SSQ-D	Simulator Sickness Questionnaire – Disorientation	Abbreviation for the disorientation components of the SSQ.
SSQ-N	Simulator Sickness Questionnaire – Nausea	Abbreviation for the nausea components of the SSQ.
SSQ-T	Simulator Sickness Questionnaire – Total	Abbreviation for the Total SSQ Score.
MSSQ	Motion Sickness Susceptibility Questionnaire	A method developed by (Reason and Brand, 1975) establishes a subject's susceptibility to the effects of motion sickness.
MSQ	Pensacola Motion Sickness Questionnaire	A method designed by (Kellogg et al., 1965) to gather qualitative data regarding a subject's motion sickness experience.
VRSQ	Virtual Reality Sickness Questionnaire	A method developed by (Kim et al., 2018) as a refinement to the SSQ, specifically for virtual reality environments.
CSQ	Cybersickness Questionnaire	A method developed by (Stone Iii, 2017) as a refinement to the SSQ for virtual reality environments.
MISC	Misery Score Rating Scale	An 11 point scale for establishing the severity of sickness a subject is experiencing.
NRT	Nausea Rating Test	A method of continuous subjective assessment of current cybersickness severity.
CS	Cybersickness	The term used by researchers to motion sickness-like symptoms experienced during the use of virtual reality environments.

MS	Motion Sickness	
DOF	Degrees of Freedom	A representative measure of the number of axes under which a device can represent motion.
	Simulator Sickness	A set of motion sickness-like symptoms triggered in virtual training apparatus utilising a real-world mock-up of the environment (such as a plane cockpit).
CNS	Central Nervous System	The communication system of the human body controls motor systems within the body.
FMS	Fast Motion Sickness Scale	A scale with values running from 0 – 20 where 0 represents no sickness at all and 20 representing frank sickness.
IVRPA	International Virtual Reality Professionals Association	An international association of photographers who create and produce 360° Panoramas and other Virtual Reality Content.
IPD	Inter-Pupillary Distance	The measured distance between an individual's pupils.
OLED	Organic Light-Emitting Diode	A type of display panel.
LCD	Liquid Crystal Display	A type of display panel.
VAC	Vergence Accommodation Conflict	The term describes the human visual system's difficulty in resolving depth in virtual scenes due to mismatches in the focus and accommodation.
VRET	Virtual reality Exposure Therapy	A form of behaviour therapy using VR experiences in place of situation exposure.
EEG	Electroencephalogram	A medical test for measuring the electrical activity within the brain.

EGG	Electrogastrogram	A medical test for measuring electrical activity in the stomach.
ECG	Electrocardiogram	A medical test for measuring electrical activity in the heart.
SUS	Slater-Usoh-Steed questionnaire	Questionnaire for establishing the amount of presence in a virtual environment.
IPQ	igroup presence questionnaire	Questionnaire for establishing the amount of presence in a virtual environment.

10 Appendices

Appendix A: - Example Forms

Simulator Sickness Questionnaire

Kennedy et al. 1993

Name ____

Date ____

Stimuli ____

Please rate the following symptoms as they are affecting you right now.

	None	Slight	Moderate	Severe
General discomfort				
Fatigue				
Headache				
Eyestrain				
Difficulty focusing				
Salivation increasing				
Sweating				
Nausea				
Difficulty concentrating				
"Fullness of head."				
Blurred Vision				
Dizziness with eyes open				
Dizziness with eyes closed				
Vertigo				
Stomach awareness				
Burping				

NASA TLX

Subject ID ________________ Test Number __________________________

Ratings Sheet

Mental Demand

How much mental and perceptual thinking was required? Was the task easy or demanding?

Low High

Physical Demand

How much physical activity was required? Was the task easy or demanding?

Low High

Temporal Demand

How much time pressure did you feel during the task? Was the rate of the task slow or fast?

Low High

Performance

How successful do you think you were at the task? How satisfied were you with your performance?

Good Poor

Effort

How hard did you have to work, both mentally and physically to achieve your level of performance?

Low High

Frustration

How insecure, discouraged, irritated, stressed and annoyed are you with the task?

Low High

Appendix B: - Experiment Forms

10.1 Briefing

Hello, thank you for agreeing to consider being part of this experiment.

Cybersickness is a significant barrier to the widespread adoption of virtual reality technologies. As part of my PhD studies, this experiment aims to investigate the impact of performing various tasks within a virtual reality environment on the onset and severity of cybersickness.

Participation in the experiment requires two sessions of virtual reality exposure lasting approximately 30 mins each. During each experiment, you will play three minigames designed to test for the factors I am investigating and answer some questions before, during and after the experiment. Each session needs to be separated by a minimum of 24 hours to ensure a return to baseline conditions, although larger gaps can be arranged if convenient. Participants must have normal vision to participate, although corrective measure may be used, such as glasses or contact lenses.

There are some known risks associated with the experiment. You will possibly experience minor discomfort in the form of disorientation, tiredness, fatigue, vertigo, dizziness and nausea. Extreme reactions to cybersickness can cause vomiting. However, this is very rare and is not the objective of the experiment. Symptoms of cybersickness usually subside within 30 minutes of exposure.

With the COVID-19 risk still present, the safety of participants in this experiment is paramount. An extensive hygiene program has been implemented to ensure minimal risk of transmission. Facemasks must be worn throughout the test, and social distancing will be maintained throughout. Virtual reality head-mounted displays will be used during the testing, which will be sanitised between usages and with a minimum of 24 hours left between each headset usage.

If you are interested in participating or have any further questions, please contact p.merritt@derby.ac.uk

Thank you for your time

Patrick Merritt, Associate Lecturer, University of Derby

10.2 Consent

Forename ____

Surname ____

Date ________ /________/________ Participant ID ______________

Hello, thank you for agreeing to participate in this experiment, aiming to establish the impact of performing tasks within a virtual reality environment on cybersickness symptoms onset and severity.

Participation requires 2 sessions lasting about 30 minutes each.

This form serves to inform you of the risks associated with participating in this experiment.

Please read the following statements and sign to acknowledge you understand each risk. If you have any questions or no longer wish to participate, please inform the experiment administrator.

I understand this experiment requires the usage of a visual display technology to participate, therefore I do not suffer from photosensitive epilepsy or any other medical condition which may preclude the usage of a VR Head mounted display or other visual display technology.

I do not suffer from severe claustrophobia.

I do not have a severe uncorrected visual impairment, such as colour blindness or astigmatism which is uncorrected by visual apparatus such as glasses.

I understand that several test will need to be conducted to establish my fitness for testing. This will include: -

> A simple eye test to verify my vision falls within normal parameters for testing.
>
> A measurement of my Interpupillary distance (distance between the eyes) for calibrating the headset.

I understand that the experiment I am participating in has been designed as such that it may induce the symptoms of cybersickness. Common symptoms include nausea, stomach discomfort, dizziness, headache, tiredness, fatigue, light headedness and disorientation. In rare cases vomiting is a possible side effect. These symptoms generally pass within 30 mins of completing the test but may persist longer than this.

I understand I have the right to withdraw from this study at any time, without reason by contacting the experiment administrator. At this time all of my data will be destroyed and no longer considered for future analysis.

I understand anonymised data pertaining to my results may be published in a PhD thesis and / or future publications.

Declaration

I have read the above statements and agree to participate in this experiment having been fully informed of the risks and issues pertaining to data collection. I also fully understand I have the right to withdraw from this study at any time.

Signature ______________________________ Date _____/_____/_____

10.3 Data Collection

Name ______________________________

Age ______________________________

Ethnicity (optional) ______________________________

Gender ______________________________

How many hours do you play video games a week? (circle answer)

<1　　1 – 2　　2 – 3　　3 – 5　　5-10　　10+

Do you have any experience using VR?

None　　Once　　Some　　Often

Tester use only

Date ____/____/____

IPD ______________

Eye test done?　　MSSQ Done?　　Instructions Given?

Test Order　　A / B　　B / A

Test A ID ______________　　Test B ID ______________

Pairwise Comparisons

Category	Maze Task	Shoot Task	Sort Task
Mental			
Physical			
Temporal			
Performance			
Effort			
Frustration			

10.4 De-briefing

Thank you for your participation in this study.

The primary focus of this study was to establish the impact of task difficulty on cybersickness symptoms, the tow sets of games you played were identical in all respects apart from difficulty. The current working theory is that an environment with higher task difficulty occupies a higher amount of working memory and as such leaves less available resources to process the environment leading to higher incidents of cybersickness.

If you have any questions, your instructor will answer them now.

Withdrawal details

To withdraw from the study, please email p.merritt@derby.ac.uk with your name and the date you participated in the study.

Appendix C: - Experiment Results

10.5 Quiet Stand Test-Path Lengths: - Initial Values Task A

Sample Number	X Motion/s	Y Motion/s	Z Motion/s	Roll degrees/s	Pitch degrees/s	Yaw degrees/s	Combined Motion/s	Combined Rotation Sum	SSQ-T Score
P 1	0.671	0.449	0.356	0.973	1.318	0.714	0.991	3.005	3.74
P 2	0.716	0.455	0.287	0.877	0.888	0.896	1.002	2.661	7.48
P 3	0.787	0.584	0.399	1.083	1.497	1.264	1.204	3.844	3.74
P 4	0.948	0.526	0.367	1.461	1.806	1.614	1.290	4.881	3.74
P 5	0.565	0.424	0.294	0.888	0.926	0.912	0.865	2.726	14.96
P 6	0.598	0.410	0.254	0.666	0.705	0.681	0.859	2.052	0.00
P 7	0.595	0.398	0.284	0.601	0.852	0.635	0.863	2.089	14.96
P 9	0.515	0.408	0.278	0.721	0.912	0.721	0.810	2.354	3.74
P 10	0.544	0.448	0.293	0.779	0.898	0.873	0.868	2.550	3.74
P 12	0.655	0.483	0.354	1.065	1.476	1.430	1.001	3.971	44.88
P 13	0.725	0.585	0.486	0.938	2.140	2.126	1.202	5.204	3.74
P 14	0.758	0.484	0.303	0.901	1.032	0.913	1.061	2.846	44.88
P 15	0.759	0.582	0.329	0.784	0.929	0.861	1.141	2.574	0.00
P 16	0.525	0.383	0.283	0.674	0.744	0.569	0.793	1.986	41.14
P 17	0.513	0.425	0.263	0.678	0.775	0.628	0.806	2.081	82.28
P 18	0.641	0.414	0.293	0.793	0.882	0.679	0.922	2.354	11.22
P 20	0.778	0.631	0.390	1.104	1.549	1.283	1.226	3.936	0.00
P 21	0.792	0.520	0.294	0.675	0.886	0.735	1.123	2.296	3.74
P 22	0.786	0.525	0.289	0.746	0.889	0.817	1.115	2.451	14.96
P 25	0.638	0.416	0.275	0.787	0.882	0.940	0.907	2.609	7.48
P 26	0.637	0.468	0.299	0.697	0.928	0.847	0.949	2.472	14.96
P 27	0.588	0.464	0.281	0.908	0.848	0.937	0.905	2.693	22.44
P 28	0.628	0.452	0.321	0.962	1.064	0.889	0.949	2.915	0.00
P 29	0.767	0.635	0.337	0.820	1.216	1.027	1.200	3.062	3.74

10.6 Quiet Stand Test-Path Lengths: - Initial Values Task B

Sample Number	X Motion/s	Y Motion/s	Z Motion/s	Roll degrees/s	Pitch degrees/s	Yaw degrees/s	Combined Motion/s	Combined Rotation Sum	SSQ-T Score
P 1	0.825	0.446	0.327	0.916	1.135	0.834	1.094	2.885	7.48
P 2	0.755	0.443	0.295	0.799	0.946	0.868	1.036	2.613	26.18
P 3	1.035	0.574	0.391	1.126	1.446	1.160	1.393	3.732	0.00
P 4	0.735	0.543	0.419	1.180	1.685	1.602	1.157	4.467	3.74
P 5	0.542	0.455	0.294	0.798	0.901	0.941	0.866	2.640	14.96
P 6	0.576	0.410	0.263	0.645	0.774	0.658	0.848	2.077	3.74
P 7	0.513	0.493	0.297	0.686	0.959	1.132	0.877	2.776	18.70
P 9	0.621	0.427	0.295	0.671	0.923	0.692	0.917	2.286	3.74
P 10	0.540	0.472	0.330	0.727	1.107	0.962	0.902	2.796	11.22
P 12	0.695	0.511	0.369	0.975	1.424	1.061	1.060	3.459	112.20
P 13	0.613	0.523	0.386	0.902	1.618	1.406	1.028	3.925	0.00
P 14	1.026	0.564	0.527	1.015	2.354	2.096	1.464	5.465	29.92
P 15	0.751	0.565	0.336	0.870	1.077	1.014	1.116	2.961	0.00
P 16	0.600	0.404	0.305	0.684	0.818	0.618	0.880	2.120	74.80
P 17	0.500	0.368	0.264	0.729	0.788	0.907	0.757	2.423	67.32
P 18	0.541	0.430	0.276	0.813	0.855	0.835	0.840	2.502	11.22
P 20	0.803	0.565	0.437	1.415	2.114	1.337	1.211	4.866	7.48
P 21	0.789	0.465	0.289	0.688	0.898	0.750	1.078	2.336	0.00
P 22	0.946	0.768	0.295	0.685	0.912	0.726	1.410	2.323	33.66
P 25	0.500	0.399	0.264	0.802	0.752	0.830	0.779	2.383	0.00
P 26	0.606	0.519	0.303	0.726	0.860	0.880	0.967	2.465	18.70
P 27	0.558	0.500	0.287	0.926	0.852	0.830	0.914	2.608	3.74
P 28	0.650	0.428	0.300	0.840	0.897	0.919	0.938	2.657	22.44
P 29	0.721	0.647	0.392	1.011	1.351	1.361	1.191	3.723	3.74

10.7 Quiet Stand Test-Path Lengths: - Post Maze Task A

Sample Number	X Motion/s	Y Motion/s	Z Motion/s	Roll degrees/s	Pitch degrees/s	Yaw degrees/s	Combined Motion/s	Combined Rotation Sum	SSQ-T Score
P 1	0.725	0.400	0.266	0.947	1.024	0.785	0.976	2.756	3.74
P 2	0.614	0.482	0.295	0.825	0.865	0.925	0.940	2.615	7.48
P 3	0.708	0.530	0.350	1.086	1.325	1.205	1.081	3.617	3.74
P 4	0.976	0.525	0.361	1.306	1.501	1.444	1.315	4.251	3.74
P 5	0.595	0.433	0.293	0.788	0.922	0.846	0.893	2.556	14.96
P 6	0.491	0.392	0.261	0.632	0.798	0.657	0.768	2.086	0.00
P 7	0.647	0.404	0.330	0.616	0.963	0.765	0.932	2.343	14.96
P 9	0.600	0.313	0.252	0.672	0.875	0.708	0.809	2.255	3.74
P 10	0.538	0.389	0.268	0.729	0.934	0.787	0.807	2.450	3.74
P 12	0.718	0.495	0.299	0.894	1.382	1.005	1.043	3.281	44.88
P 13	0.651	0.419	0.267	0.650	0.860	0.679	0.921	2.189	3.74
P 14	0.748	0.449	0.322	0.805	1.017	0.899	1.047	2.721	44.88
P 15	0.592	0.532	0.352	1.060	1.252	1.079	0.995	3.391	0.00
P 16	0.600	0.385	0.264	0.614	0.765	0.568	0.852	1.946	41.14
P 17	0.616	0.366	0.257	0.677	0.774	0.684	0.852	2.134	82.28
P 18	0.532	0.396	0.244	0.784	0.807	0.715	0.794	2.306	11.22
P 20	0.647	0.696	0.394	1.113	1.409	1.359	1.183	3.882	0.00
P 21	0.817	0.528	0.309	0.671	0.887	0.780	1.145	2.338	3.74
P 22	0.769	0.510	0.282	0.681	0.756	0.658	1.077	2.095	14.96
P 25	0.518	0.462	0.280	0.817	0.830	0.943	0.849	2.591	7.48
P 26	0.555	0.390	0.278	0.676	0.798	0.859	0.830	2.334	14.96
P 27	0.527	0.345	0.278	0.899	0.988	0.835	0.786	2.722	22.44
P 28	0.678	0.457	0.327	0.885	0.967	0.971	0.995	2.823	0.00
P 29	0.843	0.524	0.335	0.736	1.077	0.826	1.174	2.639	3.74

10.8 Quiet Stand Test-Path Lengths: - Post Maze Task B

Sample Number	X Motion/s	Y Motion/s	Z Motion/s	Roll degrees/s	Pitch degrees/s	Yaw degrees/s	Combined Motion/s	Combined Rotation Sum	SSQ-T Score
P 1	0.662	0.358	0.268	0.965	1.065	0.850	0.890	2.881	7.48
P 2	0.628	0.422	0.292	0.762	0.930	0.926	0.913	2.618	26.18
P 3	1.065	0.680	0.398	0.953	1.368	1.062	1.480	3.382	0.00
P 4	0.854	0.456	0.339	1.247	1.529	1.242	1.157	4.018	3.74
P 5	0.666	0.516	0.359	0.848	1.181	1.128	1.035	3.158	14.96
P 6	0.484	0.375	0.280	0.655	0.812	0.646	0.755	2.113	3.74
P 7	0.553	0.428	0.277	0.604	0.789	0.682	0.846	2.075	18.70
P 9	0.667	0.437	0.317	0.796	1.063	0.904	0.964	2.763	3.74
P 10	0.585	0.430	0.304	0.736	1.089	0.946	0.887	2.771	11.22
P 12	0.873	0.631	0.349	1.207	1.633	1.656	1.279	4.497	112.20
P 13	0.617	0.461	0.341	0.823	1.005	0.868	0.952	2.697	0.00
P 14	0.753	0.405	0.360	1.132	1.432	1.413	1.054	3.976	29.92
P 15	0.680	0.457	0.282	0.711	0.885	0.772	0.977	2.368	0.00
P 16	0.489	0.389	0.296	0.656	0.854	0.600	0.796	2.110	74.80
P 17	0.560	0.405	0.255	0.759	0.817	0.938	0.830	2.514	67.32
P 18	0.579	0.409	0.276	0.717	0.965	0.776	0.859	2.457	11.22
P 20	0.785	0.562	0.393	1.369	1.675	1.498	1.190	4.542	7.48
P 21	0.757	0.480	0.292	0.649	0.771	0.717	1.055	2.136	0.00
P 22	0.706	0.464	0.344	0.602	0.975	0.655	1.027	2.233	33.66
P 25	0.708	0.432	0.279	0.749	0.792	0.759	0.980	2.301	0.00
P 26	0.587	0.506	0.284	0.727	0.755	0.913	0.934	2.395	18.70
P 27	0.745	0.528	0.318	0.947	1.010	0.882	1.089	2.839	3.74
P 28	0.667	0.431	0.331	0.929	1.013	1.018	0.973	2.961	22.44
P 29	0.900	0.638	0.410	0.978	1.342	1.285	1.333	3.605	3.74

10.9 Quiet Stand Test-Path Lengths: - Post Shooting Task A

Sample Number	X Motion/s	Y Motion/s	Z Motion/s	Roll degrees/s	Pitch degrees/s	Yaw degrees/s	Combined Motion/s	Combined Rotation Sum	SSQ-T Score
P 1	0.679	0.497	0.282	1.040	1.088	0.808	1.002	2.936	3.74
P 2	0.561	0.476	0.289	0.918	0.940	0.983	0.890	2.841	7.48
P 3	0.829	0.546	0.418	1.262	1.621	1.249	1.226	4.132	3.74
P 4	1.035	0.616	0.452	1.421	2.089	1.618	1.428	5.128	3.74
P 5	0.584	0.478	0.312	0.965	0.990	0.929	0.918	2.884	14.96
P 6	0.616	0.451	0.257	0.871	0.875	0.729	0.909	2.475	0.00
P 7	0.550	0.382	0.267	0.632	0.703	0.595	0.809	1.930	14.96
P 9	0.668	0.372	0.276	0.788	1.022	0.796	0.910	2.606	3.74
P 10	0.657	0.459	0.382	1.018	1.592	1.230	1.011	3.840	3.74
P 12	0.694	0.488	0.300	0.965	1.133	1.120	1.018	3.218	44.88
P 13	0.582	0.433	0.261	0.713	0.811	0.812	0.866	2.336	3.74
P 14	0.778	0.458	0.320	0.983	1.134	1.172	1.075	3.288	44.88
P 15	0.771	0.545	0.361	1.055	1.130	1.027	1.146	3.212	0.00
P 16	0.581	0.361	0.251	0.669	0.729	0.623	0.818	2.021	41.14
P 17	0.672	0.439	0.268	0.766	0.868	0.866	0.947	2.500	82.28
P 18	0.540	0.551	0.283	1.035	1.297	1.510	0.938	3.842	11.22
P 20	0.869	0.690	0.435	1.269	1.838	1.631	1.361	4.738	0.00
P 21	0.968	0.484	0.314	0.676	1.020	0.795	1.239	2.490	3.74
P 22	0.897	0.524	0.294	1.149	0.895	0.896	1.208	2.940	14.96
P 25	0.500	0.487	0.280	0.828	0.952	0.986	0.853	2.766	7.48
P 26	0.581	0.560	0.336	0.828	0.929	1.029	0.992	2.786	14.96
P 27	0.602	0.416	0.262	0.932	0.774	0.812	0.877	2.517	22.44
P 28	0.664	0.498	0.339	1.128	1.209	1.225	1.021	3.563	0.00
P 29	1.012	0.677	0.345	0.844	1.133	0.866	1.417	2.844	3.74

10.10 Quiet Stand Test-Path Lengths: - Post Shooting Task B

Sample Number	X Motion/s	Y Motion/s	Z Motion/s	Roll degrees/s	Pitch degrees/s	Yaw degrees/s	Combined Motion/s	Combined Rotation Sum	SSQ-T Score
P 1	0.745	0.513	0.343	1.115	1.080	0.888	1.094	3.084	7.48
P 2	0.730	0.434	0.298	0.840	1.076	0.979	1.013	2.895	26.18
P 3	0.953	0.378	0.338	0.951	1.250	0.993	1.181	3.194	0.00
P 4	1.051	0.494	0.412	1.392	1.855	1.444	1.371	4.690	3.74
P 5	0.609	0.481	0.330	0.846	0.935	0.918	0.953	2.699	14.96
P 6	0.622	0.455	0.310	0.796	0.994	0.814	0.939	2.605	3.74
P 7	0.525	0.471	0.365	0.857	1.492	0.990	0.904	3.339	18.70
P 9	0.609	0.421	0.276	0.684	0.846	0.836	0.889	2.365	3.74
P 10	0.825	0.928	0.734	1.359	3.690	3.751	1.659	8.800	11.22
P 12	0.862	0.595	0.380	1.330	1.726	1.644	1.264	4.700	112.20
P 13	0.764	0.539	0.357	1.277	1.630	1.463	1.124	4.370	0.00
P 14	0.797	0.452	0.373	1.017	1.471	1.130	1.114	3.618	29.92
P 15	0.682	0.533	0.326	0.842	1.008	0.868	1.049	2.718	0.00
P 16	0.727	0.457	0.264	0.650	0.822	0.593	1.007	2.064	74.80
P 17	0.598	0.391	0.265	0.814	0.903	0.902	0.855	2.619	67.32
P 18	0.587	0.454	0.249	0.786	0.775	0.730	0.883	2.292	11.22
P 20	0.862	0.685	0.483	1.659	2.112	1.724	1.358	5.495	7.48
P 21	0.915	0.534	0.316	0.780	0.901	0.840	1.236	2.521	0.00
P 22	1.475	0.774	0.389	0.807	1.383	1.068	1.915	3.257	33.66
P 25	0.685	0.460	0.294	0.851	0.854	0.942	0.989	2.646	0.00
P 26	0.779	0.572	0.374	0.854	1.314	1.036	1.173	3.204	18.70
P 27	0.565	0.501	0.287	1.077	0.939	0.841	0.919	2.857	3.74
P 28	0.765	0.466	0.338	1.031	1.144	1.199	1.078	3.374	22.44
P 29	0.646	0.535	0.368	0.911	1.363	1.243	1.036	3.516	3.74

10.11 Quiet Stand Test-Path Lengths: - Post Sorting Task A

Sample Number	X Motion/s	Y Motion/s	Z Motion/s	Roll degrees/s	Pitch degrees/s	Yaw degrees/s	Combined Motion/s	Combined Rotation Sum	SSQ-T Score
P 1	0.660	0.453	0.278	0.948	0.973	0.913	0.955	2.834	3.74
P 2	0.532	0.456	0.298	0.902	0.879	1.009	0.857	2.789	7.48
P 3	0.820	0.564	0.477	1.311	1.711	1.195	1.260	4.218	3.74
P 4	1.061	0.592	0.360	1.483	1.516	1.800	1.410	4.799	3.74
P 5	0.636	0.456	0.329	0.852	1.082	1.000	0.952	2.934	14.96
P 6	0.524	0.443	0.267	0.794	0.840	0.855	0.831	2.489	0.00
P 7	0.510	0.443	0.277	0.658	0.743	0.771	0.819	2.172	14.96
P 9	0.657	0.414	0.268	0.850	1.006	0.838	0.929	2.694	3.74
P 10	0.652	0.461	0.319	0.973	1.192	1.174	0.974	3.339	3.74
P 12	0.650	0.483	0.304	0.943	1.133	1.137	0.990	3.213	44.88
P 13	0.695	0.450	0.285	0.817	0.983	0.824	0.983	2.624	3.74
P 14	0.629	0.480	0.361	0.871	1.172	1.073	0.993	3.116	44.88
P 15	0.963	0.670	0.438	1.237	1.672	1.546	1.425	4.455	0.00
P 16	0.715	0.471	0.298	0.799	0.951	0.724	1.023	2.474	41.14
P 17	0.617	0.424	0.282	0.774	1.073	0.889	0.904	2.737	82.28
P 18	0.632	0.462	0.292	0.951	1.250	1.023	0.937	3.223	11.22
P 20	0.863	0.555	0.386	1.195	1.606	1.289	1.243	4.091	0.00
P 21	0.831	0.493	0.353	0.780	0.996	0.929	1.146	2.705	3.74
P 22	0.722	0.500	0.302	0.671	0.819	0.769	1.039	2.259	14.96
P 25	0.616	0.464	0.279	0.837	0.822	0.855	0.924	2.514	7.48
P 26	0.671	0.556	0.365	0.772	0.828	0.985	1.061	2.585	14.96
P 27	0.544	0.412	0.254	1.057	0.780	0.824	0.826	2.660	22.44
P 28	0.686	0.507	0.351	1.220	1.366	1.517	1.043	4.103	0.00
P 29	0.826	0.539	0.320	0.841	1.180	0.972	1.163	2.994	3.74

10.12 Quiet Stand Test-Path Lengths: - Post Sorting Task B

Sample Number	X Motion/s	Y Motion/s	Z Motion/s	Roll degrees/s	Pitch degrees/s	Yaw degrees/s	Combined Motion/s	Combined Rotation Sum	SSQ-T Score
P 1	0.834	0.584	0.329	1.085	1.113	0.987	1.207	3.184	7.48
P 2	0.629	0.476	0.304	1.011	1.067	1.002	0.948	3.080	26.18
P 3	1.016	0.543	0.393	1.232	1.358	1.411	1.366	4.002	0.00
P 4	0.932	0.607	0.437	1.661	2.000	1.591	1.350	5.252	3.74
P 5	0.672	0.466	0.300	0.879	0.855	0.988	0.980	2.722	14.96
P 6	0.646	0.460	0.270	0.829	0.986	0.829	0.942	2.643	3.74
P 7	0.547	0.461	0.294	0.722	0.781	0.813	0.868	2.316	18.70
P 9	0.607	0.427	0.279	0.776	0.859	0.889	0.896	2.524	3.74
P 10	0.891	0.646	0.436	1.089	1.964	2.062	1.356	5.116	11.22
P 12	1.182	0.695	0.472	1.667	2.392	2.020	1.626	6.079	112.20
P 13	0.676	0.459	0.343	0.840	1.216	0.941	0.998	2.997	0.00
P 14	0.776	0.501	0.357	1.059	1.458	1.328	1.115	3.845	29.92
P 15	0.678	0.550	0.336	0.942	1.182	1.028	1.061	3.152	0.00
P 16	0.755	0.463	0.290	0.796	0.875	0.683	1.052	2.354	74.80
P 17	0.628	0.390	0.275	0.842	0.966	1.063	0.884	2.870	67.32
P 18	0.652	0.347	0.271	0.806	0.785	0.811	0.879	2.401	11.22
P 20	0.842	0.648	0.465	1.595	1.978	1.979	1.318	5.553	7.48
P 21	0.722	0.599	0.346	0.864	1.104	1.019	1.140	2.987	0.00
P 22	1.270	0.858	0.416	1.117	1.686	1.690	1.805	4.493	33.66
P 25	0.581	0.422	0.303	0.854	0.958	0.876	0.882	2.687	0.00
P 26	0.644	0.578	0.359	0.926	1.036	1.050	1.060	3.011	18.70
P 27	0.855	0.558	0.340	1.149	1.320	0.899	1.206	3.369	3.74
P 28	0.739	0.435	0.326	0.983	1.027	1.016	1.025	3.026	22.44
P 29	0.918	0.755	0.468	1.180	1.651	1.525	1.451	4.356	3.74

10.13 Quiet Stand Test-Path Lengths: - Means and STD

		Test Scenario							
		Start A	Start B	Maze A	Maze B	Shoot A	Shoot B	Sort A	Sort B
X Motion / s	Mean	0.672003	0.685035	0.654409	0.690463	0.703812	0.765704	0.696323	0.77881
	STD	0.110299	0.158374	0.117349	0.136772	0.156156	0.200544	0.136769	0.184765
Y Motion / s	Mean	0.481992	0.49656	0.450915	0.470845	0.495341	0.521815	0.489455	0.53861
	STD	0.075012	0.089649	0.083036	0.084263	0.08445	0.123387	0.061988	0.121476
Z Motion / s	Mean	0.317062	0.330928	0.298594	0.318439	0.316093	0.352876	0.322617	0.350276
	STD	0.053046	0.065419	0.039717	0.043757	0.057608	0.097012	0.05537	0.066516
Roll Motion / s	Mean	0.857498	0.85939	0.815084	0.855077	0.94807	0.980177	0.938977	1.03764
	STD	0.190622	0.18895	0.181043	0.210775	0.202651	0.254758	0.209798	0.271341
Pitch Motion / s	Mean	1.084997	1.14354	0.990748	1.072974	1.11555	1.315146	1.107234	1.275666
	STD	0.364266	0.430933	0.221778	0.280332	0.350098	0.619743	0.286889	0.447173
Yaw Motion / s	Mean	0.957973	1.01743	0.874282	0.96403	1.012779	1.159734	1.037924	1.187467
	STD	0.35924	0.340874	0.218858	0.280263	0.285155	0.619499	0.27072	0.412374
Combined X, Y, Z Motion / s	Mean	1.002235	1.03003	0.961035	1.010628	1.036589	1.125107	1.028662	1.142313
	STD	0.151779	0.19657	0.147703	0.175573	0.186566	0.252406	0.168401	0.250857
Sum Rotation Motion / s	Mean	2.900468	3.020361	2.680115	2.89208	3.0764	3.455058	3.084135	3.500772
	STD	0.861117	0.90337	0.597466	0.746197	0.793643	1.417414	0.724201	1.089397
Mean SSQT	Mean	14.64833	19.79083	14.64833	19.79083	14.64833	19.79083	14.64833	19.79083
	STD	19.87965	27.88919	19.87965	27.88919	19.87965	27.88919	19.87965	27.88919

Appendix D:- Number of Participants per Study

Study name	Author	Number of participants
Correlating reaction time and nausea measures with traditional measures of cybersickness.	Nesbitt et.al.	24
A comparative study of cybersickness during exposure to virtual reality and “classic” motion sickness: are they different?	Gavgani, et.al.	30
A Metric to Quantify Virtual Scene Movement for the Study of Cybersickness: Definition, Implementation, and Verification.	So, et.al.	36
A Study on Cybersickness Reduction Method using Oculomotor Exercise.	Ho Kim, et.al.	14
Assessing Postural Instability and Cybersickness Through Linear and Angular Displacement.	Widdowson, et.al.	24
Automatic Prediction of Cybersickness for Virtual Reality Games.	Jin, et.al.	24
Effects of postural stability, active control, exposure duration and repeated exposures on HMD induced cybersickness	Risi and Palmisano	20
Combined Pitch and Roll and Cybersickness in a Virtual Environment.	Bonato, Bubka and Palmisano	19
Cybersickness during VR gaming undermines game enjoyment: A mediation model	Yildirim, Caglar	32

Cybersickness in the presence of scene rotational movements along different axes	Lo and So	16
Cybersickness without the wobble: Experimental results speak against postural instability theory	Dennison and D’Zmura	15
Don’t make me sick: investigating the incidence of cybersickness in commercial virtual reality headsets	Yildirim, Caglar	Experiment 1) 45 Experiment 2) 35
Effect of economically friendly acustimulation approach against cybersickness in video-watching tasks using consumer virtual reality devices	Liu, et.al.	29
Effect of Visual Realism on Cybersickness in Virtual Reality	Arttu Tiiro	33 phase 1 20 phase 2
Effects of Presence on Causing Cybersickness in the Elderly within a 3D Virtual Store	Jacko, et.al.	60
Effects of steering locomotion and teleporting on cybersickness and presence in HMD-based virtual reality	Clifton and Palmisano	25
Effects of Visual Realism and Moving Detail on Cybersickness	Pouke, et.al.	25
Identifying Cybersickness through Heart Rate Variability alterations	Garcia-Agundez, et.al.	13
Identifying Severity Level of Cybersickness from EEG signals using CN2 Rule Induction Algorithm	Pane, et.al	9
Impact of air flow and a hybrid locomotion system on cybersickness	Paroz and Potter	12
Monocular Viewing Protects Against Cybersickness Produced by Head Movements in the Oculus Rift	Palmisano, Szalla and Kim	14

Postural stability predicts the likelihood of cybersickness in active HMD-based virtual reality	Arcioni, et.al.	21
Profiling subjective symptoms and autonomic changes associated with cybersickness	Gavgani, et.al	14
Rotation Blurring: Use of Artificial Blurring to Reduce Cybersickness in Virtual Reality First Person Shooters	Budhiraja, et.al	15
Use of physiological signals to predict cybersickness	Dennison, Wisti and D’Zmura	20
Viewpoint Snapping to Reduce Cybersickness in Virtual Reality	Farmani and Teather	12
VR after effect and the relation of cybersickness and cognitive performance	Mittelstaedt, Wacker and Stelling	60

Appendix E:- Ethical Approval

Ethics ETH2021-0165: Patrick Merritt

Date Created	29 Sep 2020
Date Submitted	29 Sep 2020
Date forwarded to committee	29 Sep 2020
Researcher	Patrick Merritt
Student ID	
Category	Postgraduate research student
Supervisor	Chris Windmill
Project	Evaluation of detecting cybersickness via VR HMD positional measurements under realistic usage conditions.
College	College of Science and Engineering
Current status	Approved

Ethics application

Project information

Project title

Cybersickness in Virtual Enviroments

What is the aim of your study?

To identify the role of cognitive load in the onset of cybersickness, and detect the onset of cybersickness in sufferers during VR usage.

What are the objectives for your study?

1) Identify the impact of cognitive load during VR usage.

2) Identify factors useable to identify onset of cybersickness in subjects from headset data.

Are there any research partners (specific staff members) within the University of Derby involved in the project?

Are there any research partners external to the University of Derby involved in the project?

No

Initial screening

Does this project involve human participants? Yes

If yes, should your research adhere to the British Psychology Society (BPS) code of ethics and conduct?

No

Does your study involve data collection with any persons who could be considered vulnerable (under 18 years or the elderly, or those with physical or mental disabilities)?

No

Does your project involve collecting data within NHS organisations or from any NHS employees or patients? No

Does it involve collecting or analysing primary or unpublished data about people who have died, other than data that is already in the public domain?

No

Does your study involve direct access to an external organisation?

No

Does your study involve species not covered by the Animals Scientific Procedures Act (1993)? No

Does your study involve ionising radiation?

No

Does your study involve the evaluation of medical devices, or the testing of medicinal and pharmaceutical products? No

Does your study involve Her Majesty's Prison and Probation Service?

No

Does your study involve serving offenders, professionals who work with them, or questions relating to criminal offences?

No

Does your study involve a need to see, acquire or store material that could be viewed as illegal or that may attract the interest of the police, security or intelligence services?

No

Will your study have any impact on the natural or built environment?

No

Funding and previous applications

Has this research been funded by an external organisation (e.g. a research council or public sector body)?

No

If yes, please provide the name of funder:

Has this research been funded internally?

No

Name of internal fund

Funding amount

Term of funding

Date funding agreed

Have you submitted previous requests for ethical approval to the Committee that relate to this research project? Yes

If yes, please provide previous application reference:

ETH1920-2451

Study

Brief review of relevant literature and rationale for study

Cybersickness is widely understood to be a form of motion sickness affecting users of computerised systems, particularly virtual reality systems (Gianaros & Stern, 2010) (Mazloumi Gavgani et al., 2018). The physiological reasons for humans getting motion sick is currently unknown, the currently accepted theory is Sensory Conflict Theory (Oman, 1990; Reason, 1978; Zhang et al., 2016), where the passive movements experienced by a subject are not matched by vestibular and visual systems. The exact cause of Cybersickness or “trigger” for an individual to suffer from cyber sickness is different, while some users immediately get sick in VR environments others can handle extreme experiences without being phased.

Developers of XR equipment are well aware of the potential negative impact their systems may have on users, and as the full impact of this is not fully understood guidelines tend to be aggressive (Lewis, 2015; Oculus, 2020; Valve Corporation, 2019), advising the immediate discontinuation of using the device upon the first signs of cybersickness symptoms. However these advice guidelines are inconsistent with differing suggestions offered by the major manufacturers as to usage limits and not generally based upon fact.

Recently (Mittelstaedt et al., 2019) demonstrated the impact of utilising different VR display methods (Large VR TV display and VR HMD devices) and differing control methodologies (Bike locomotion and traditional game pad) on task effectiveness after VR exposure. Deficiencies were observed in Reaction times after VR immersion; however, the severity of this deterioration was found to have low correlation with severity of cybersickness symptoms experienced by the subject. Previous studies such as (Nalivaiko et al., 2015; Nesbitt et al., 2017) have previously suggested that cybersickness causes the detrimental impacts on cognitive functions. This is reinforced by (Mittelstaedt et al., 2019) which shows that the utilisation of the VR headset impacts cognitive performance, in many areas such as reaction time (Nalivaiko et al., 2015) and mental rotation (Levine & Stern, 2002).

This experiment aims to identify the impact of cognitive load on cybersickness onset, allowing a determination to be made as to whether task load influences the rate at which individuals get sick. This will provide evidence as to whether normalization of the task in VR Cybersickness experiments is necessary to obtain a consistent result. Further more analysis of headset motion will allow for determination if the onset of sickness can be predicted from variations in user stability. this data will be used to propose an algorithm for real time evaluation of cybersickness onset.

Cited references for any sources in the sections on rationale, methods etc.

Gianaros, P. J., & Stern, R. M. (2010). A Questionnaire for the Assessment of the Multiple Dimensions of Motion Sickness. 10.

Lewis, T. (2015, March 2). Samsung Gear VR: Virtual Reality Tech May Have Nasty Side Effects | Live Science. https://www.livescience.com/49669-virtual-reality-health-effects.html

Mazloumi Gavgani, A., Walker, F. R., Hodgson, D. M., & Nalivaiko, E. (2018). A comparative study of cybersickness during exposure to virtual reality and “classic” motion sickness: are they different? Journal of Applied Physiology, 125(6), 1670–1680. https://doi.org/10.1152/japplphysiol.00338.2018

Oculus. (2020). Oculus, Health and Safety Warnings. https://www.oculus.com/legal/health-andsafety-warnings/?locale=en_GB

Oman, C. M. (1990). Oman-1990-Motion-sickness-a-synthesis-and-eva.pdf. CANADIAN JOURNAL OF PHYSIOLOGY AND PHARMACOLOGY, 68(2), 294–303. https://doi.org/10.1139/y90-044

Reason, J. T. (1978). Motion Sickness Adaptation: A Neural Mismatch Model <sup/>. Journal of the Royal Society of Medicine, 71(11), 819–829. https://doi.org/10.1177/014107687807101109

Valve Corporation. (2019). Safety, Maintenece, Regulatory and Warranty & Agreement Information. https://steamcdn-a.akamaihd.net/store/valve_index/ValveIndexSafetyAndMaintenanceInfo.pdf

Zhang, L.-L., Wang, J.-Q., Qi, R.-R., Pan, L.-L., Li, M., & Cai, Y.-L. (2016). Motion Sickness: Current Knowledge and Recent Advance. CNS Neuroscience & Therapeutics, 22(1), 15–24. https://doi.org/10.1111/cns.12468

Levine, M. E., & Stern, R. M. (2002). Spatial Task Performance, Sex Differences, and Motion Sickness Susceptibility. Perceptual and Motor Skills, 95(2), 425–431. https://doi.org/10.2466/pms.2002.95.2.425

Mittelstaedt, J. M., Wacker, J., & Stelling, D. (2019). VR aftereffect and the relation of cybersickness and cognitive performance. Virtual Reality, 23(2), 143–154. https://doi.org/10.1007/s10055-018-03703

Nalivaiko, E., Davis, S. L., Blackmore, K. L., & Nesbitt, K. V. (2015). Cybersickness provoked by head-mounted display affects cutaneous vascular tone, heart rate and reaction time. Autonomic Neuroscience, 192, 63. https://doi.org/10.1016/j.autneu.2015.07.032

Nesbitt, K., Davis, S., Blackmore, K., & Nalivaiko, E. (2017). Correlating reaction time and nausea measures with traditional measures of cybersickness. Displays, 48, 1–8. https://doi.org/10.1016/j.displa.2017.01.002

Outline of study design

Subjects will be asked to perform several small tasks in a VR environment over 2 sessions with differing levels of task complexity in each.

Before testing users will be informed as to the intent of the test and have the full process explained to them in writing. At this point informed consent will be obtained.

Users will then be screened for uncorrected visual deficiencies which may interfere with the test results, through simple non invasive eye tests and verbal questioning. Subjects demonstrating an unsuitable visual quality or known defect will be discarded and not included in the study. at this point a measurement will be taken for the inter pupillary distance to allow calibration of the headset.

From here participants will fill out the MSSQ to establish a history of motion sickness, and instructions will be given for the test. Subjects will be remined at this point to stop the test should they become unwell to the point they cannot continue.

Subjects will then be exposed to 2 sets of 3 tasks (Task A and Task B). Tasks in groups A and B will be functionally identical however the tasks in group b will have their cognitive demands increased.

Task 1 is a memory test requiring participants to memorise a code, navigate a short maze and enter the code on the other side. Task 2 is a simple shooting gallery requiring subjects to shoot targets matching a sample. Task 3 requires the subject to sort shapes into bins by throwing them. After each task, and before the first task, subjects will be required to stand motionless to provide a period to capture balance data. after each task subjects will provide NASA TLX data to asses the difficulty of the task they just completed.

Pre and post test, users will record balance data by standing on a balance board for a duration of 30 seconds, once with eyes open once with eyes closed. subjects will also complete the SSQ pre and post experiment to verify the amount of sickness induced by the enviroment.

Outline of study methods

The Motion Sickness Susceptibility Questionnaire (MSSQ) is a widely used questionnaire aimed to establish a user's susceptibility to motion sickness. The Simulator Sickness Questionnaire (SSQ) is the standard method used in VR cybersickness studies to establish the intensity of cybersickness symptoms a user is experiencing.

NASA TLX data will be captured to assess the difficulty of the tasks the user has performed.

The headset will be recording the position of the subject throughout the experiment to monitor the subjects balance and posture throughout. A back senor will also be used for this purpose.

Postural data will be captured by using a balance board to assess the subjects pre and post balance stability.

Please provide a detailed description of the study sample, covering recruitment, selection, number, age and if appropriate, inclusion and exclusion criteria.

Study sample will be healthy adults aged 18 – 65, with normal vision with or without corrective methods such as glasses or contact lenses.

Are payments or rewards/incentives (e.g. participant points) going to be made to the participants?

No

If yes, please provide details

Do you propose to carry out your project partly in a non-English language?

No

If yes, please provide details

Ethical considerations

Consent

Full consent and warning as to risks of the study (nausea and possible vomiting) will be disclosed at the start of the study

Deception

No deception will be performed

Debriefing

Full disclosure of the purpose of the study will be disclosed at the end of the study.

Withdrawal from the investigation

Users will have the right to withdraw at anytime. At this time results from the subject will be destroyed, and not used within any future publications. Previously published results cannot be recalled and thus will be unaffected by withdrawal.

Anonymity and confidentiality

All data will be stored anonymously, and securely in line with data protection act requirements.

Protection of participants

A full risk assessment of the test environment will be conducted before testing commences.

Experiment includes a known risk of vomiting, SOP and risk assessment will fully document how to manage this and will be in place before study commences in line with sports science maximal exercise testing.

No known risks regarding psychological issues.

Subjects will be wearing HMD's and therefore effectively blind in some cases during the experiment.

Therefore subjects will be isolated during testing and seated to prevent potential accidents. researchers will be on hand to assist in cases of device malfunction or emergency.

Observation research

Subjects will be observed by researchers but not filmed, this is to ensure safety of the participants.

Giving advice

Advice will not be given during the testing.

Research undertaken in public places

N/a research will be conducted under isolated controlled conditions.

GDPR - collecting personal data

All data will be stored anonymously, and securely in line with data protection act requirements.

No unnecessary data will be stored and all data will be destroyed 7 years after study finishes.

Basis for collecting data

Consent

Data retention

No unnecessary data will be stored and all data will be destroyed 7 years after study finishes.

The data will be stored on a secure drive and encrypted when not being analysed.

Data will not be shared.

Rights of data subject

Subjects will be informed as to the risks of the study and the purpose of the study before beginning.

The subject will be allowed to access and view their data on request.

Any data corrections required will be performed upon request.

All data related to a subject will be deleted on request. data already published will be unaffected by this request

Requests to restrict processing will be taken as requests to withdraw.

Data can be made readable via documentation

Objections of participants will be treated as withdrawal requests.

Commercial sensitivity N / A

Are you using non-standard software to store or analyse data?

No

Are there other ethical implications that are additional to this list?

No

If yes, please provide details

Have/do you intend to request ethical approval from any other body/organisation?

No

If yes, please provide details

Do you intend to publish your research?

Yes

Have the activities associated with this research project been risk-assessed?

Yes

Appendix F:- Parmetric Test Results

10.14 Test A SSQ-T

	Error Df	Df.res	F value	Pr(>F)
1 TestOrder	SmplN 1	12	1.1751	0.2996534
2 SickorWell	SmplN 1	12	12.5218	0.0040812 **
3 UserExp	SmplN 3	12	3.5135	0.0491462 *
4 TestOrder:SickorWell	SmplN 1	12	1.2329	0.2885973
5 TestOrder:UserExp	SmplN 3	12	3.4686	0.0508154 .
6 SickorWell:UserExp	SmplN 2	12	1.1660	0.3445313

10.15 Test A Start Quiet Stand

	Error Df	Df.res	F value	Pr(>F)
1 TestOrder	SmplN 1	12	0.01605	0.90129
2 SickorWell	SmplN 1	12	2.49859	0.13993
3 UserExp	SmplN 3	12	0.34575	0.79287
4 TestOrder:SickorWell	SmplN 1	12	0.65015	0.43575
5 TestOrder:UserExp	SmplN 3	12	1.49666	0.26542
6 SickorWell:UserExp	SmplN 2	12	1.28205	0.31288

10.16 Test A Maze Quiet Stand

	Error Df	Df.res	F value	Pr(>F)
1 TestOrder	SmplN 1	12	0.42406	0.52719
2 SickorWell	SmplN 1	12	1.63364	0.22537
3 UserExp	SmplN 3	12	1.87324	0.18792
4 TestOrder:SickorWell	SmplN 1	12	2.01074	0.18163
5 TestOrder:UserExp	SmplN 3	12	3.28128	0.05852 .
6 SickorWell:UserExp	SmplN 2	12	1.16074	0.34607

10.17 Test A Shoot Quiet Stand

	Error Df	Df.res	F value	Pr(>F)
1 TestOrder	SmplN 1	12	0.683820	0.42441
2 SickorWell	SmplN 1	12	1.064272	0.32259
3 UserExp	SmplN 3	12	0.436259	0.73108
4 TestOrder:SickorWell	SmplN 1	12	0.044677	0.83615
5 TestOrder:UserExp	SmplN 3	12	2.094547	0.15442
6 SickorWell:UserExp	SmplN 2	12	0.718887	0.50714

10.18 Test A Sort Quiet Stand

	Error Df	Df.res	F value	Pr(>F)
1 TestOrder	SmplN 1	12	0.18593	0.67397
2 SickorWell	SmplN 1	12	1.72981	0.21301
3 UserExp	SmplN 3	12	0.41732	0.74380
4 TestOrder:SickorWell	SmplN 1	12	1.25774	0.28402
5 TestOrder:UserExp	SmplN 3	12	1.46138	0.27433
6 SickorWell:UserExp	SmplN 2	12	0.31843	0.73325

10.19 Test B SSQ-T

	Error Df	Df.res	F value	Pr(>F)
1 TestOrder	SmplN 1	12	14.2700	0.0026352 **
2 SickorWell	SmplN 1	12	15.3258	0.0020543 **
3 UserExp	SmplN 3	12	4.0341	0.0337722 *
4 TestOrder:SickorWell	SmplN 1	12	1.0228	0.3318141
5 TestOrder:UserExp	SmplN 2	12	1.3301	0.3007767
6 SickorWell:UserExp	SmplN 3	12	2.0240	0.1643115

10.20 Test B Start Quiet Stand

	Error Df	Df.res	F value	Pr(>F)
1 TestOrder	SmplN 1	12	0.680170	0.42562
2 SickorWell	SmplN 1	12	1.179127	0.29887
3 UserExp	SmplN 3	12	0.087072	0.96578
4 TestOrder:SickorWell	SmplN 1	12	2.278179	0.15708
5 TestOrder:UserExp	SmplN 2	12	1.211404	0.33173
6 SickorWell:UserExp	SmplN 3	12	1.527167	0.25797

10.21 Test B Maze Quiet Stand

	Error Df	Df.res	F value	Pr(>F)
1 TestOrder	SmplN 1	12	0.19541	0.66631
2 SickorWell	SmplN 1	12	0.77117	0.39710
3 UserExp	SmplN 3	12	1.70813	0.21827
4 TestOrder:SickorWell	SmplN 1	12	2.02840	0.17987
5 TestOrder:UserExp	SmplN 2	12	2.66578	0.11017
6 SickorWell:UserExp	SmplN 3	12	0.77635	0.52939

10.22 Test B Shoot Quiet Stand

	Error Df	Df.res	F value	Pr(>F)
1 TestOrder	SmplN 1	12	6.02849	0.030298 *
2 SickorWell	SmplN 1	12	0.73009	0.409589
3 UserExp	SmplN 3	12	1.13505	0.374131
4 TestOrder:SickorWell	SmplN 1	12	0.35632	0.561650
5 TestOrder:UserExp	SmplN 2	12	1.33539	0.299481
6 SickorWell:UserExp	SmplN 3	12	0.48350	0.699934

10.23 Test B Sort Quiet Stand

	Error Df	Df.res	F value	Pr(>F)
1 TestOrder	SmplN 1	12	2.70818	0.125755
2 SickorWell	SmplN 1	12	0.81458	0.384529
3 UserExp	SmplN 3	12	0.53824	0.665006
4 TestOrder:SickorWell	SmplN 1	12	0.96920	0.344313
5 TestOrder:UserExp	SmplN 2	12	3.38421	0.068316 .
6 SickorWell:UserExp	SmplN 3	12	0.81750	0.508703