The role of multisensory feedback in the objective and subjective evaluations of fidelity in virtual reality environments.

Thesis submitted in accordance with the requirements of the
University of Liverpool for the degree of
Doctor of Philosophy by

Natalia Cooper

April 2017

Supervisors: Dr Georg Meyer, School of Psychology
Dr Mark White, School of Engineering

Outcomes of this thesis

Awards

The Best Paper Award (October 2015) - European Conference on Virtual Reality (EURO VR) 2015 in Lecco, Italy. The Effects of Multisensory Cues on the Sense of Presence and Task Performance in a Virtual Reality Environment.

Best Online Poster Award (June 2015) - University of Liverpool, Faculty of Health and Life Sciences. The Effects of Multisensory Cues on the Sense of Presence and Task Performance in a Virtual Reality Environment.

Journal Publications

{Chapter 3] Cooper N, Cant I, White M and Meyer G. *But My Eyes are Fixed on You: Visually Evoked Postural Responses are Modulated by the Perceived Stability of Foreground Objects.* Submitted to Plos ONE

[Chapter 4] Cooper N, Millela F, Cant I, Pinto C, White M. and Meyer G. *The Effects of Substitute Multisensory Cues on the Task Performance and Sense of Presence in Virtual Environment.* Submitted to Plos ONE

[Chapter 4] Cooper N, Millela F, Cant I, Pinto C, White M. and Meyer G. *Track Me if You Can: Substitute Multisensory Cues Modulate Sickness During Inaccurate Motion Tracking.* In preparation for submission to Plos ONE.

[Chapter 5] Cooper N, Millela F, Cant I, Pinto C, White M. and Meyer G. *Virtually Tyred: User Experience and Sensory Feedback Modulate Training Transfer from Virtual to Real Environments*. Submitted to Manufacturing and Virtuality workshop, 29-30th June 2017; publication in the International Journal on Interactive Design and Manufacturing (in preparation)

[Chapter 6] Cooper N, Roscoe J, White M and Meyer G. *Side to side: Motion Cuing and Cognitive Workload Shape User Experience in VR.* In preparation for submission to Journal of Aviation Psychology.

Conference presentations (posters and talks)

[Poster] *15th IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2016, Mexico* **-** poster presentation titled: Augmented Cues in Virtual Environment Facilitates Learning Transfer.

[Poster] *European Conference on Visual Perception (ECVP) 2016, Barcelona, Spain* – poster presentation titled: Virtual Training with Augmented Cues Facilitates Learning Transfer from Virtual to Real Environment.

[Poster] *Eurohaptics 2016, London, UK* - poster presentation tilted: Tactile Feedback Facilitates Performance and Presence.

[Talk] *European Conference on Virtual Reality (EURO VR) 2015, Lecco, Italy* - presented talk and poster titled: The Effects of Multisensory Cues on the Sense of Presence and Task Performance in a Virtual Reality Environment.

[Poster] *International Multisensory Research Forum (IMRF) 2015, Pisa, Italy* - presented poster titled Multisensory Cues in VR Facilitate the Sense of Presence and Performance.

[Talk] *Ergonomics and Human Factor conference (IEHF) 2015* - presented talk at Doctoral Consortium and a poster at the conference titled: Additional Sensory Cues Facilitate on Presence and Performance in a VR.

[Poster] *International Multisensory Research Forum (IMRF) 2014, Amsterdam,* Netherlands - presented poster titled: “But My Eyes are Fixed on You: Visually Evoked Postural Responses are Modulated by the Perceived Stability of Foreground Objects” and as a second author presented poster “Simulation Fidelity Affects Perceived Sound Comfort.”

[Talk] *Action Group AG-21GARTEUR 2013* - presented talk about my future studies at the conference meeting about Rotorcraft Simulation Fidelity Assessment: Predicted and Perceived Measures of Fidelity in the National Aerospace Laboratory (NLR), Amsterdam, Netherlands.

[Poster] *European Conference in Visual Perception ECVP 2013,* presented poster “The Effects of Stress on Body Ownership and The Rubber Hand Illusion” Currently published in Perception, 2013 http://www.perceptionweb.com/ecvp.cgi?year=2013

Abstract

The use of virtual reality in academic and industrial research has been rapidly expanding in recent years therefore evaluations of the quality and effectiveness of virtual environments are required. The assessment process is usually done through user evaluation that is being measured whilst the user engages with the system. The limitations of this method in terms of its variability and user bias of pre and post-experience have been recognised in the research literature. Therefore, there is a need to design more objective measures of system effectiveness that could complement subjective measures and provide a conceptual framework for the fidelity assessment in VR.

There are many technological and perceptual factors that can influence the overall experience in virtual environments. The focus of this thesis was to investigate how multisensory feedback, provided during VR exposure, can modulate a user's qualitative and quantitative experience in the virtual environment. In a series of experimental studies, the role of visual, audio, haptic and motion cues on objective and subjective evaluations of fidelity in VR was investigated. In all studies, objective measures of performance were collected and compared to the subjective measures of user perception.

The results showed that the explicit evaluation of environmental and perceptual factors available within VR environments modulated user experience. In particular, the results shown that a user's postural responses can be used as a basis for the objective measure of fidelity. Additionally, the role of augmented sensory cues was investigated during a manual assembly task. By recording and analysing the objective and subjective measures it was shown that augmented multisensory feedback modulated the user's acceptability of the virtual environment in a positive

manner and increased overall task performance. Furthermore, the presence of augmented cues mitigated the negative effects of inaccurate motion tracking and simulation sickness. In the follow up study, the beneficial effects of virtual training with augmented sensory cues were observed in the transfer of learning when the same task was performed in a real environment. Similarly, when the effects of 6 degrees of freedom motion cuing on user experience were investigated in a high fidelity flight simulator, the consistent findings between objective and subjective data were recorded. By measuring the pilot’s accuracy to follow the desired path during a slalom manoeuvre while perceived task demand was increased, it was shown that motion cuing is related to effective task performance and modulates the levels of workload, sickness and presence.

The overall findings revealed that multisensory feedback plays an important role in the overall perception and fidelity evaluations of VR systems and as such user experience needs to be included when investigating the effectiveness of sensory feedback signals. Throughout this thesis it was consistently shown that subjective measures of user perception in VR are directly comparable to the objective measures of performance and therefore both should be used in order to obtain a robust results when investigating the effectiveness of VR systems. This conceptual framework can provide an effective method to study human perception, which can in turn provide a deeper understanding of the environmental and cognitive factors that can influence the overall user experience, in terms of fidelity requirements, in virtual reality environments.

Acknowledgments

Firstly, I would like to thank my supervisors Dr Georg Meyer and Dr Mark White whose continuous support, motivation and scientific guidance helped me in every step in this process. Your advice and encouragement during this process was invaluable and without it none of this would be possible. I hope we can keep in touch and continue to discuss our research endeavours. I would also like to thank my examiners Professor Walterio Mayol-Cuevas from the University of Bristol and Dr Maria Limniou from the University of Liverpool for their advice and guidance on how to improve this dissertation.

My huge thanks go to the people at the Virtual Engineering Centre who supported me all the way and provided me with their technological creativity and expertise in regards to virtual environments. I will always remain grateful for the opportunity I've had to learn and progress as a research scientist in their company.

I have been very fortunate to have a support of my family and friends here in England and in Slovakia throughout this process and to all of them I give my thanks. In particularly I would like to thank, my mother-in-law Rachel and her partner Keith, my father-in-law Phil, my brother-in-law Ben and his wife Chloe and their children Alice, Scarlett, Max and Leo, my other brother-in-law Matthew and his wife Rachel and their children Sam and Louis (and now also baby Ehtan), who all have been there with their love and support and on many occasions helped to take care and entertain my children when I needed time and space to study.

I would also like to thank my friends and family in Slovakia, in particular my mum and dad, my sister Eva and her husband Slavo and their children Bianka and Eliska and my brother Lubos, his wife Lenka and their son Jakub, who all have always listened patiently to my moans and accomplishments, and my continuous efforts in trying to explain to them what I actually do. Although far away, they continuously supported me, through skype calls, phone calls and occasional visits, during my undergraduate and postgraduate studies and provided me with endless love and encouragement.

And last but not least, I would like to thank from the bottom of my heart my children Liana, Michael and Abigail who always gave me a lots of hugs and kisses that helped me unwind and made me smile every time I needed a break from my research work. Finaly I would like to thank from the bottom of my heart to my husband Zakeus, whose continuous support and encouragement gave me the strength to accomplish this. I am very grateful for your grammatical correctness, our many constructive discussions, but mostly for your constant love and caring.

I love you all and I am truly happy that you are all with me here to share this moment and celebrate my accomplishments, which I hope mark the start of my new career in research.

List of Contents

Chapter 1: Fidelity assessments in the virtual reality environments

Chapter 2: Environmental and perceptual factors in the fidelity assessments

Chapter 3: The effects of visual and environmental factors on postural stability

3.1 **Introduction**

3.2 **Methods**

3.3 **Results**

Chapter 4: The effects of multisensory cueing on fidelity assessments

Chapter 5: The effects of virtual training on transfer of skills from virtual to real environment

Chapter 6: The effects of motion cuing on fidelity assessments in simulated flight task

Chapter 7: Overall discussion and future recommendations

Chapter 1.

Fidelity assessments in virtual reality environments

1.1 General introduction

Virtual reality (VR) environments are commonly used as tools for training, research, interpersonal communication, data visualisation and many other purposes. Previous research studies have suggested that the user experience needs to be at the forefront of any evaluation process when investigating the usability of feedback signals in VR (de Korte, Huysmans, De Jong, Van de Ven & Ruijsendaa, 2014; Meyer, Clarke & Robotham, 2012, Meyer, Shao, White, Hopkins & Robotham, 2013). The feedback signals that are present during the VR interaction, determine the acceptability of the virtual simulation and any other devices or products that are associated with this technology (Pietschmann & Rusdorf, 2014; Meyer at al., 2012, 2013).

User evaluation of VR systems often includes all qualitative and quantitative experience that is measured whilst a user engages with a given system (Pietschmann & Rusdorf, 2014). The evaluation of the effectiveness of virtual environments is usually achieved through the assessments of fidelity. The definition of fidelity adopted in this research is the same as the definition proposed by Meyer et al. (2012) who suggested that the term fidelity can be referred to as *"a measure of the degree to which a simulation system represents a real-world system"* (Meyer et al., 2012, p.1). In general, fidelity can be thought of as the faithfulness of a simulation. The experimental research has suggested that well-developed models of fidelity can be used to predict outcomes and enhance statistical power in research applications

(Mowbray et al., 2003). In addition, the fidelity measures are recognised as important metrics of systems effectiveness as they can assist in producing a meaningful comparison of treatments and although the term fidelity is often applied to training devices, it has been also used in reference to equipment and the environment (Meyer et al., 2012; 2013; Keyson, 2007; Hamstra, Brydges, Hatala, Zendejas & Cook, 2014).

1.1.2 Fidelity of a virtual environment

The concept of simulation fidelity is usually understood as the degree to which a simulation looks, feels and acts like in a real life environment. It has been highlighted that fidelity is a multifactorial concept and that the fidelity requirements vary according to learning and task contexts (Hamstra et al., 2014). For example within the medical research, Screbo and Dawson (2007) suggested that fidelity could be described in physical and functional terms. Physical fidelity concerns the degree of similarity between the equipment, materials, displays, and controls used in the operational environment and those available in the simulation. On the other hand functional fidelity is concern with how the processes are implemented i.e. how the information requirements are accomplished with response requirements. For example, some VR-based medical simulations differ in how different instruments are selected as compared to the real environment and therefore they lack functional fidelity (Screbo & Dawson, 2007). Within aviation research, Perfect and colleagues (2010) defined two complementary subgroups of fidelity: predicted fidelity and perceptual fidelity. Predicted fidelity refers to the degree to which a flight simulator matches the characteristics of a real aircraft and is assessed by comparing quantitative data from aircraft and simulator. Perceptual fidelity refers to a simulator’s ability to induce the same pilot behaviour that is essential for the control and operation of a real aircraft, in a simulated environment (Padfield, Hodge & White, 2010).

It seems that there are many different descriptors used to define fidelity; examples reflecting the type of fidelity can be fitted into two groups: one that relates to the hardware, such as equipment fidelity, objective fidelity, physical fidelity, predicted fidelity; and one that relates to its functionality, such as behavioural fidelity, functional fidelity, environmental fidelity, informational fidelity and perceptual fidelity. All of these terms are used within the research literature and as such there is no formally recognised terminology for the term fidelity (Timson, 2013).

As the main theme of this research work is around fidelity assessments, it is necessary to define what terminology will be used. The main focus of this thesis is on the effects of multisensory cues, in particular how their presentation and salient availability can influence objective and subjective fidelity assessments. Therefore the term *physical fidelity* will be used for the faithfulness of the simulated cues and the term *informational fidelity* will refer to the nature and saliency of these cues.

1.1.2.1 High or low fidelity

For a long time it was generally assumed that higher fidelity results in higher performance; however the beneficial effects and the necessity for high fidelity simulations have been questioned in many empirical studies (Hamstra et al., 2014). Empirical research investigating the quantifiable benefits of high fidelity systems showed that rather than enhancing the overall VR set up, the enhancement of individual components can also be beneficial to performance (McMahan, Gorton, McConnell and Bowman, 2006; Dahlstrom, Dekker, Van Vinsen and Nyce, 2012; Screbo & Dawson, 2007; Salas, Wildman and Piccolo, 2009). This supports the notion that the overall fidelity of the virtual environment may not be as important in some tasks as it was previously thought. For example, McMahan et al. (2006) provided empirical evidence of the benefits of immersion, but found that object manipulation can be successfully performed with less immersive and less costly displays, with no loss of efficiency. Similarly, Dahlstrom et al. (2012) found that a high level of simulator fidelity has little or no effect on skill transfer. They suggested that lower fidelity simulation can reduce complexity and enhance focus on training and should be used to complement higher fidelity simulations. Screbo and Dawson (2007) suggested that there are three main reasons why high-fidelity VR systems do not always lead to better learning: individual based distortions in perception of stimulus; distorted reality in virtual displays due to different viewing angles; and the multimodal nature of perception. Salas et al. (2009) suggested that low fidelity simulation can often provide several advantages over more complex simulations, and provide just as rich of a learning opportunity. They concluded that as long as an appropriate level of cognitive (perceptual) fidelity is achieved, physical (predicted) fidelity does not necessarily need to be high when using VR as a management or educational tool.

Many studies have been conducted to investigate which factors are most important to the overall fidelity of the virtual environment and under which conditions the levels of fidelity can be degraded but still be effective for overall task performance (Hamstra et al, 2014, Salas et al., 2009, Screbo & Dawson, 2007, McMahan et al., 2006). These empirical studies have served as a main motivation for some of the research questions addressed in this thesis. The main aim of this thesis is to investigate the role of multisensory cues during VR exposure and its effects on user experience in terms of objective and subjective performance metrics. Additionally, I wanted to examine whether the nature and saliency of these multisensory cues will affect the overall fidelity assessments of VR. The experimental study described in Chapter 4 is focusing on the effects of the high level of informational fidelity on user experience when the overall physical fidelity of the simulation is decreased.

1.1.3 User experience

User experience (UX) is often defined as an umbrella term for all qualitative experiences a user has while interacting with a given product (Pietschmann and Rusdorf, 2014)**.** The ISO definition of UX focuses on a “user's perception and responses resulting from the use or anticipated use of a product, system, service or game” (ISO FDIS 9241-210:2010 as cited in Pietschmann and Rusdorf, 2014). Within the virtual environment, the user acts and interacts within a space generated by the computer. By being enveloped in this environment the user experiences presence in the computer-generated world which surrounds the user with ever-changing sensations, while simultaneously responding to the user’s actions (Witmer and Singer, 1998).As oppose to usability , which refers to ability of the user to carry out certain task, user experience takes a broad view looking at the individual interactions with the system as well as the thoughts, feelings and perceptions that result from that interaction (Albert & Tullis, 2013).

The key elements of the virtual reality experience are the virtual world, immersion, sensory feedback and interactivity (Faas et al, 2014). Keyson (2007) suggested that the key factors for enhancing user experience in VR are the sense of being in control, the feeling of involvement, emotional engagement and expected and functional performance of VR. In the similar way, Witmer and Singer (1998) suggested that selective attention, involvement and immersive response are necessary to experience the sense of presence. The terms immersion, presence and involvement have become the key concepts that are being studied during the user evaluation process of the effectiveness of VR systems (Faas et al, 2014; Witmer and Singer, 1998, Ma and Kaber, 2006, Meehan et al., 2006).

1.1.4 Objective and subjective measures of fidelity

Task performance is the most common objective measure of system effectiveness. The measures that are obtained to evaluate task performance are task completion time, error rate, task accuracy or response time. The main advantages of objective measures is that they offer less variability across participants and they do not require extra attention from users, however they are less sensitive to users' preferences and habituation (Screbo & Dawson, 2007).

The subjective measures are usually collected through the administration of various questionnaire-based rating scales. The reasons for the frequent use of subjective measures include their practical advantages, such as the ease of implementation and their non-intrusiveness. However, it has been pointed out that they are time consuming and as they are filled in after the experience they might not capture the constructs that they attempt to measure (Rubio et al., 2004).

The above-mentioned research has provided another theme to the research work described in this thesis, in particular in the way how to use, evaluate and compare objective and subjective metrics used for the assessments of fidelity. A significant amount of work has been conducted in order to bridge the gap between the subjective opinion and the formal objective metrics of fidelity. The effectiveness of virtual reality environments is traditionally assessed by subjective measures of presence and immersion. However, many studies have pointed out that there is a need for more objective measures of fidelity. For example, Wong (2014) tested participants in a high fidelity helicopter flight simulator in a target-tracking task whilst manipulating the available audio cues. She compared objective measures after training with subjective self-evaluation measures. Whilst some correlations between objective and subjective measures were found, she reported that sometimes there was no correlation between the measures. She concluded that although subjective measures are a good indicator of self-performance, objective data offer a valuable task-orientated perspective on simulator fidelity. Similarly, Meyer et al. (2013) argued that objective fidelity evaluations of virtual environments should be centred on human-performance and be task-specific, rather than measure the match between simulation and physical reality. In their study they showed how multisensory perception measures could be used to form an evaluation framework that uses human performance as a referent and is designed to evaluate the contribution of individual cues to task performance. The same assumption is undertaken in the studies described in this thesis as the presented studies where designed in order to investigate the relative contribution of individual sensory cues: visual, audio, tactile and motion.

The above mentioned studies provided an overall theme for the research work presented in this thesis as they suggest that there is a need to further investigate and evaluate factors that influence the fidelity assessments of virtual reality environments. Therefore, the overall aim of this thesis is to provide additional knowledge into the understanding of the role of multisensory cuing in user experience and the fidelity assessments. In the section bellow I will describe the main aims of my research work and present research questions that were addressed in the studies that form the chapters of this thesis.

1.2 Main aims of this thesis

The overall focus of this thesis is centered on fidelity - in particular how the fidelity is defined and what is the best ways to assess the fidelity in VR environments. The work presented in this thesis is mainly concerned with the role of sensory feedback that is available during the VR exposure and its influence on the overall experience of a user during the fidelity assessments. The main implication of this research is to provide a deeper understanding of the fundamental aspects of perception, performance and interaction within the real and virtual reality environments and through the knowledge gained to provide an evaluative framework for the assessment of the fidelity in the virtual environments.

This thesis will present a series of experimental studies that were designed to investigate how humans perceive and interact within virtual reality (VR) environments. The main aim of this thesis is to explore the role of multisensory feedback presented during VR exposure. The main questions driving this research are centred on the sensory cues available in virtual environments during the VR interaction. In particular, I am going to explore and investigate the role of visual, audio, tactile and motion sensory feedback during the experimental task on objective and subjective fidelity assessments. In order to do this, I designed and conducted a set of experimental studies where the main focus was to investigate the role of multisensory perception for the effective assessment of the fidelity of virtual environments. By presenting multisensory cues in various modes of feedback, such as unimodal, bimodal and multimodal, and designing various performance tasks, I hoped to provide a greater understanding of the advances and limitations of multisensory presentation in VR on user's perception, behaviour and acceptability of virtual environments.

In the section below, I will present the main research questions in three sections. Each section is concerned with the assessments of fidelity in VR, especially focusing on the ways the sensory cues are presented in virtual environments. For each section a set of experimental studies were designed and conducted. The studies have been separated to form the chapters of this thesis.

1.2.1 Fidelity of the environmental factors

The first question that I wanted to address is concerned with how important the presentation of sensory cues is in VR. In particular, does it matter to what degree the visual sensory cues presented in VR represented the perception in the real world? Does the perception of depth influence the fidelity of VR? To answer these questions, two experiments were conducted where the fidelity of the virtual environment was manipulated and the effects it had on user's behaviour were recorded. In the first study I examined whether visual disparity cues, used to provide depth perception through 3D stereoscopic displays, affect user perception. In the second study I investigated whether the environmental factors, such as the physical properties of environmental reference anchor points presented in a scene during VR exposure will modulate the fidelity of the simulation. In both studies the magnitude of postural responses was recorded as an objective measure of fidelity. As the postural responses are usually performed unconsciously (Meyer et al., 2013), the main aim was to investigate whether the changes in the presentation of the visual sensory cues affects users postural stability. These studies are described in greater detail in Chapter 3.

1.2.2 Fidelity of the sensory feedback in virtual environment

In my following studies I have focused on interaction techniques used in VR. My research questions were focused on the informational fidelity of presented sensory cues. I was interested to explore what happens when the available sensory feedback in a virtual environment is not portraying the faithful representation of the real world? What if the sensory information presented during the VR task lacks physical fidelity but provides useful information? Will the varying levels of fidelity influence user's objective performance? Will it modulate user's subjective evaluations of the VR environment? To answer these questions, I designed and conducted two experimental studies where sensory cues presented during the task were available as augmented cues, i.e. they lacked physical fidelity but provided additional task-relevant information and therefore had high informational fidelity. Additionally, the virtual environment was modulated by the introduction of motion tracking discrepancy during the VR simulation. The reason for this manipulation was to investigate whether the augmented cues will influence user's behaviour and performance. The follow-up study, with the upgraded VR set up, investigated whether the virtual training using these augmented sensory cues can translate into performance improvements in the real environment. The first study is described in more detail in Chapter 4. The second study, which focuses on the effects of virtual training with additional sensory cues and the subsequent transfer of learning from virtual to the real task, is described in Chapter 5.

1.2.3 Fidelity assessment during the whole body motion simulation

In the next study the focus was again on the interaction in VR but this time the main emphasis was on the motion cues, in particular the effects of whole body motion, such as the motion platform used in aviation training. The main questions driving this research were: Do whole body motion simulators guarantee improvements in fidelity requirements? Can we observe beneficial effects of the motion cuing when the perceived demand of the task is increased? Is there a relationship between the motion cuing and user’s perceptual workload? To answer this question, a 6 degrees-of-freedom (DOF) helicopter flight simulator was used in order to provide sufficient whole body motion feedback. I designed and conducted a study where the effects of motion cuing were examined whilst pilots performed a simulated flight task. The slalom manoeuvre was flown in two motion conditions whilst at the same time the perceived demand of the task was manipulated by the pole width separation in three experimental conditions. The study, which is described in Chapter 6, focused on assessing pilot’s flying strategies and task accuracy during the task as well as on the pilot’s psychological state including presence and immersion, simulation sickness and cognitive workload.

1.3 Experimental approach

In all of the studies presented in this thesis, a user-centric design was adopted where behavioural and psychological measures were collected, analysed and compared. In all cases the objective measures of performance were collected and at the same time the ratings from various subjective questionnaire-based measures were also obtained.

1.3.1 Ethical approval

Prior to running the experimental studies described in this thesis, ethical approvals were obtained from the University of Liverpool's Institute of Psychology and Health Sciences Ethics Committee (PSYC-1112–049A). The ethical guidelines for qualitative and quantitative approaches were as followed:

- Every participant who took part in the studies described in this thesis was given prior information about the study and after making an informed decision about participation they signed an informed consent form.
- Every participant was informed that they can withdraw at any time during the task, should the task become too uncomfortable or too unpleasant for the participants to continue.
- Participants were informed about the confidentiality guidelines and requirements of the information collected; all collected data were securely stored in line with the guidelines.
- After the task every participant was debriefed and made aware of the main aim and the purpose of the study.

1.3.2 Statistical analyses

For the statistical analysis of the collected data and measures, a factorial within subject repeated measures design (ANOVA) was adopted in all of the studies presented in this thesis, except for one study (Chapter 5) where different groups of people were subjected to different experimental conditions and therefore a factorial mixed design (ANOVA) was adopted. Additionally, pairwise comparison tests (dependent and independent) were conducted when the difference between the conditions was investigated. The reliability and validity analyses of the questionnaire-based metrics used in these studies were also conducted.

Prior to every experimental study described in this thesis, a power analysis G-power (Faul, Erdfelder, Buchner & Lang, 2009) was conducted to determine the sample size required to detect an effect of a given size with a given degree of confidence. Statistical power is a measure of likelihood that a researcher will find statistical significance in a sample if the effect exists in the full population (Field, 2013). The power of a statistical test is the probability that its null hypothesis (H_0) will be rejected given that it is in fact false. G power is a power analysis program for statistical tests commonly used in social and behavioural research. It provides improved effect size calculators and graphic options, supports both distribution-based and design-based input modes, and offers all types of power analyses. Statistical tests were specified through a distribution-based approach and through a design-based approach. The type of power analysis conducted for each experimental study described in this thesis was a priory test which was done to compute required sample size – given α, power and effect size as input parameters. The output parameters consisted of critical *t*, degrees of freedom (Df), total sample size and actual power. For the detailed description about G power analysis please refer to Faul et al. (2009).

1.4 Overall findings and contributions

The work described in this thesis is concerned about the role the multisensory cues play in the fidelity assessments of the virtual environment (VE). In this thesis objective and subjective measures were used to determine the effects of multisensory feedback on overall user experience assessed through the fidelity metrics. Visual, audio, tactile and motion sensory feedback was provided in a series of experimental studies whilst participants performed a task in highly immersive virtual environments. The main findings of the conducted research are described bellow

- The manipulation of visual disparity cues that enables depth perception in VR environments can influence the fidelity of the environment as assessed through the postural adjustments. Environmental factors, such as physical properties of the real reference points within the scene were also shown to influence postural responses. Therefore it can be concluded that the way the sensory cues are presented in VR can impact the fidelity of the VR environment. Additionally, the results confirmed the suitability of the postural responses to multisensory stimulation as a viable option for the objective assessments of fidelity in virtual environments.

- Increasing the informational fidelity of the multisensory cues during the manual assembly task in VR enhances performance and users’ overall experience, even when the overall physical fidelity of the virtual environment is disrupted. The results showed that the additional augmented audio, tactile and visual sensory feedback that is relevant to the task in hand, improves task performance and participants’ sense of presence and immersion. In addition, the beneficial effects of audio and tactile feedback were recorded in objective performance. The additional multisensory feedback was also shown to mitigate the negative effects of motion tracking inaccuracy.
- The virtual training with augmented sensory cues translates into performance improvements in the real task scenarios. Virtual training groups showed better performance on the real task than the control group. Augmented sensory cuing in VR stimulated attention to error prone events on the real task – the group that received virtual training with additional augmented cues conducted significantly less errors that the group that was trained in VR without any additional sensory cues provided.
- The availability of motion feedback during the simulated flight task has beneficial effects on the objective and subjective fidelity metrics. The task

performance and users’ comfort was improved when motion cues were available. The results showed that the perceived task demand during flight task influenced the beneficial effects of motion feedback in low workload conditions. Therefore it can be concluded that a certain amount of cognitive workload is necessary to attain alertness in order to achieve the desired accuracy of performance.

1.5 Structure of the thesis

The structure of this thesis is as follows:

Chapter 2 will present a background literature review that was conducted to provide motivation and insight for the presented research work by discussing self-motion, multimodal perception, motion tracking techniques in VR, and the use of motion platforms. The psychological measures that are used to investigate user experience during VR exposure and interaction will be also introduced together with the usability metrics used in the experimental studies described in this thesis.

Chapter 3 will then presents the first study that investigated the effects of visual disparity cues that enable depth perception through 3D stereoscopic displays and the effects of environmental factors on users’ acceptability of VR environments as measured through postural balance.

Following this, Chapter 4 will outline the second study that consists of two experiments that were set up to explore the effects of augmented sensory cues on users’ perception and performance during the manual assembly task in VR. In this chapter the effects of the positional inaccuracy of the motion tracking systems is also investigated.

Chapter 5 then describes a follow up study from Chapter 4 that was designed to investigate whether the beneficial effects of virtual training with additional augmented sensory cues can be translated into performance improvements in the real task scenarios.

Chapter 6 will present the last study of this thesis that was conducted in a high fidelity 6 degrees of freedom helicopter flight simulator. The main aim of this study was to investigate the role of motion sensory feedback on objective and subjective fidelity assessments, such as task performance and cognitive workload.

Finally, Chapter 7 summarises the overall findings, presents main conclusions of this work and provides recommendations for future research.

Chapter 2.

Environmental and perceptual factors in the fidelity assessments

Chapter overview

The fidelity assessment of the virtual environment can be influenced by many environmental and perceptual factors. This chapter will present a background review of empirical research studies that investigated environmental and perceptual factors in the assessment of the fidelity in VR. These included the investigation of postural stability, multisensory perception, multimodal interfaces, motion tracking systems and motion cuing available in virtual environments. Most of the research studies mentioned in this chapter were concerned with the fidelity evaluations in terms of user performance and perception, whilst using virtual reality as a research platform.

2.1 Self-motion perception in virtual environments

The human balance system relies upon sensory information from the surrounding environment to interpret how the body is moving and is continuously adjusting for the state of instability (Day & Gueraz, 2007). To navigate effectively through our everyday environment we must estimate accurately our own motion relative to people and objects around us. Sometimes the signals that help control our postural sway can become ambiguous, especially during self-motion where the moving environment or visual motion can create self-motion illusion, also called vection (Riecke, Schulte-Pelkum, Avraamides, Heyde & Bülthoff, 2006; Meyer et al., 2013). When this happens, our nervous system must combine and resolve this ambiguity in order to control and stabilise our body posture (Dokka, Kenyon, Keshner & Kording, 2010). Perhaps one of the most known examples of self-motion illusion is experienced on a train station when a person who is sitting on a stationary train perceives the motion of an adjacent train as the motion of the their stationary train: visual information indicates motion however no vestibular signals are present since the train is still stationary therefore causing a self-motion illusion (Riecke et al. 2006).

When the virtual environments are configured to a high level of fidelity to simulate the real world, the same postural responses as experienced in the real world should be observed in high fidelity virtual settings (Riecke et al., 2006; Meyer et al., 2013). The self-motion illusion in is a very convincing and compelling experience; it is sometimes desired but sometime not and in this cases it can have negative consequences in terms of participant’s wellbeing. Self-motion illusions are particularly desired in the entertainment industry as it can improve the overall user experience. The main goal of many VR applications used in the entertainment industry, such as 3D/IMAX theatres and theme parks, is to place a person into a mediated environment and create an experience combing all multisensory information, including motion, to provide realistic sensations (Atkins, 2008). However, many VR technologies are used in commercial and industrial research settings where the VR is mostly used for visualisation, teaching and training purposes (Salas et al., 2009; Van der Linden, Johnson, Bird, Rogers & Schoonderwaldt, 2011; Laha, Sensharma, Schiffbauer & Bowman, 2012; Stevens & Kinciad, 2015; Menzies, Rogers, Phillips, Chiarovano … & MacDougal, 2016). The self-motion illusion in these applications is not desired however it is often experienced. This can have adverse effects of performance and user experience, mainly because it can give rise to the symptoms of motion sickness (Seay, Krum, Hodges & Rybarsky, 2001; Sparto, Whitney, Hodges, Furman & Redfen, 2004; Stein & Robinski, 2012). Empirical studies have shown that postural instability precedes motion sickness, which is the main concern of VR technologies (Murata, 2004; Villard et al., 2008) and it is therefore important to investigate factors that can affect postural stability. Ricco and Stoffregen (1991) propose a theory of postural stability in which that defined postural control as the coordinated stabilisation of all body segments. This theory holds that when human or animal encounters a destabilising environment it must try to regain and maintain postural control. If the strategy for marinating postural control is not learned a postural instability occurs which can give rise to motion sickness (Ricco & Stoffregen, 1991). These studies have all provided motivation for some of the themes and research questions addressed in this thesis.

It has been suggested that postural stability can be used as an objective measure of fidelity in virtual reality environments that could complement subjective measures. Meyer et al. (2013) set out to investigate what factors can modulate the amount of postural sway in response to visual motion. They explored the effects of virtual and real environmental anchors on visually evoked postural responses (VEPRs) that are part of natural postural sway, in high fidelity virtual environment. They hypothesize that as these responses are performed in unconscious manner they should mimic those seen in real situations. Their results revealed that lateral postural sway can be modulated by the presence of visual, audio and tactile reference points, however the anterior-posterior postural sway was unaffected by the presence of these anchors. They concluded that postural sway can be used as an objective measure of fidelity, which could complement user’s subjective evaluations of VR (Meyer et al., 2013). Similarly, Menzies and colleagues (2016) suggested that postural stability could be used to objectively assess visual fidelity of VR headset. In their study three

VR devices and a stable VR visual stimulus were evaluated in regards to postural stability. They findings showed that the eyes-open condition allowed for significantly greater postural stability than the other conditions, which supports the validity of posturography as a measure of fidelity in VR (Menzies et al., 2016).

The above-mentioned research has directly influenced my research work presented in Chapter 3 that focused on how presentation of visual information in VR environments can influence postural stability and possible motion sickness. The discussion of previous research work allowed me to identify two key concerns; the need to further investigate the role of the fidelity of visual information, and the need to identify factors that can reduce the negative effects of self-motion illusion. Informed by the findings from previous research, I designed and conducted two experimental studies: the first study focused on the effects of the fidelity of visual cues in VR, in particular the role of visual disparity cues that enable depth perception; in the second follow up study the main focus was to investigate how the physical properties of environmental anchors can modulate the negative effects of self-motion illusion. The main motivation behind this research was to extend the previous research and provide additional knowledge about the factors that can influence postural stability during the VR exposure.

2.2 Multisensory perception in virtual reality environments

Multisensory feedback presented in VR is another way how the fidelity of the VR technology can be assessed. As mentioned previously our environment delivers a large amount of sensory stimulation from all modalities and as the human limitations reduce our ability to pay attention to multiple stimuli, the brain needs to choose which information is necessary in terms of its characteristics and relevance for any given moment. To provide a clearer understanding of this process, Wickens (1998, 2002) developed Multiple Resource Theory (MRT) that is considered the guiding principle of multimodal perception. MRT theory assumes that people have different cognitive resources for processing information. Some of these resources are more suited for simultaneous use (parallel processing) than others. When the information is perceived through different modalities, fewer interfaces should occur than compared to the case where the same amount of information is presented through a single modality (Wickens, 2002). Sometimes, many resources are available for one modality whilst others are limited in processing information, which can eventually deteriorate performance, especially when one modality needs to handle more than one task (Oskarsson, Eriksson & Carlander, 2012).

2.2.1 Visual and audio feedback

As a human, we have a bias towards vision as we take most of our information about the environment around us through visual modality (Jerome, 2006). In regards to the VR environment, it is usually believed that the more sensory feedback is present, the more effective the VR application is (Jia, Bhatti & Nahavandi, 2011, Lee & Bilinghurst, 2013). However, visual modality can easily become overwhelmed, with a huge amount of information in the real and virtual environments, which may decrease overall task performance. In a bid to overcome this limitation, experimental research in multisensory stimulation suggested potential benefits of other modalities, such as audio feedback. Indeed, studies have shown that visual and audio sensory feedback presented alone or in combination, improved task performance in target localisation (Jerome, 2006), target accuracy (Jacko, Emery, Edwards, Ashok, … Sainfort, 2004) and spatial attention tasks, without affecting perceptual workload

(Jeon, Davison, Nees, Wilson & Walker, 2009; Santangelo & Spence, 2007). Meyer et al. (2012) found that visual and auditory cues (with training) can lead to improved and transferable performance.

2.2.2 Haptic/tactile feedback

The advantages of additional haptic feedback in VR environments have also been explored, even though this feedback is slightly more difficult to operationalize in VR. In 3D virtual interactions haptic feedback can be presented as force feedback or vibrotactile feedback. Force feedback conveys information to the user by the generation of forces on the mechanical interface, which a user can move (for example Phantom limb) whereas vibrotactile feedback refers to the skin stimulation that uses vibrations to transmit sensations (Martinez, Martinez, Molina & Garcia, 2011) but not, for example, mechanical resistance to movements.

The use of force feedback in the simulation is often very costly and it has been argued that vibrotactile feedback is a viable alternative to force feedback with minimal cost and complexity (Martinez et al., 2011; Schoomaker & Cao, 2006). For example, Martinez et al. (2011) compared force and vibrotactile feedback in texture detection study and showed that vibrotactile feedback is more efficient than force feedback in the textures recognition from different patterns and shapes. Similarly, Schoonmaker and Cao (2006) suggested that the distortions in force feedback devices used during minimal invasive surgery make the task more difficult for surgeons. They investigated whether vibrotactile feedback can enhance performance and control the forces that are being applied. Their results showed that vibrotactile feedback aided probing depth error and the control of the forces applied. They concluded that vibrotactile stimulation can serve as a viable substitute for force feedback in simulated minimal invasive surgery. The findings from these studies provided motivation for inclusion of haptic feedback during my studies described in Chapter 4 and 5.

The empirical research investigating the benefits of vibrotactile feedback has shown potential advantages during mulimodal stimulation. For example, it was suggested that the main contribution of vibrotactile cues to effective interaction is seen when the visual or auditory modalities are fully engaged (Vitensen, Jacko & Emery, 2003; Akamatsu, MacKenzie & Hasbroucq, 1995; Hopp, Smith, Clegg & Heggestad, 2005) especially when the task involves close range interaction or simple manipulations (Van Erp & Van Veen, 2004, Ramsamy, Haffegee, Jamieson & Alexandrov, 2006; Burke, Prewett, Gray, Yang, Stilson… & Redden, 2006; Hopp, Smith, Clegg & Heggestad, 2005; Adams, Klowden & Hannaford, 2001; Feintuch, Raz, Hwang, Josman… & Weiss, 2006). The presence of vibrotactile feedback was also shown to enhance performance in spatial guidance tasks (Ho, Tan & Spence, 2007), in the teaching and learning processes of a new physical activity (Oakley, 2009; Faas, Bao, Frey & Yang, 2014) and this feedback also served as an efficient warning signal in complex tasks, such as driving a car or air traffic control (Ho et al., 2005). In the meta-analysis of 43 studies that that examined the effect of multimodal feedback on user performance, Burke et al. (2006) found that visual-auditory (VA) and visual-tactile (VT) signals reduced reaction time and improved performance; in particularly VA feedback was shown to be most effective in a single task with normal workload, and VT feedback was shown to be more effective when multiple tasks are being performed and the workload conditions are high. Both of them were effective for target acquisition tasks but they varied in the effectiveness for other task types.

They concluded that adding an additional modality to visual feedback improves overall performance (Burke et al., 2006).

This literature review provided a great motivation and the main theme for my research work presented in this thesis, which is described in Chapter 4 and 5. The above-mentioned studies have investigated the potential beneficial effects of multisensory presentation during the task, in particular visual, audio and tactile sensory cuing. This research directly influenced the main focus of this thesis being on the role of multisensory cuing in the fidelity assessments of VR. In the studies presented in Chapter 4 and 5 additional sensory cues were provided during a manual assembly task to see whether the research findings from the previous studies that suggested potential advantages of visual, audio and tactile sensory cuing can be replicated. However, my research work extends previous studies by investigating how additional task-relevant information can be utilised by the user. By conducting this research I hope to provide additional knowledge about the beneficial effects of additional sensory cuing on user performance and overall subjective experience in VR.

2.2.3 Motion feedback

In the same way as visual, audio and tactile sensory cues, motion cuing is also investigated in a great detail by the experimental researchers especially within the transport industry, for example with aviation and automotive technologies. The beneficial effects of motion cuing have been observed in many studies, mainly because the presence of motion feedback improves the realism of the simulated flight task and thus have been linked to user’s sense of presence. For example, Mulder, Verlinden and Dukalski (2012) investigated the effects of motion on the sense of presence during virtual sailing. Five experienced sailors completed the same course several times while subjected to various motion conditions. Their results showed that adding simulated motion resulted in a positive effect on the subject's presence and immersion. Steven and Kincaid (2015) examined whether the higher fidelity of visual information contributes to the perceived sense of presence and performance. Their findings showed that visual information presented during the simulated task had a significant effect on performance and presence ratings. Moreover, a significant moderate relationship between performance and presence was found, which suggests that increasing user immersion and sense of presence could result in an improved performance.

These research studies directly influenced the research work presented in Chapter 6 of this thesis; here the effects of motion cuing are investigated during various conditions that differentiate in the levels of perceived task demand on cognitive workload

2.3 Multimodal interfaces and performance

The sensory information in VR is usually presented multimodally and as it was suggested that multimodal perception can be sometimes limiting (Wickens, 1998, 2002) a significant amount of empirical research have been conducted to investigate its usefulness in VR. The advantages of multisensory interfaces have been explored and reported in empirical studies. For example, Jai et al. (2011) investigated the role of multimodal interfaces during a virtual assembly task and found that multimodal feedback significantly improved performance. Similarly, Lee and Bilinghurst (2013) showed that a multimodal interface used within augmented reality increased the levels of user satisfaction and improve performance.

Augmented reality (AR) is one of the recent technological developments in human-computer interactions and has been implemented in a wide variety of application domains, including medicine, entertainment, education and engineering (Navab, Feuerstein & Bichlmeier, 2007; Billinghurst, Clark & Lee, 2015; Mayol-Cuevas, Davison, Tordoff & Murray, 2005, Billinghurst, 2002; Torrez-Gomez & Mayol-Cuevas, 2014, Henrysson, Billinghurst & Ollila, 2005; Fjeld & Voegtli, 2002; Piekarski & Thomas, 2002). The main aim of AR application is to provide additional information within the real world through the use of AR interfaces. Sensory information presented with VR and AR technologies have the potential to improve users experience including performance, therefore there is a need to investigate the effects of additional sensory cues on task performance and user perception in VR (Van Krevelen & Poelman, 2007). This research directly motivated my choice of sensory cues for the research studies described in Chapter 4 and 5. In these studies the sensory cuing was augmented in a way that provided additional task-relevant information and its utility was evaluated through the objective and subjective performance metrics.

When examining the beneficial effects of multimodal interfaces on user's cognitive workload, Shi, Ruiz, Taib, Choi and Chen (2007) and Oviatt, Coulston and Lunsford (2004) all found that users reported a lower cognitive load when operating a multimodal interface compared to operating a unimodal interface, especially when performing complex tasks. Despite this, other studies have reported that the availability of sensory modalities, in particular tactile and audio feedback, can decrease overall performance; they were shown to have an adverse effect in accuracy tasks and were also perceived as distracting and annoying (Oakley, 2009; Vitense, jacko & Emery, 2003). Moreover, it has been reported that in some cases the interaction with the multimodal interfaces may cause a sensory overload, which can decrease performance and the overall effectiveness of the virtual environment (Viaud-Delmon, Warusfel, Seguelas, Rio & Jouvent, 2006). The results from above mentioned studies suggest that information presented in other modalities (audio and haptic) can take pressure of the visual channel and subsequently reduce cognitive processing efforts that are required for effective task performance (Shi et al, 2007; Oviatt et al., 2004). Thus, it can be concluded that the information from multiple modalities improves performance more than information competing for attention in the one modality (Jai et al., 2011). Similarly, the sensory cues presented in an AR application were also shown to improve performance (Lee & Billinghurst, 2013).

The discussion of previous research findings suggests that additional modalities can be utilised during the task and as such they provide another theme for this thesis by looking into how the sensory modalities work together during the interaction in VR. As the augmented cuing is usually used to present additional information over the real world scenarios, this research extend previous studies by using augmented cuing in virtual environments.

2.4 Motion tracking in virtual environments

In virtual environments, a high quality motion tracking is necessary for effective interaction as it records the position and the orientation of real objects in physical space and allows spatial consistency between real and virtual objects (Lugrin, Weisbush, Latoschik & Strehler, 2013). Effective motion tracking during the VR interaction is also important for fidelity assessments (Steed, 2008). The best way to evaluate the effectiveness of a motion tracking system is usually through the measures of perceived latency or lags of the system (Papadakis, Mania & Koutroulis, 2011; Meehan, Razzaque, Whitton & Brooks, 2003; Steed, 2008). Inaccuracy of a motion tracking system is an important parameter in understanding the effectiveness of VR systems; therefore a continuous measurement in controlled conditions is necessary and has been encouraged by many researchers (Meehan et al., 2003). The VR research community is yet to agree on a common definition of latency; it is sometime referred to as a lag in the frame rate, refresh rate or update rate. Meehan et al. (2003) defined the term latency as '*the delay between a user's action and motion and when that action is visible on the display*'*(p.345)*. Empirical research, such as the work of Ellis et al. (2002) shows that latency and update rate can have a negative impact on performance and user responses in VR and can also lead to simulation sickness (Steed, 2008). Friston, Karlström and Steed (2016) found that latency begins to affect performance at 16ms into the task and this effect was found to be non-linear.

2.4.1 Motion tracking systems

To achieve a high tracking accuracy, a three-dimensional high-quality motion tracking system needs to be implemented as it was argued that an accurate motion tracking is necessary to create and update the viewpoint of the user and allow natural interaction in VR (Greuter & Roberts, 2014).

VICON motion capture system (Vicon, 2016) is recognised to be one of the best available systems for accurately recording three-dimensional movement, particularly the movements of a human body (Dobrian & Bevilacqua, 2003). The system have been extensively used for motion tracking in academic and industry research, medical applications as well as in entertainment and sport industries (Vicon, 2016). Due to high speed and high resolution of the VICON Bonita infrared cameras the VICON system has been shown to provide an accurate method for capturing human and object motion at high frame rate (Dobrian & Bevilacqua, 2003). The system uses marker-based motion capture: spheres covered with reflective tape, known as markers are placed on visual reference point on different parts of human body of objects. The VICON system is designed to track and reconstruct these markers in three-dimensional space. During the capture the coordinates of all markers ion each camera's view are stored in a data-station. The VICON system then links the correct positions of each marker together to form continuous trajectories (Kapur, Tzanetakis, Virji-Babul, Wang & Cook, 2005).

However, as an accurate and precise motion tracking generally relies on expensive hardware, recent availability of inexpensive commercial systems for tracking of body movement enabled many researchers to effectively record positional movements of subjects during the experimental studies. One of these systems is Microsoft Kinect, originally developed for the interaction in a computer game environment, but due to its low-cost range sensors it has become a widely used alternative to expensive motion tracking softwares like VICON system in many applications (Khoshelham & Elbernik, 2012, Shin et al., 2013). The Kinect sensor consists of infrared laser emitter, an infrared camera and RGB camera. The laser beam from the camera is projected onto a scene which is captured and correlated back to the reference pattern. The Kinect sensor captures depth and color images simultaneously at a frame rate of up to 30 fps (frames per second), provides angular field of view (57°) and has estimated covering area of about 1 - 3.5 meter.

Due to its ability to store and retrieve a vast amount of data, Microsoft Kinect has been used in many research applications (Khoshelham & Elbernik, 2012). For example, Microsoft Kinect was used as a reasonably accurate 3D imaging device in medical settings (Shin et al., 2013), in mapping applications (Khoshelham &

Elberink, 2012), in 3D modeling (Oikonomidis, Kyriazis & Argyros, 2011) and 3D tracing (Lee, Bonebrake, Bowman & Höllerer, 2010). However, when used for human motion tracking, Kinect sensors have shown few drawbacks especially in the error measurements, latency, and limited tracking area and resolution quality. Khoshelham and Elberink (2012) found that the error of depth measurement increases with increasing distance to the sensor and ranges from a few millimetres up to about 4cm at the maximum range of the sensor. Furthermore, they found that the overall quality of the data was dramatically decreased by the low resolution and depth measurements. Similarly, when the spatial accuracy and precision of Microsoft Kinect in creating 3D images used for medical applications was investigated and compared to actual ruler measurements on flat and curved objects, the results showed a 2mm error on flat and a 3 mm error on curved objects (Shin et al., 2013). Therefore, a continuous evaluation of low cost motion tracking systems is needed in order to determine its effects and identify any possible shortcomings arising from the use of these systems.

As the studies presented in this thesis were conducted with a high fidelity motion tracking systems it provided another theme for the work conducted in this thesis. By manipulating the accuracy of the motion tracking system which was done in order to simulate error measurements reported with low-cost motion tracking systems, comparison analyses were possible. This provided another objective to the studies investigating the effects of augmented sensory cuing.

2.5 Motion feedback in virtual environments

Motion systems have been used to support training for several decades; the technological advances of visual systems based on model boards were later transformed into a simple version of computer image and following the enormous advancements of image systems, complex aircraft models and motion control systems were later developed (McCauley, 2006).

Motion platforms have been used in real time vehicle simulations to simulate motion, especially in driving and flight simulators. The goal of these platforms is to provide user with the most accurate perception of driving or piloting the vehicle (Casas, Coma, Riera & Fernandez, 2012). A simulation motion base needs to generate a certain amount of acceleration cues to support realistic motion perception. The acceleration cues from the motion platform and the visual cues provided by outside world scenery are the main cues that will make the pilot aware of disturbances in pitch, roll and turbulence (Groen, Valenti Clari & Hosman, 2001).

In general, motion is proposed to be more beneficial in rotary-wing flight training than fixed-wing flight training (Pasma, Grant, Gamble, Kruk & Herdman 2011). But despite many technological advances and the use of state-of-the-art equipment, it has been recognised that motion platforms still have many constrains, such as limited motion cuing in terms of duration, amplitude and realism (McCauley, 2006; Perfect, White, Padfield, Gubbles & Berryman, 2010; White, Pefect, Padfield, Gubbles & Berryman, 2012). The effects of motion feedback available through motion platform in VR environments provided basis for another theme of research questions addressed within this thesis.

2.5.1 Motion cuing in aviation

The necessity of motion platforms within the aerospace industry has been investigated in great depth (Hall, 1989; Schroeder, 1999; Wang, White, Owen, Hodge & Barakos, 2013; Hodge, Perfect, Padfield & White, 2015); however the empirical research investigating the effects of motion cuing has produced some contrasting findings. For example, Wang et al. (2013) investigated the effects of motion cuing in various environmental conditions and found that in the degraded visual conditions, such as fog or night-time, motion feedback facilitated performance. Motion cuing has also been found effective in achieving desired positional and target accuracy; pilots performed manoeuvres more successfully without the need for heavy over-controlling i.e. too many inputs through the control mechanisms such as the cyclic stick, the collective lever and the anti-torque pedals) (Hodge et al., 2015). Schroeder (1999) explored the effects of different levels of degree of freedoms in motion platforms: roll rotation, yaw rotation, lateral translation, and vertical translation and found that lateral and vertical translational motion had a significant effect on simulation fidelity. Apart from the motion cues, other sensory information available during the simulation tasks such as optic flow, horizon and whole body rotation where also shown to facilitate performance when subjects were asked to stabilise helicopter motion simulators during a hover manoeuvre (Berger, Terzibar, Bykirch & Bülthoff, 2007). In the study that investigated the effects of multisensory feedback provided during the simulated task, Meyer et al. (2013) found that audio cues (after training) and physical motion cues improved performance on a simulated tracking task. They pointed out the importance of investigating the role of individual sensory cues in the assessments of fidelity of the virtual environment.

In contrast to the abovementioned studies, it has been suggested that many other empirical studies have not consistently shown that a pilot's performance benefits from motion feedback (Pasma et al., 2011). For example, Horey (1992) investigated whether restricted motion capacity would have a negative influence on training by completing sets of manoeuvres in three motion cuing condition (no motion, restricted motion and full motion). The results showed that only 1 out of 11 performance measures was significantly related to motion cuing; the performance improved in restricted and no motion conditions, but not in the full motion condition. Groen et al. (2001) evaluated perceived motion in a simulated take-off task and found that pilots performed the manoeuvre more efficiently with attenuated motion cuing. Go, Bürki-Cohen, Chung, Schroeder, Saillant, Jacobs and Longridge (2003) investigated the effects of motion base on airline pilot training and their results showed that pilots in the no motion condition flew the task more precisely with less effort as compared to the group in the full motion condition; however the full motion group had a faster response to disturbance cues. Vaden and Hall (2005) adopted a meta-analysis approach in order to obtain a true mean effect size for motion platform with regard to fixed wing simulators. Their review of 7 studies provided a small positive effect of motion platform on pilot training effectiveness; however these findings need to be interpreted with caution due to the small sample size of review papers. In the meta-analysis of 24 effect sizes from the transfer-of-training studies, De Winter, Dodou and Mulder (2012) reported that motion cuing seems to be more important for flight-naïve individuals in learning tasks with low dynamic stability, but not for experts in fixed-wing aircraft manoeuvring tasks.

The discussion of the previous research work again provided motivation for one of the research study described in this thesis. One of the aims in this study was to investigate the differences in task performance when the motion cuing is provided and when it is disabled. This study is described in more detail in Chapter 6.

2.5.2 Motion cuing in driving simulations

Studies conducted in the driving simulators also extensively investigated the role of visual-vestibular interaction. Kemeny and Panerai (2003) investigated and evaluated the role of visual and vestibular cues and their interaction during driving simulation and concluded that non-visual sensory modality, such as the vestibular system is important in the perception of motion. The importance of motion platform for effective performance was shown in previous research. For example, Hogema, Wentink and Bertollini (2012) found that the presence of motion cues during low speed turning manoeuvres caused more cautious behaviour during driving and made the driving task more realistic. Feenstra, Bos and van Gent (2011) investigated the effects of motion cueing on steering behaviour and found that the presence of motion feedback decreased the magnitude of the steering correction inputs from the drivers. They concluded that motion feedback can improve drivers' control performance in extreme scenarios, such as a slalom task.

Siegler, Reymond, Kemeny and Berthoz (2001) investigated the contribution of kinaesthetic cues provided by the motion platform when executing driving tasks, such as braking and cornering at the intersections. Motion platform was activated on half of the trials where subjects were asked to perform a simple driving sequence. Their results show that motion cuing modulated breaking behaviour in a positive manner; it prevented drivers to reach overly high decelerations and helped them to achieve a more accurate position by signposts as compared to no motion cuing. However, others have noted that vehicle dynamics such as washout algorithms could negatively influence drivers' performance and subjective feelings of discomfort (Aykent, Paillot, Merienne, Fang & Kemeny, 2011). The conflicting findings from the abovementioned studies indicate that there is a continuous need to evaluate the effects of motion platform in flight and driving simulators.

The literature review of the studies that investigated the effects of motion platforms has provided further theme and motivation for the research work presented in Chapter 6. In this study the effects of motion cuing were investigated trough the recording of the accuracy of the tracking performance during varied task demand conditions. As previous research provided contrasting findings in regards to the effectiveness of motion platforms in aviation training, this research work could provide additional knowledge and support for the use of motion platforms for training.

Through conducting a literature review I identified that there is a lack of studies that conducted a comparative data analysis of objective and subjective measures within aviation research. For this reason the main focus of this research was to use and consider both objective and subjective metrics when investigating user's overall experience during the fidelity assessments. In this way the results from the research studies described in this thesis can serve as a comparative data that could be used to inform future research into the effectiveness of motion platforms.

In the section below I will provide the description of objective and subjective measures that were used during the research work presented in this thesis.

2.6 Performance metrics

A metric is a way of measuring or evaluating a particular phenomenon or a thing. In the field of user experience there are sets of metric specific for certain purpose such as task success, user satisfaction, and error. Using sets of measurements each time something is measured should result in comparable outcomes (Albert & Tullis, 2013).

User experience metrics are used to reveal something about the personal experience of a user whilst using a product or system. They are useful tools as they can reveal something about the interaction in terms of effectiveness (being able to complete a task), efficiency (the amount of effort required to complete the task), or satisfaction (the degree to which the user was happy with his or her experience while performing a task) (Albert & Tullis, 2013). By measuring user experience the metrics can provide structure and design for the evaluation process and give insight into the findings by providing information to decision makers (Albert & Tullis, 2013).

In the sections below I will provide an overview of the objective and subjective metrics used in the experimental studies described in this thesis.

2.6.1 Objective performance measures

The common objective measure of performance in the virtual environment are overall task time (completion time), accuracy of the performance, error rate, response times, dexterity tests, checklists and global rating scales (Moorthy, Munz, Sarker & Darzi, 2003). In the research work described in this thesis the task performance was measured objectively by recording the postural responses (Chapter 3), the overall completion times (Chapter 4 and 5), the error rates during manual assembly task (Chapter 5) and the accuracy of the performance (Chapter 6). However, not all measures were used in every study.

2.6.2 Subjective performance measures

The common subjective measures of performance in VR are usually concerned with user experience, particularly with how the user feels and subjectively perceives VR environment during the interaction. It was suggested that multimodal interfaces used in virtual environments can strengthen the communication between users and computers as these interfaces can effectively stimulate variety of human sensory channels to enhance users' interaction in VR (Santangelo & Spence, 2007). The subjective measures used for the fidelity evaluation in this thesis include the investigation of the sense of presence and immersion, the levels of simulation sickness and the levels of cognitive workload. All of these constructs are common measures that were considered in previous studies when investigating the system effectiveness and an individual's sense of engagement in VR activity (Meehan et al., 2006; Ma & Kaber, 2006; Santangelo & Spence, 2007; Aykent et al., 2011; Faas et al., 2014). In the sections below, each measure will be discussed in more detail in relation to previous empirical research.

2.6.2.1 Sense of presence

An increasing number of researchers starting to accept that the sense of presence experienced during the VR exposure is an important factor in the evaluation process of the effectiveness of the virtual systems (Meehan et al., 2006). Most of the researchers agree that the term *'presence'* refers to the subjective experience of the user. For example, researcher defined presence as the sense of being in one place or environment, even when one is physically situated in another (Witmer & Singer, 1998, Witmer et al., 2005, Faas et al., 2014); a state of consciousness, the (psychological) sense of being in the virtual environment (Slater et al., 1996) the sense of being physically present within a computer-generated or remote environment (Ma & Kaber, 2006); user's subjective psychological response to a VR system (Bowman & McMahan, 2007).

2.6.2.2 Immersion

The definition of the term *'immersion'* however has been subjected to a more controversial debate within the research community. Immersion is either defined as a subjective experience of a user i.e. immersion is a psychological state where one perceives himself as being included in and interacting with an environment that provides a continuous stream of stimuli and experience (Witmer & Singer, 1998; Witmer et al., 2005; Faas et al., 2014; Ma & Kaber, 2006), or as a technological aspect of virtual environment i.e. the objective level of sensory fidelity a VR system provides that is measurable and has many levels such as field of regard, field of view, stereoscopy, display size, display resolution, head based rendering, frame rate and refresh rate (Slater, Linakis, Usoh, Kooper & Street, 1996; Bowman & McMahan, 2007; Schuchardt & Bowman, 2007).

Witmer and colleagues (1998, 2005) suggested that to experience increased levels of the presence, user needs to perceive a sense of involvement and immersion within the VR. In their paper they defined the term presence as *'a psychological state of "being there" mediated by an environment that engages our senses, captures our attention, and fosters our active involvement' [6, p.298]*; the degree of presence depends on many factors such the fidelity of sensory components, the ease of use, the nature of interaction, user's attention, previous experiences and current state. The term immersion is defined as '*a psychological state that is characterized by perceiving oneself to be enveloped by, included in, and interacting with an environment that provides a continuous stream of stimuli and experiences [6, p.299]*.

They suggested that immersion can be increased by interactivity and reduced by extraneous distractors. The term involvement is defined as '*a psychological state experienced as a consequence of focusing one's mental energy and attention on a coherent set of stimuli or meaningfully related activities or events [6, p.298].* The involvement is increased by interactivity and participation and can be decreased by the absence of these factors. Witmer and colleagues concluded that both, immersion and involvement, collectively contribute to the construct called presence (Witmer et al., 1998, 2005). In this thesis, I follow the definitions of the sense of presence, immersion and involvement as suggested by Witmer and colleagues (Witmer et al., 1998, 2005).

The general assumption is that the greater number of human senses stimulated, the greater the capability of the stimulus to produce a sense of presence (Fass et al., 2014; Ma & Kaber, 2006). Empirical research have found that high levels of immersion can increase sense of presence, which can in turn enhance spatial understanding and makes some applications more effective (Bowman & McMahan, 2006; Laha et al., 2012). Meyer et al. (2012) showed that the auditory, visual and motion spatial references could be all used as a measure of presence. Multimodal sensory input was also found to enhance the sense of presence and task performance, but also memory for the objects in VR environment (Meyer at al., 2012, Dinh, Walker, Hodges, Song & Kobavashi, 1999; Nam, Shu & Chung, 2008). Additionally, the presence and immersion were found to have some positive relationship with performance in VR (Ma & Kaber, 2006, Millinel-Pivel & Charron, 2015). For example, Stevens and Kinciad (2015) investigated whether higher presence in virtual simulation training was associated with higher performance by a trainee. They found a moderate relationship between the degree of presence experienced in the simulation and an individual’s performance over three training trials, suggesting that high presence ratings contribute to improve performance. A significant correlation between performance and presence was also found in the studies that used VR as a training platform for US Army soldiers (Yougbult & Hue, 2003) as well as in the architectural design training (Faas et al., 2014).

This research review provided the main incentive to use the sense of presence as one of the construct that is being assessed during the studies described in this research. A more detailed description about the metric used to record the levels of presence is provided in the section 2.7.1.

2.6.2.3 Simulation sickness

One of the main drawbacks in the usability of virtual reality as a research and training platform is the system’s ability to induce undesirable side effects, such as the feeling of discomfort and nausea, also known as simulation sickness. The occurrence of simulation sickness can be explained through the Sensory Conflict theory (Reason & Brand, 1975) which predicts that inconsistencies in what is perceived, what is not perceived and what is expected, create a conflict in vestibular, visual and proprioceptive sensors and as a consequence the feeling of motion sickness is induced (Dahlman, Falkmer & Forsman, 2012). This theory has been supported by research findings showing that latencies in tracking devices create discrepancies between visual displays, vestibular signal, and actual motion, and as such stimulate onset of motion sickness (Meehan et al., 2003; Lubeck, Bos & Stins, 2015).

Apart from displaying three-dimensional images in virtual environments, visual displays are also used to display motion. Sometimes no additional vestibular cues are provided which, according to the Sensory Conflict theory, can give rise to sickness symptoms (Reason & Brand, 1975). Previous experimental research confirmed this in a number of studies where simulated motion significantly increased cognitive and perceptual processing; participants reported increased sickness symptoms such as headache, nausea, eyestrain and difficulty focusing and concentrating (Bruck & Watters, 2009; Dalhman et al., 2012). Lubeck et al. (2015) investigated sickness symptoms between still and moving images and concluded that motion in images is essential for the occurrence of sickness symptoms.

Numerous studies have been conducted in order to investigate which factors can influence the levels of simulation sickness. Extensive research efforts have been made to identify potential factors that can influence susceptibility and occurrence of motion sickness. A research report investigating simulation sickness in virtual environments suggested that the characteristics of the given task, such as degree of control, exposure time, vection, head movements, unusual tasks or manoeuvres are all associated with simulation sickness (Koliansky, 1995). Other factors relate to the technological aspects of VR systems, such as the type of VR devices (Allen, Hanley, Rokers & Green, 2016), visual graphics and the latency of the system in terms of update or refresh rate (Zielinski, Rao, Sommer & Kopper, 2015). Similarly, the individual differences such as age, gender, postural instability, expectancies and previous experience (Dalhman et al., 2012) were also shown to influence the levels of motion sickness.

Findings in some studies showed that it is not only visual motion during the VR exposure but also the types of devices used as a virtual reality platform affected the levels of simulation sickness (Treleaven, Battershill, Cole, Fadelli, Freestone, Lang & Sarig-Bahat, 2015). Simulation sickness has been investigated in many areas where users' actions, altered due to sickness symptoms, affected performance. Stein and Robinsky (2012) found numerous consequences for training in flight operations due to simulation sickness, such as avoidance behaviour in certain flight manoeuvres and poor habit patterns that can be acquired by avoidance strategies and could deteriorate performance and adversely affect airworthiness and flight safety (Stein & Robinsky, 2012). The levels of sickness symptoms arising from exposure to VR systems have been associated with impaired performance in driving simulators (Klüver, Herrigel, Schöner & Hecht, 2016), however when the influence of different washout algorithms (tuning processes) during driving was investigated, the results showed that motion cuing can, in fact, reduce the levels of simulation sickness (Ayken, Paillot, Merienne, Fang & Kemeny, 2011).

Previous research findings have provided direct incentive to include the simulation sickness measure as one of the dependent variable that was measured during the exposure in VR in most of the studies described in this thesis. The levels of simulation sickness were recorded on the questionnaire-based metric that is described in more detail in the section 2.7.2.

2.6.2.4 Cognitive workload

Cognitive workload is another construct that is considered to be an important metric when investigating the effectiveness of virtual simulation in high and low workload conditions. Mental workload, which can be defined as the depletion of human internal resources to accomplish the presented work has long been recognized as an important factor in human performance in complex systems (Leung, Yucel & Duffy, 2010). Many factors have been identified that could influence the levels of cognitive workload during a task in VR. Seaborn, Riecke and Antle (2010) suggested that presenting information in different sensory modalities could reduce the cognitive load during a task. They investigated the effects of visual and haptic modality and although they did not report any significant results in performance between visual and visual-haptic modes, their results indicate that there is a strong coupling between visual and haptic channels during workload conditions.

Within aviation research, an early focus of workload studies was mainly on high demand conditions; the main concern for high demand conditions was to investigate whether the imposed demand will exceed the pilot's capability and response efficiency during these conditions (Hancock et al., 1995). In the studies of cognitive load during simulated flying tasks, it has been argued that when one or more sensory inputs in flight are unavailable, pilots expend more effort and report increased workload. Bell and Grant (2011) examined workload scores with three different cuing technologies (no motion, motion cuing seat and full 6 DOF motion). No significant differences in the overall workload scores were observed, however when examining the individual components of workload they found that pilots experienced significantly lower workload and lower effort using motion seat as compared to 6 degrees motion platform and no motion cuing. They suggested that this might be due to the fact that the motion cues provided by motion seat could provide greater fidelity to those experienced in real systems. Similarly, Pasma et al. (2011) investigated the utility of three flight simulator motion conditions (fixed base, motion-cuing seat and full motion) in order to support training on in-flight rotary wing emergency recovery procedures. The analysis of collected workload and sickness ratings revealed that full motion contributed more to the levels of sickness than motion seat, however the levels of workload were unaffected by motion configurations. When investigating the effect of motion cues on performance during the task with high workload, Wang et al.'s (2013) analysis of the collected workload ratings during the day and night visual environments showed that the presence of motion cuing reduced pilot workload in high demand conditions; without the motion cues, the pilot workload ratings and control activities on control stick and pedals were higher.

This research work prompted the use of cognitive workload metric in the research studies described in Chapter 5 and 6. In these studies participants perform a task in conditions that can be described as high load conditions and therefore the workload metric was administrated to all participants. The metric used to record the levels of cognitive workload is described in more detail in section 2.7.3.

2.6.2.5 Interplay between presence, sickness and workload

All of the psychological measures mentioned in previous section can also be influenced by each other. For example, the level of sickness reported during VR interactions can modulate the level of presence and involvement in VR. However, this association is two-fold. Stanney, Kingdon, Graeber and Kennedy (2002) investigated relationships between exposure time, control presence and sickness and showed that with higher control over their movement, participant's performance and sense of presence improved, but so did the levels of nausea and discomfort (Stanney et al., 2002). Similarly, Saey et al. (2001) found that high FOV produced higher sense of presence but also higher sickness scores. Lin, Duh, Parker, Abi-Rahel and Furness (2002) investigated the effect of four different field-of-views and also found a positive correlation between the SSQ and presence scores. Ma and Kaber (2006) found positive relationship between presence and workload; in their study all VR factors settings that led to greater perception of presence also led to increased perception of mental demand. They concluded that some minimal level of workload may be necessary to develop a sense of presence.

On the other hand, the experimental research investigating the influence of psychological factors, such as effects of presence and workload on simulator sickness in simulated and real driving task showed that increased presence ratings were associated with decreased sickness symptoms whilst workload had no effect on sickness scores (Milleville-Pennel & Charron, 2015). These contrasting findings suggest that the interrelations between these parameters need to be more closely examined and further investigated and thus some of the research work presented in this thesis considers all of these construct (Chapter 5 and 6).

The above described literature review about psychological constructs that can influence user experience such as presence, immersion, involvement, simulation sickness and cognitive workload served as a direct motivation for my research work described in this thesis. The findings from the above mentioned studies suggests that further investigation of the factors that can influence user experience in VR as well as fidelity assessments about the system effectiveness is needed. In the studies described in this thesis I directly investigate how users feel during the exposure and interaction in VR environments through the use of objective performance and subjective questionnaire-based metrics. My main interest was to explore how sensory feedback can affect user experience and therefore I decided to use a range of usability metrics that are described in more detail in the following section.

2.7 Assessment metrics

2.7.1 Assessing the sense of presence

There are many factors that were found to influence the levels of presence and the experimental research literature has proposed several questionnaire-based metrics on how to assess the sense of presence (Yougbult & Hue, 2003). In this thesis the metric chosen for the assessment of the sense of presence was Presence Questionnaire (PQ) (Witmer & Singer, 1998), and therefore only this metric will be described in this section. Witmer and Singer (1998) identified four main factors that can affect the level of presence in VR. These factors include ‘control’ that refers to anticipation and mode of control in a given environment; ‘sensory modality’ that includes environmental richness of the sensory signals; ‘realism’ that refers to graphical realism and consistency with the real world; and ‘distraction’ such as the feeling of isolation and interface awareness. Based on these factors, Witmer and Singer (1998) developed a Presence Questionnaire (PQ) that consists of 32 items that measure a user’s sense of presence and immersion with the virtual environment. It aims to identify the degree of involvement in the virtual experience and allows for the effects of different aspects of the environment and system to be obtained (Nichols, Haldane & Wilson, 2000).

The PQ questionnaire has been subjected to many principal component analyses and most of them produced four factor models. Within the questionnaire four subscales were identified, which include Involvement/Control subscale, Sensory subscale, Adaptation/Immersion subscale and Interface Quality subscale. As the tendency to experience presence is considered a personal characteristic, Witmer and Singer (1998) suggested another factor that could contribute to the overall levels of presence. This measurement was called immersive tendency, which can be a personal tendency to be drawn into an activity, such as reading a book or watching television. As the immersive tendency is considered as a trait Witmer and Singer (1998) developed the Immersive Tendencies Questionnaire (ITQ) that contains 29 items with three subscales: Involvement, Focus and Games. Witmer, Jerome and Singer (2005) have shown that both of these scales are internally consistent measures with high reliability: they reported internal consistency measures of reliability of 0.75 for ITQ and 0.81 PQ using Cronbach's Alpha. Both of the questionnaires were found to be interrelated, i.e. higher scores on ITQ mean higher levels of immersion and presence.

The ITQ and PQ questionnaires (Appendix B and E) have been used extensively to assess immersion and presence in many applications across the range of academic, research and commercial industries (Ma & Kaber, 2006; Faas et al., 2014; Swindells, Po, Hajshirmohammadi, Corrie, Dill, Fisher & Booth, 2004; Stothard, 2008; Steven & Kincaid, 2015). Even though some criticism of the questionnaire has been proposed in terms of definitions (Slater, 2003), Witmer and Singer's Presence Questionnaire remains a standard tool used in the evaluation process of virtual environments (Faas et al., 2014). The ITQ and PQ questionnaires were also chosen for this study, as they are a statistically validated tools by previous research that found the ITQ and PQ to be internally consistent measures with high reliability, which allows for more powerful statistical analysis to be performed (Witmer & Singer, 1998; 2005).

2.7.2 Assessing simulation sickness

Simulation sickness has been identified as one of the main drawbacks of the current VR systems and thus a detailed identification of factors, that can cause or facilitate the levels of simulation sickness, is necessary (Kluver et al., 2016; Lin et al., 2002; Lubeck et al., 2015; Millevel-Pennel & Charron, 2015; Stanney et al., 2002). It has been suggested that simulator sickness can be quantified as a multidimensional construct with several symptom components. Kennedy, Lane, Berbaum and Lilienthal (1993) proposed a multidimensional approach to motion sickness by developing a Simulation Sickness Questionnaire (SSQ) with three main objectives a) to provide a more valid index of simulator sickness severity as distinguished from motion sickness; b) to provide subscale scores that are more diagnostics of the locus of simulator sickness; and c) to provide a scoring approach to make monitoring and tracking straightforward (Kennedy et al., 1993).

The SSQ contains 16 items rated by participants as “none”, “slight”, “moderate”, or “severe”, which assess the oculomotor (eyestrain, difficulty focusing, blurred vision, headache), disorientation (dizziness, vertigo), and nausea (general discomfort, stomach awareness, increased salivation, burping) dimensions of simulator sickness. The scores for each symptom in each subscale are combined by a series of mathematical computations to produce an overall SS score (Brooks, Goodenough, Crisler, Klein, Alley… & Wills, 2010). The meta-analysis of 2100 questionnaires showed that nausea is the most prevalent symptom of motion sickness that is caused by dominant frequencies less than 1Hz (Kennedy, Stanney & Dunlap, 2000). The questionnaire has been used extensively in a variety of settings and applications and is currently the most frequently used measure of sickness symptoms in VR (Kluver et al., 2016; Lin et al., 2002; Lubect et al., 2015; Millevel-Pennel &

Charron, 2015; Stanney et al., 2002; Sparto, Whitney, Hodges, Furman & Redfen, 2004; Seay et al., 2001; Koliansky, 1995; Webb, bass, Johnson, Kelley, Martin & Wildzunas, 2009). The questionnaire can be seen in Appendix C.

2.7.3 Assessing cognitive workload

As the task performed in the studies described in this thesis can be classified as a high load task, the measures of mental and physical workload was needed. The empirical research literature contains a number of scales that were designed to measure cognitive workload during the task. The scale chosen to measures cognitive workload in the studies described in this thesis is the NASA TXL rating scale and thus this section describes this metric in more detail.

The most common measure of mental workload is National Aeronautics and Space Administration Task Load Index also referred to as NASA Task Load Index (NASA TXL) that was first developed by Hart and Staveland (1988). Since then this multidimensional subjective measure with high reliability (r = .83) has been used in a variety of settings and is considered to be one of the most effective measures of perceived workload (Singh, Sharma and Singh, 2005). The NASA-TXL assesses mental workload along six dimensions on scales ranging from 0 to 10 measured in five-point intervals. Three of those dimensions reflect the demands which experimental tasks place on the operator (mental, physical, and temporal demand), whereas the remaining three dimensions characterize the interaction between the operator and the task (effort, frustration, and performance) (Singh et al., 2005). In particular, mental demand measures the mental and perceptual activity required for the task; it refers to the easiness or simplicity of the task. Physical demand measures the amount of physical activity, such as pushing, pulling, turning and controlling that is required for the task. Temporal demand investigates time pressure and the pace of work; it refers to the time pressure experienced during the task, such as how slow or fast the pace is. Effort combines mental and physical strenuousness and refers to how hard participant had to work to accomplish the task. Performance is a self-reported measure of task success; it refers to the user's satisfaction with their performance. Finally, the frustration scale quantifies irritation, stress and discouragement during the task; it refers to the levels of insecurity, irritation or annoyance experienced during the task (Hart, 2006). The full scale of NASA TXL can be found in Appendix F.

The scale has been utilized to provide the global workload score in many areas of research including the aviation and transportation sectors, as well as in medical, training and industrial applications (Klein, Lio, Grant, Carswell & Strup, 2009; Curtis, Dawson, Jackson, Litwin, Meusel …& Winter, 2015; Balaji, Singh, Sodergren, Corker … & Paraskeva, 2015; Singh et al, 2005; Stone, Watts, Zhong & Wei, 2011; Suma, Finkelstein, Reid, Babu, Ulinski & Hodges, 2010; Leung et al, 2010). Rubio, Diaz, Martin and Puente (2004) evaluated and compared NASA TXL to other types of workload scales, such as SWAT and Workload Profile based on intrusiveness, sensitivity, diagnosticity, convergent validity, concurrent validity, implementation requirements, and acceptability. Based on their findings, they suggested that NASA TXL was the most effective in predicting the performance of a particular individual. NASA TXL is therefore used for the assessment of cognitive workload in some of the studies described in this thesis.

2.8 Overview of the research studies in this thesis

In the following chapters, I will describe and present experimental studies that were designed and conducted to evaluate the relative contribution of individual sensory cues as well as investigate the interaction between them. The first study described in Chapter 3, focuses on the presentation of visual information within VR in terms of depth perception by recording postural responses. The second study then directly follows the first study to show in what way the environmental reference points influence postural responses to visual motion and how these could be used to eliminate the negative effects of self-motion illusion.

Chapter 4 focuses on the ways the sensory cues are presented in VR; however this time visual, audio and tactile sensory cues are available as unimodal, bimodal and multimodal feedback during manual assembly task. The presented cues have high informational fidelity however they lack overall physical fidelity and thus the effects on users' objective and subjective fidelity assessments are investigated. Additionally, the ability of additional sensory cuing to mitigate negative effects of motion tracking inaccuracy is also investigated.

Chapter 5 then describes a study where the additional sensory cues where used during virtual training. The main focus of this study was to show whether there are any transferable performance improvements in terms of overall performance. Chapter 6 then focuses on investigating how motion feedback, provided by motion platform in VR, can affect the accuracy performance when various levels of perceived task demand is introduced.

Chapter 3

The effects of visual and environmental factors on postural stability

Chapter overview

This chapter present the first study of this thesis that focused on investigating how the user perception of virtual environments can affect postural stability. The study consists of two experiments where the main focus was to establish which factors can influence the levels of immersion as assessed by the magnitude of postural responses. A short introduction is given first, followed by the results from each experiment and discussion of the overall findings.

3.1 Introduction

3.1.1 Postural stability

The postural control system is ruled by the visual, somatosensory and vestibular inputs from our surroundings that our central nervous system combines and coordinated in order for us to maintain postural stability. Sometimes postural balance can be disrupted by the conflicting information from sensory modalities and cause postural instability, which can give rise to motion sickness (Murata, 2004). Virtual reality applications can serve as a great platform where factors that can affect postural responses can be manipulated and investigated in a controlled environment. Studies have shown that postural instability precedes motion sickness and it is therefore important to investigate factors that can have an effect on postural sway (Villard et al., 2008). Behavioural data recorded in the studies of postural sway show that the magnitude of the visually evoked postural responses can be influenced by many factor. Some of the main factors are the stimulus characteristics such as area, velocity and spatial frequency. Environmental factors such as the base of support, fixation point and the information coming from other sensory modalities were also shown to influence the postural stability (Berthoz et al., 1979, Peterka, 2000, Bronstein & Buckwell, 1997, Meyer et al., 2012; 2013). As the postural adjustments to the real environment are performed unconsciously, experimental research has suggested that the investigation of postural stability during VR exposure could serve as a complementary objective measure of presence and immersion (Meyer et al., 2013).

Bronstein and Buckwell (1997) first suggested that visual stimulus is a key factor, which can contribute to postural instability. In their study participants maintained a stable stance during background motion in parallel with the screen, however when foreground object was introduced the postural sway in opposite direction was recorded. They argued that the control of postural sway does not result from rigidly wired-up, optokinetic reflexes but is modulated by fixation point and the configuration of the environment. To test how environmental factors can affect postural sway, Meyer et al. (2013) conducted an experiment similar to Bronstein and Buckwell, (1997) to investigate visually induced postural responses (VEPRs) and added another two conditions where virtual foreground object was presented during background and foreground fixation. In the ‘real’ condition a physical teapot was presented on a stand; in the virtual condition a virtual reality-matching counterpart (virtual teapot) was presented on the stand in the same position as the real one (Figure 1). They matched previously reported findings for real objects (Bronstein & Buckwell, 1997) and further observed significant systematic difference in postural sway between real and virtual objects. As the postural responses are unconscious that postulated than investigation of VEPRs can be done as an objective evaluation of presence to compliment subjective measures. This study provided a direct motivation for conducting the research work described in this chapter. To allow for the diret comparison with the previous research the same experimental design was used. However, this study extend previous research by looking closely at how the visual information, in particular visual disparity cues that enable depth perception in VR, influence the postural responses to visual motion. As the Meyer et al. (2013) study suggested that visual reference points can modulate self-motion illusion, the second study directly explores in more detail in what way can the environmental reference points be used to mitigate the negative effects of self-motion illusion.

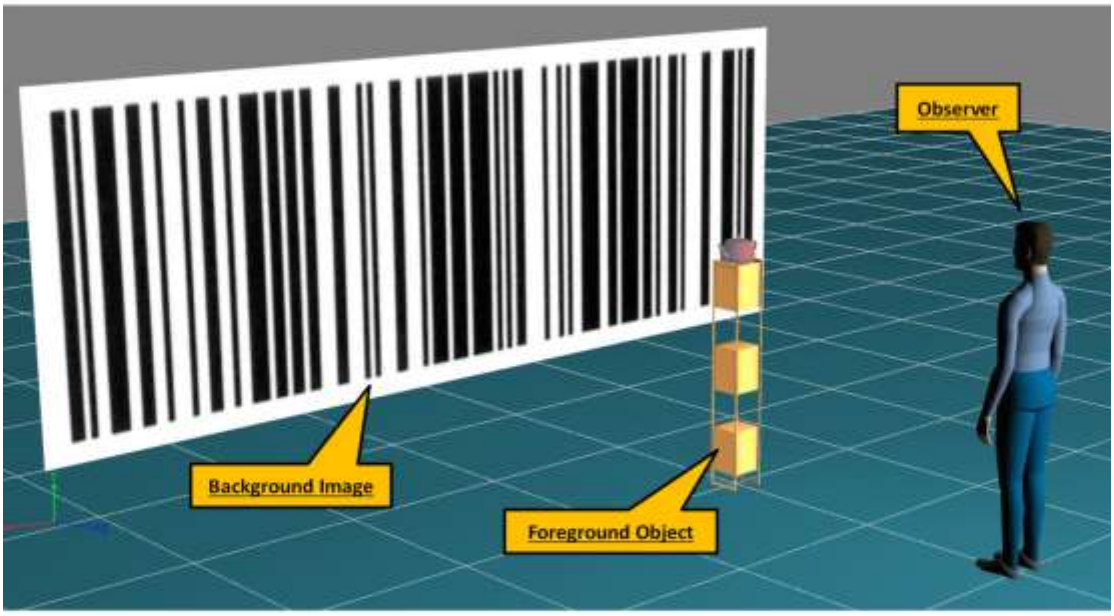


Figure 1. Virtual set up in Meyer et al., (2013) study – the same set was used in the experiments described in this chapter. The main elements of the VR set up were the translating background image (a barcode), the foreground object (a virtual or real teapot (or helium balloon in this study) on a stand and an avatar representing the observer in the scene.

The main aim of this study was to extend the Meyer et al. (2013) study and further explore how the environmental factors in the virtual and real environments can influence the levels of immersion and postural stability. The first part of the study described in this chapter investigates how visual disparity cues that are present in 3D stereoscopic displays can affect user's levels of immersion. In the second part, the main objective is to examine how the explicit assessment in the positional stability of the real objects presented in the scene during VR exposure modulates the levels of immersion. In both of these studies, postural stability, measured by the visually evoked postural responses, is used as a measure of immersion.

3.1.2 The disparity cue presentation: accommodation vs. vergence

Virtual Reality (VR) environments, almost without fail, provide stereoscopic visual cues, either via projection systems or head-mounted displays. These 3D displays present images on a single plane, typically the projection screen. Depth perception of objects in the display is achieved by introducing disparities in the retinal images of the left and right eye that would be observed is the simulated scene were observed in reality (Figure 2). The perception of depth however, is quite different in a natural environment. In a real world, in order to see objects clearly the eyes focus and converge by the amount that depends on the distance between the object and us. Vergence is defined as the movement of the eyes in opposite direction (convergence) in order to locate the object of interest. Accommodation can be defined as the movement of the lens that is adjusted and driven by the blur in the retina, in order to focus on the object (Lambooij, IJsselsteijn & Heynderickx, 2010). In natural world, both of these processes are synchronised: the viewer adjusts the vergence of the eyes to look at the object and the eyes focus to sharpen the retinal image (Shibata, Kim, Hoffman & Banks, 2011).

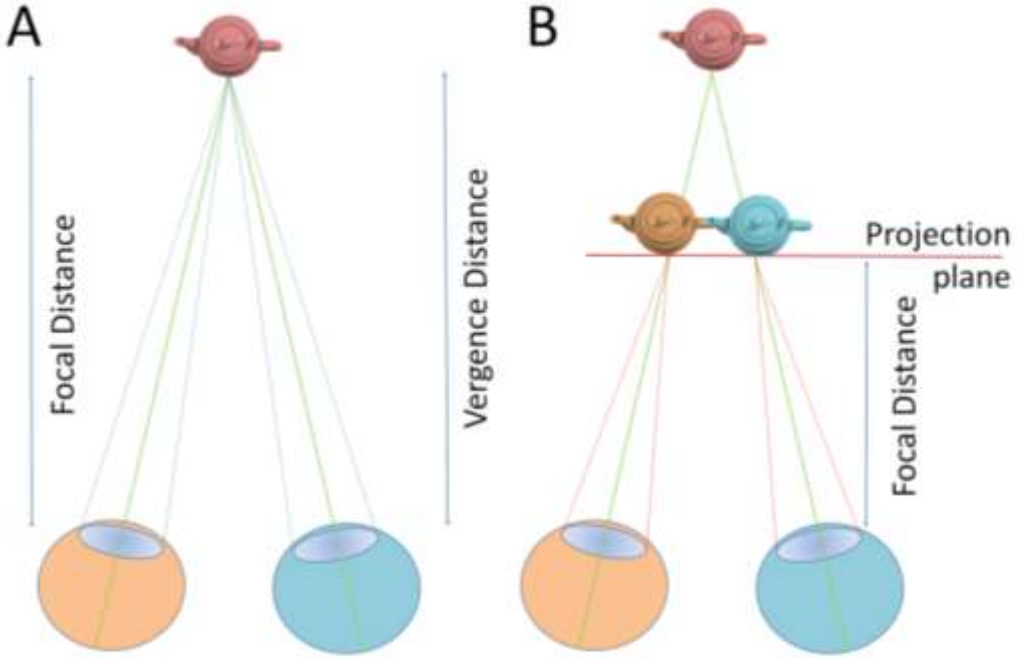


Figure 2. In natural environments accommodation (determined by focal distance) and vergence are linked (a). In most VR environments depth cues are generated by providing disparity cues such that objects are rendered separately for each eye. Accommodation stays on the projection plane, so that conflicting cues are presented which affect performance and comfort.

On the other hand, a visual stimulus provided by 3D stereoscopic display differs from that of the real world because the image provided is produced on a flat surface (Hoffman, Girshick, Akeley & Banks, 2008). The distance from the eyes to the screen is always the same however when disparity cues are introduced virtual objects are rendered as being either in front or behind the projection screen (Figure 2). While this procedure provides appropriate disparity cues, it does not provide the necessary accommodation cues. Typical systems, additionally, render all objects in a scene ‘in focus’ so that additional blur cues that would be seen for objects outside of the focal plane are also missing in VR systems. As eyes try to view the virtual objects and structures they need to converge and focus at different distances. This discrepancy in viewing distances is referred to as vergence-accommodation (VA) conflict that was found to be a major drawback of many new VR technologies, in particular in head mounted displays (HMD) (Lambooij et al., 2010; Shibatta et al., 2011; Hoffman et al., 2008). For example, Shibata et al. (2011) investigated the effects of viewing distance during VA conflict and reported that subjects experienced increased discomfort and fatigues during cross and uncrossed disparities at different distances. Similarly, Hoffmann et al. (2008) showed that vergence–accommodation conflicts arising in conventional 3D displays hinder visual performance and cause fatigue.

Kramida (2016) proposed a series of possible solutions for vergence-accommodation conflict in head mounted displays (HMD) however he concluded that as VAC still remains a major factor contributing to discomfort and visual fatigue, especially in the near task in VR and AR applications. Therefore, it is necessary to further investigate and evaluate VR interfaces in order to minimise the negative effects of VAC on user's behaviour and performance. The vergence-accommodation conflict is directly considered as a possible explanation of the findings reported in this study, which is addressed in the section 3.4.1.

3.1.3 Summary of experimental aims

Both of the studies presented in this chapter follow the same experimental set up as that of Meyer et al. (2013) where visual motion of a background stimuli, presented as black and white image of a barcode, is presented whilst real and virtual foreground objects are presented in the foreground. In the previous experiment Meyer et al. (2013) showed that postural responses differed for virtual and real foreground objects in a non-optimised VR environment. This study extends the previous study by manipulating the ways in which visual disparity cues are presented and by introducing another real object into the environment as a reference point.

Experiment 1 explores the effect of 3D parameters; in particular the study focus is on how the position of visual disparity cues that are provided in 3D stereoscopic displays modulated user’s behaviour and comfort. As it was previously mentioned that visual disparity cues might cause visual fatigue and discomfort due to the vergence-accommodation conflict, the purpose of this study is to explore whether the rendered position of visual disparity cues influences user’s levels of immersion by examining postural stability. Two experimental blocks are designed: they differ in the way the visual disparity cues are presented. In the first experimental block (Experiment 1a) the position of the background visual stimulus is designed to lie exactly on the plane of the projection system, whilst real and virtual foreground objects are presented and rendered as being 2 meters in front of the observer. In the second experimental block (Experiment 1b), background visual motion is rendered as being 2 meters behind the projection screen whilst virtual and real foreground objects are presented directly on the projection screen (Figure 1). Consistent and inconsistent vergence and accommodation cues are present in both experimental blocks.

In Experiment 2 the main focus is to investigate whether the physical properties of environmental anchors influence user’s postural responses. In the first experimental block (Experiment 2a) a real teapot is presented as a stable reference point in the scene and the magnitude of the postural sway is recorded. In the second experimental block (Experiment 2b) a helium balloon is presented as a ‘stable’ reference point. The physical properties of the two objects are quite different: teapot is stable object that will remain in the same place unless is moved with physical force; the physical properties of a helium balloon on the other hand are different: balloon can move freely in the air and as such might not be perceived as stable object. The main aim of this study is to investigate whether participants consider these differences when using foreground objects as environmental anchors during the perceived self-motion. The magnitudes of postural sway in response to the visual motion during the teapot and the helium balloon presentation are computed and compared. The data are partially consistent with previous research and show that the positional stability of the environmental anchors might play a role in the assessments of environments during the adjustment of postural responses.

3.2 Methods

3.2.1 Participants

A power analysis, using G*Power 3.0 software tool (Faul et al., 2009), was conducted with previously recorded data (Meyer et al., 2013) and showed that a sample size of 8 participants was required to achieve a power of 0.30 and a significance level of 0.05 with a large (0.50) effect size. In Experiment 1, 8 participants were recruited via opportunity sampling with the age ranging from 18 to 48, (M = 25.3, SD = 12.19), 6 males and 2 females. For the second experiment a power analysis revealed that a sample size of 13 participants is required to achieve a power of 0.6 and a significance level of 0.05 with a large effect size (0.50). In Experiment 2, 13 participants were recruited via opportunity sampling with the age ranging from 20 to 58, (M = 29.9, SD = 15.45), 4 females and 9 males. All participants signed a consent form and reported normal or corrected-to-normal vision and hearing. The main reason for the low sample sizes in these studies were time constrains for the use of VR set up at VEC.

3.2.2 Virtual reality set up

The laboratory consists of a planar display screen of 6.0m in length and 2.1m in height behind which are two active stereo projectors that create 3390 x 1200 resolution images at a rate of 120Hz. 3D stereo images are produced by an NVIDIA Quadro K6000 GPU. Observers wear wireless LCD shutter glasses that are synchronized with the projectors to provide stereoscopic images. The position of the shutter glasses is tracked using 16 VICON Bonita B10 infrared cameras (250 fps capture speed, motion resolution of 0.5mm of translation and 0.5 degrees of rotation in a 4m x 4m volume using 9mm markers). Position data, computed using VICON Tracker software, is broadcasted in real-time across the internal network using a VRPN protocol at a rate of 200Hz and used to update the virtual environment. The position of the 3D-glasses in space was also tracked and recorded to provide the head position data used in the analysis of postural sway.

The virtual reality environment (Figure 1) consists of three major elements:

The background image: an image of a barcode extending across the 6m x 2.1m display screen, at a distance of 4m in front of the observer (86˚ visual angle horizontally).

The foreground object: a 3D geometric model of a teapot sitting on top of a stand, 2m away from the observer subtending 4.6˚ visual angle (Figure 3).

The observer: a virtual camera whose position and orientation are directly linked to that of the real-life observer through the tracking system to dynamically render images on the display screen that will appear to the observer as a true perspective of the scene. The avatar shown in Fig 2 is not visible (Meyer et al., 2013).

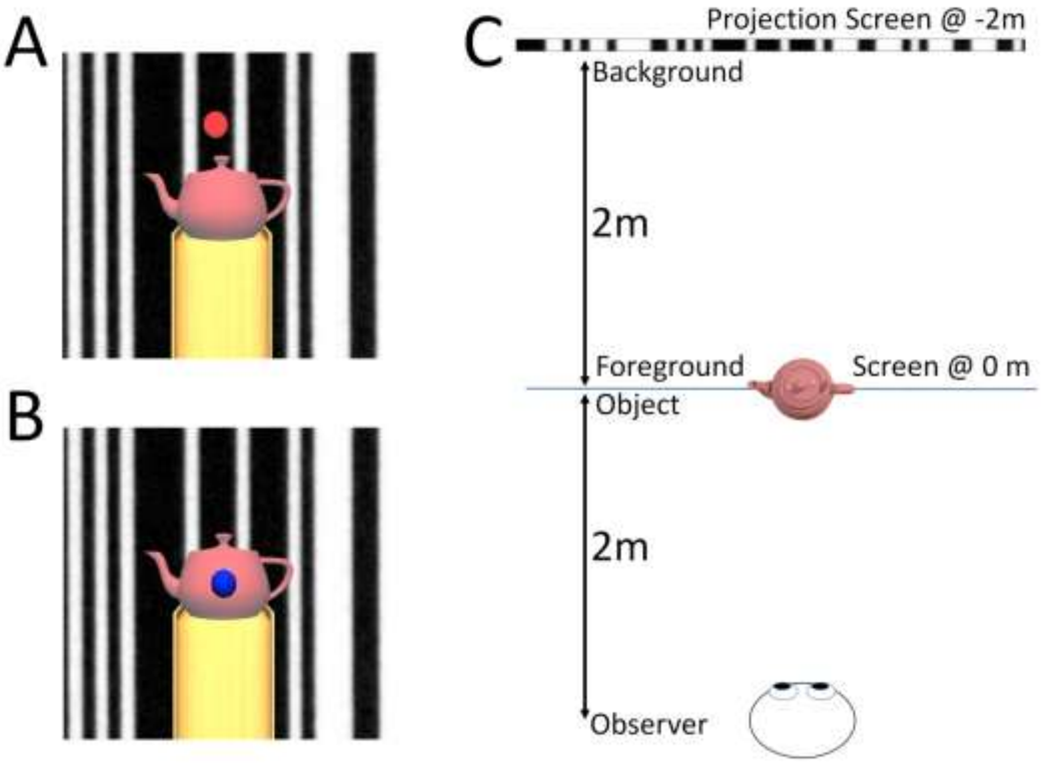


Figure 3. The stimuli from the viewer's perspective. Participants were asked to press a button when the fixation target (either a red dot on the background image (A) or a blue dot on the foreground object (B)) transiently disappeared. Panel C shows a plan view of the room layout: Participants viewed a moving image of a barcode 4m ahead of them. A real or virtual teapot on a stand 2m from the observer was used as a foreground object. We tested two conditions: in one condition the projection screen was placed 2.0 metres in front of the participants, such that the foreground object had matching accommodation and vergence cues while in a second condition the projection screen was 4.0 m from the observer such that the background was 'correctly' displayed.

3.2.3 Stimulus

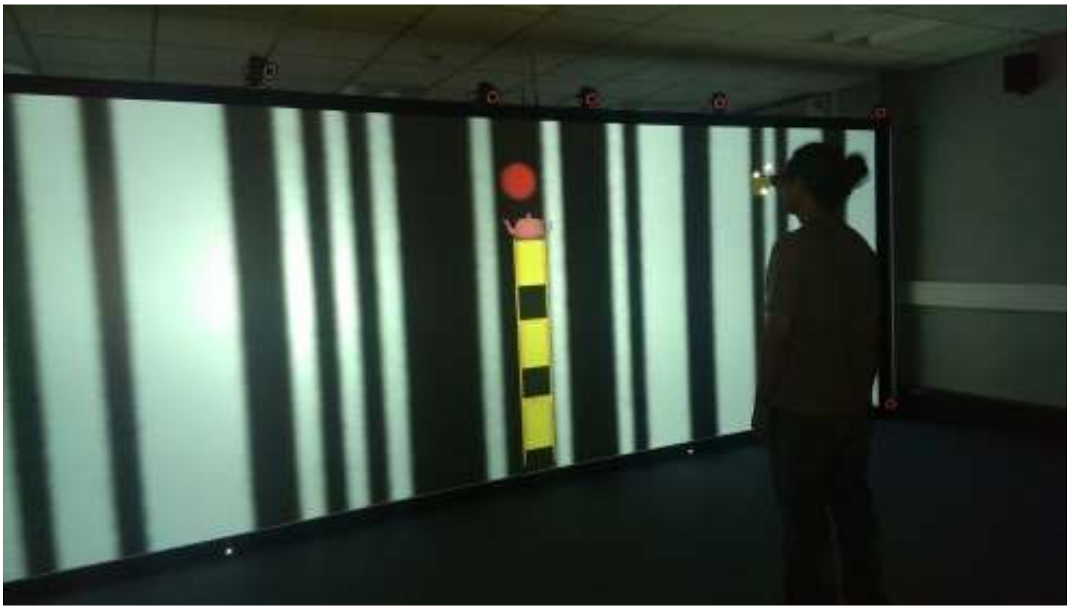
In a previous experiment, Meyer et al. (2013) showed that the movement of the background image in the lateral motion causes highly stimulus dependent VEPRs. In this study, the virtual environment was manipulated in two experiments each consisting of two experimental blocks containing medio-lateral visual motion of the background image and an image of a barcode to provide a rich motion stimulus. The order of presentation was counterbalanced across the participant's pool. Each block lasted 20 minutes and participants had a minimum of 20 minutes rest between the two successive blocks. Within each block, five conditions were run that consisted of 20 presentations of a 6 second visual motion pattern that preceded 6 seconds of a static display (rest condition) during which natural postural sway was measured (Bronstein & Buckwell, 1997; Meyer et al., 2013). The dynamic characteristics of the visual stimulus follows previous studies (Meyer et al., 2013) and was chosen to be within the range of those present in spontaneous body sway, consistent with the aim to present motion signals that would be interpreted as being caused by self-motion. The motion signals consisted of three raised-cosine oscillations at a frequency of 0.5Hz and maximum amplitude of 50mm (0.7˚ visual angle in the lateral motion condition) relative to the origin.

3.2.4 Task procedure

Upon arrival to the Virtual Engineering Centre (VEC), each participant read an information sheet and signed a consent form. Then the participants were led into a room where the experiment took place and the whole procedure of the task was explained. The participants were required to put on LCD shutter glasses and were asked to stand on two foam paddings with both feet approximately 20cm apart. The reason for the use of foam padding was to enhance postural sway adjustments to the visual information that has been previously shown to serve as a reference for postural stability (Mergner, Schweigart, Maurer & Blümle, 2005). Participants were asked to keep as still as possible and fixate on the target. The distance from the projection screen was varied: during the 0m experimental block they were standing 2m from the projection screen; during the -2m experimental block they were standing 4m from the projection screen (Figure 4). Streepey, Kenyon and Keshner (2007) have previously reported that the information in peripheral vision was used to stabilise posture even when visual motion promoted postural instability. When participants were standing 4m from the projection screen, the edges of the screen were visible in the peripheral field of view. The field of view was restricted by “blinkers” attached to the 3D shutter glasses to ensure that the lateral extend of the VR display was comparable in both conditions.

The fixation target was either a red or blue dot and participants were asked to click the clicker when the fixation target transiently disappeared. The disappearance of the fixation point was elicited at random intervals by the experimenter (about 25 times per condition) in all conditions during moving and static display. There was approximately 1 event every 5 seconds and with at least 1 second minimum gap between events. The main reason for asking participants to fixate and report disappearance of the target was to ensure that they maintain focus on the appropriate area of the virtual display.

A



B

Figure 4. Study set up. Picture A shows a participant fixating on the background whilst a virtual foreground is present (-2m condition). Picture B shows a participant fixating on the background whilst a real foreground object is present (0m condition).

3.2.5 Data analysis of VEPRs

Head position data was recorded using a VICON motion tracking system by tracking IR reflective markers on the 3D shutter glasses worn by the participants. To quantify the VEPRs from the recorded postural way data, a natural sway component needs to be accounted for. It has been demonstrated that modulation by lower-limb proprioceptive signals leads to very low frequency sway (<1Hz) and 95% of the energy in natural sway is at a frequency less than 2Hz (Meyer et al., 2013; Dakin, Son, Inglis & Blouin, 2010). To account for this the recorded data was filtered off-line using a second order zero-lag Butterworth band-pass filter with cut off frequencies of 0.125Hz and 2.5Hz and sampled at 10Hz (Meyer et al., 2013). The standard deviation of the filtered data from each trial were then calculated for medio-lateral position so that outlying and thus potentially contaminating data might be removed before continuing with further analysis, however it was not necessary to disregard any trials.

To quantify the VEPRs, the proportion of motion energy at the visual stimulus frequency (0.5Hz) relative to the total energy in the spectrum was computed (eqn. 1) during the stimulus condition and is compared to the magnitude of natural postural sway at this frequency during the rest. The analysis windows in both cases extended over the full 6 seconds of stimulus or rest (Meyer et al., 2013).

To determine which sway was visually evoked and while accounting for the natural sway component of the recorded data, the VEPRs were quantified by comparing the values for visual motion and static display periods of each condition, using:

$$VEPR_v = \frac{F(v)}{\sum_{w=0}^{Fs/2} F(w)}$$

where *VEPR* is the amplitude of the Fourier component, *F(v),* at the visual stimulus frequency, $v = 0.5Hz$, relative to the total energy in the spectrum between 0Hz and half the sampling frequency ($F_s/2 = 5Hz$). To quantify visually evoked postural responses we compared this measure for the 6 seconds stimulus presentations to the 6 second pause immediately following the stimulus. To estimate the mean and variance of the population responses we used standard bootstrapping to estimate the sampling distribution (Efron & Tibshirani, 1986). The means and variance estimates are based on 1000 resamples of the observed dataset. All data reported hereafter are derived from the estimate means and standard deviations produced by this procedure (Meyer et al., 2013).

To estimate the mean and variance of the population responses we used standard bootstrapping (Efron & Tibshirani, 1986) to estimate the sampling distribution by resampling the postural sway data of our participant pool. For all data reported here the means and variance estimates are based on 1000 resamples of the observed dataset, obtained by random sampling with replacement of the population data (Meyer et al., 2013).

3.2.6 Statistical analysis

Both of the experiments shared same designs and analysis. Experiment 1 tests VEPRs in five experimental conditions while Experiment 2 measures VEPRs in three experimental conditions. Visual background motion provided in medio-lateral direction is presented and visually evoked postural responses are measured in both directions.

The analysis is conducted in two stages:

1st stage: VEPRs should match the direction of the visual stimulation: lateral head motion is expected for laterally moving visual stimuli.

Experiment 1: An ANOVA with the factors ‘experimental condition’ (5) x ‘visual stimulus motion’ (on/off, 2) and ‘subject’ as a random factor was computed for each of the two orthogonal head motion directions. Planned comparisons were only carried out where significant main effects of ‘visual stimulus motion’ or interactions were evident.

Experiment 2: An ANOVA with the factors ‘experimental condition” (3) x ‘visual stimulus motion’ (on/off, 2) and ‘subject’ as a random factor was computed for each of the two orthogonal head motion directions. Planned comparisons were carried out where significant main effects of ‘visual stimulus motion’ or interactions were evident.

For both experiments, after the initial sets of ANOVAs, a bootstrapping technique was used to estimate the population response from the individual data (Efron & Tibshirani, 1986). Bootstrapping is a statistical technique that allows estimation of population parameters when the shape of sampling distribution in not known. The sample data are treated as a population from which smaller samples (bootstrap sample) are taken. By performing this procedure a number of times, the sampling distribution can be estimated and subsequent statistical analysis can be performed (Field, 2013).

2nd stage: Visual stimulus motion should cause significant increases of the VEPR relative to the pause immediately following the stimuli (our control condition). Planned comparisons, using one-tailed t-tests, are therefore used.

Experiment 1: Seven planned comparisons were conducted; one for each of the five experimental conditions to investigate the difference between X VEPRs and pause, and two additional direct comparisons of the VEPR for matching real and virtual foreground conditions. Bonferroni adjusted alpha levels of 0.007 per test (.05/7) were used for hypothesis testing. All statistical tests on these population estimates were performed using independent (two-sample) t-tests, because bootstrapping was used to estimate the sample population mean and variance (Field, 2013).

Experiment 2: Four planned comparisons were conducted: one for each of the three experimental conditions to investigate difference between X VEPRs and pause, and one for direct comparison of the VEPR for real foreground object. Bonferroni adjusted alpha levels of 0.0125 per test (0.05/4) were used for hypothesis testing. All statistical tests on these population estimates were performed using independent (two-sample) t-tests (Field, 2013) because bootstrapping was used to estimate the sample population mean and variance (Efron & Tibshirani, 1986).

3.3 Results

The result of Experiment 1 will be presented first, followed by a short summary. Then the results of the Experiment 2 will be presented and a short summary will be also provided. In the next section a discussion of the findings from both experiments (Experiment 1 and 2) will be presented.

3.3.1 Experiment 1. The effect of visual disparity cues on postural sway presented through the manipulation of the position of convergent plane.

3.3.1.1 Main aims

The main aim of Experiment 1 was to investigate whether visually evoked postural responses (VEPRs) for real and virtual foreground objects will differ when the location of the fixation point is projected either on or behind the convergent screen plane. The order of the experimental block was counterbalanced across participants. All participants completed all of the conditions in both experimental blocks. In the first experimental block, background stimulus was presented directly on the projection screen (0m condition) whilst participants were standing 4 meters from the projection screen. A real object was placed 2 meters from the screen and 2 meters in front of the observer whilst the virtual foreground object was rendered as being 2 meters in front of the screen to match its real counterpart (Figure 4 B). In the second experimental block, participants were standing 2 meters from the screen and the background image was projected as being 2 meters behind the projection plane (-2m condition). The virtual foreground object was rendered directly on the projection screen and the real object was positioned as close to the projection screen as possible (Figure 4 A).

All participants completed all conditions in both experimental blocks. Due to the virtual reality set up the order of the conditions could not be changed, so all participants completed the conditions in the same order. The order of the experimental blocks was counterbalanced across participants.

3.3.1.2 Experimental conditions

Each experimental block consisted of five conditions:

1. BG Only - fixate a target on the background (BG), no foreground object present.
2. BG virtual - fixate a target on the background, a virtual foreground object is present.
3. BG real - fixate a target on the background, a real foreground object is present.
4. FG virtual - fixate a target on the virtual foreground (FG) object.
5. FG real - fixate a target on the real foreground object.

3.3.1.3 Experimental results

The behavioural performance of participants was recorded as the amount of correct clicker responses when the fixation point transiently disappeared. Overall participants correctly identified more than 95% of the events (M = 98.15, SD = 1.55). This means that the attention or fixation position as a cofounding variable can be rejected.

3.3.1.4 Experiment 1a. Postural response at 0 meters convergence plane.

The postural sway was investigated in lateral direction in both experimental blocks. In the first experiment the position of the moving background coincided with the projection screen (Figure 4 B). This condition provides no vergence-accommodation conflicts at the moving background pattern, except for the FG virtual condition. In conditions where the real foreground object was presented as a visual anchor, VEPRs that follow the original data described in Meyer et al. (2013) and Bronstein and Buckwell (1997) were expected and observed.

The repeated measures analysis of variance (ANOVA) was conducted with the factors ‘condition’, ‘visual motion’ and ‘participant’ as a random factor. The analysis showed significant main effects of visual motion (F (1,63) = 10.83, p = 0.0016, η^2 = 0.52), significant main effect of condition (F (4,63) = 2.93, p = 0.027, η^2 = 0.48) and significant interaction between visual motion and condition (F(4,63) = 3.78, p = 0.008, η^2 = 0.68). An analysis of the VEPR component of postural sway was conducted at 0.5Hz in comparison to the same frequency component when no motion is present (pause). Seven planned comparisons, with Bonferroni correction, where conducted to investigate difference in VEPRs between X VEPR and pause. Figure 5 shows that postural responses to visual motion differ between the points of fixation.

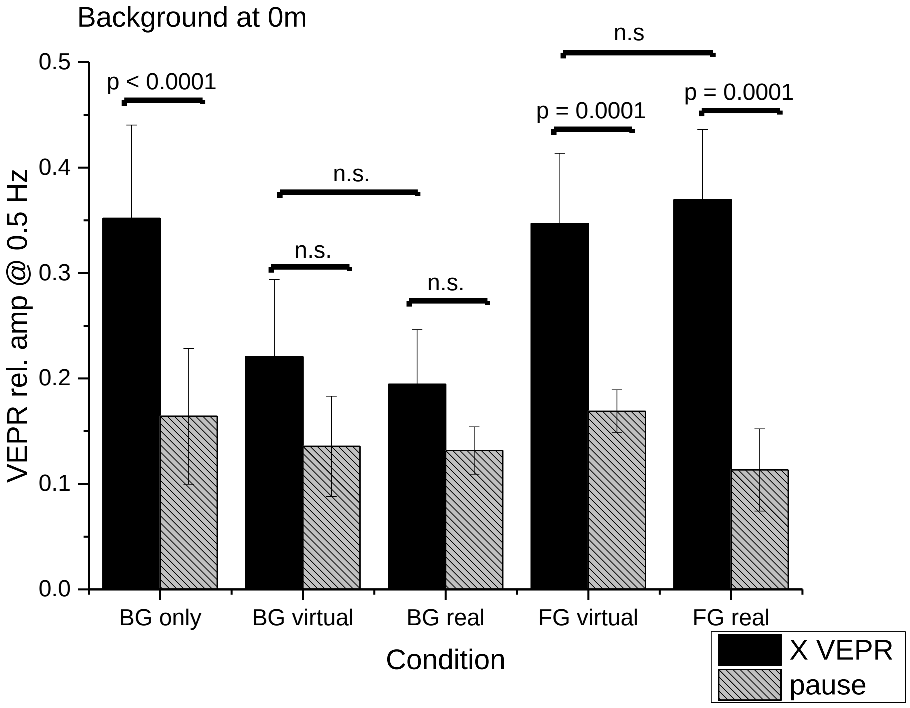


Figure 5. Relative amplitude of postural responses to visual stimulus in motion (black bars) and in pause (grey bars) in 0m experimental block in every condition. The labels identify the fixation point and whether the foreground object was virtual or real.

The summary in Table 1 shows the results of planned comparisons between visual motion and pause in each condition. Comparisons where there are significant differences between the visual stimulation and rest condition are highlighted in bold. Significant differences were recorded in VEPRs during the background fixation when no foreground objects were presented (BG only). No significant differences between postural sway and pause during background fixation were recorded when real or virtual foreground objects were presented but not attended (BG virtual and BG real). The amplitude of the VEPRs in both conditions was significantly reduced relative to BG only condition. No significant difference in the amplitude of the postural sway was observed between BG virtual and BG real condition. The phase in both conditions was similar and in the anti-phase to the phase observed in the BG only condition. When fixation was shifted to the foreground objects significant differences were observed during virtual and real object fixation (FG virtual and FG real). When participants fixated on the foreground object whilst the background was moving, significant VEPRs were recorded in anti-phase to the BG only condition. The sway amplitude did not differ between real and virtual foreground objects or that observed in BG only condition.

Table 1

Lateral sway energy at 0.5Hz during visual stimulation in 0 meter condition

Condition	% Energy at 0.5 Hz Visual Motion Mean (SD)	% Energy at 0.5 Hz Rest Mean (SD)	Difference	Effect size Cohen's d
BG only	35.19% (8.8%)	16.4% (6.4%)	$t_{(14)} = 4.85$, $p < 0.0001$	2.34
BG virtual	22.07% (7.3%)	13.6% (6.0%)	$t_{(14)} = 2.75$, $p = 0.0155$	1.33
BG real	19.45% (5.2%)	13.17% (5.2%)	$t_{(14)} = 3.14$, $p = 0.0081$	1.57
FG virtual	34.70% (6.7%)	16.9.0% (4.4%)	$t_{(14)} = 7.23$, $p = 0.0001$	3.61
FG real	36.95% (6.7%)	11.3% (4.9%)	$t_{(14)} = 9.38$, $p < 0.0001$	4.69

Table 1: Lateral Sway energy at 0.5Hz during visual stimulation and rest and comparison statistics for the target and control conditions in Experiment 1, background at 0m. Comparisons where there are significant differences between the visual stimulation and rest condition (Bonferroni corrected alpha levels of $p < 0.007$, (0.05/7)) are highlighted in bold. The effect sizes (Cohen's d) indicate the magnitude of the observed differences in means.

3.3.1.5 Experiment 1b. Postural response at -2 meters convergence plane

Changing the environment configuration, such that the foreground object is on the projection screen while the moving background image is presented 2 meters behind the screen, changes VEPR patterns recorded for the same participants. In Experiment 1b the configuration of the virtual reality set up was manipulated: the background visual motion was rendered on the projection screen as being 2 meters behind (-2m condition). The virtual and real foreground objects were presented directly on the projection screen (Figure 4 A). A repeated measures ANOVA for the lateral sway data was conducted with factors ‘condition’, ‘visual motion’ and ‘participant’ as a random factor and showed a significant main effect of visual motion ($F(1, 63) = 11.98$, $p = 0.001$, $\eta^2 = 0.6$) for postural sway in the direction the stimulus (lateral sway). No significant main effect for ‘condition’ was observed ($p = 0.549$, $\eta^2 = 0.15$) and no significant interaction was found between visual motion and condition ($p = 0.327$, $\eta^2 = 0.23$). Overall, these findings suggest that visually evoked postural responses were reduced in the -2m experimental block. The graphical representation of the results between the visual motion and pause are shown on Figure 6.

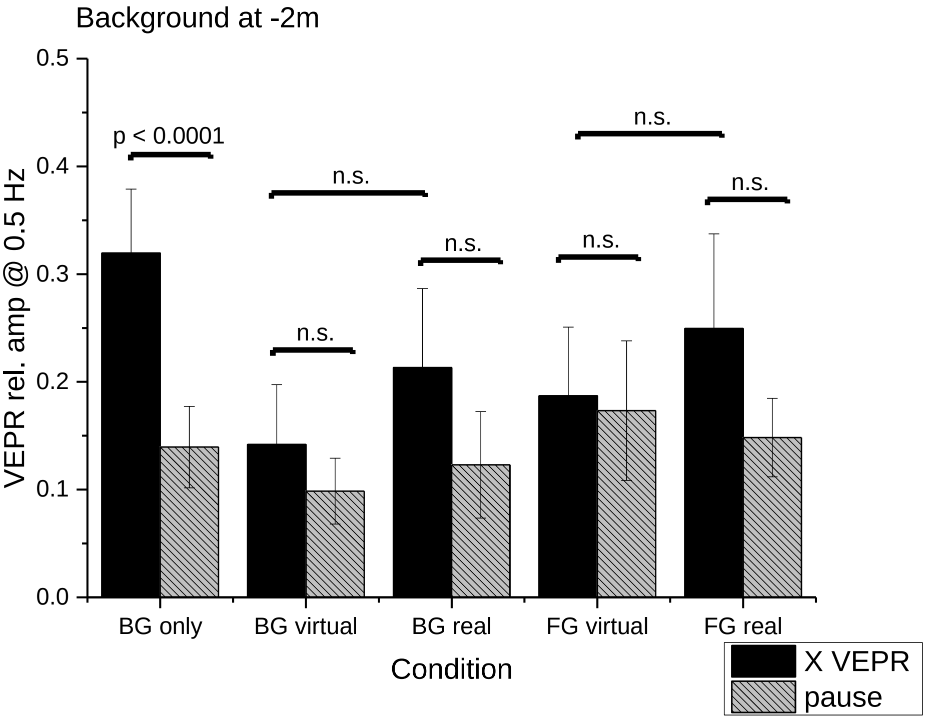


Figure 6. Relative amplitude of postural responses to visual stimulus in motion (black bars) and in pause (grey bars) in -2m experimental block in every condition. The labels identify the fixation point and whether the foreground object was virtual or real.

The results of planned comparison are summarised on Table 2. In this environmental configuration, our results show significant VEPRs during background fixation whilst no foreground object is present. The VEPR component of postural sway was reduced during background fixation whilst real and virtual foreground objects were present in the environment. The amplitude of the VEPRs was significantly reduced as compared to BG only condition. The phase in both conditions was similar and in anti-phase to the phase observed in the BG only condition. No significant VEPRs were recorded when fixation was shifted to the real and virtual foreground objects. There were significant VEPRs observed in anti-phase to the BG only condition. The sway amplitude did not differ between real and virtual foreground object and for the BG only condition.

Table 2

Lateral sway energy at 0.5Hz during visual stimulation in -2 meter condition

Condition	% Energy at 0.5 Hz Visual Motion Mean (SD)	% Energy at 0.5 Hz Rest Mean (SD)	Difference	Effect size (Cohen's d)
BG only	31.95% (5.9%)	13.95% (3.8%)	$t_{(14)} = 7.25$, $p < 0.0001$	3.61
BG virtual	14.18% (5.7%)	9.86% (3.1%)	$t_{(14)} = 1.95$, $p = 0.074$	0.96
BG real	21.22% (7.3%)	12.3% (4.9%)	$t_{(14)} = 2.88$, $p = 0.012$	1.44
FG virtual	18.71% (6.4%)	17.3% (6.5%)	$t_{(14)} = 0.42$, $p = 0.642$	0.21
FG real	24.95% (8.8%)	14.8% (3.6%)	$t_{(14)} = 3.01$, $p = 0.009$	1.5

Table 2: Lateral Sway energy at 0.5Hz during visual stimulation and rest and comparison statistics for the target and control conditions in Experiment 1, background at 0m. Comparisons where there are significant differences between the visual stimulation and rest condition (Bonferroni corrected alpha levels of $p < 0.007$, (0.05/7)) are highlighted in bold. The effect sizes (Cohen's d) indicate the magnitude of the observed differences in means.

3.3.1.6 Comparison between two experimental blocks: Experiment 1a and Experiment 1b

The comparison between two experimental blocks was performed in each condition, in order to investigate which had significantly larger postural responses to visual motion. The graphical representations of the results between the visual motion are shown on Figure 7. Due to the bootstrapping procedure planned comparisons were performed using two samples t-tests.

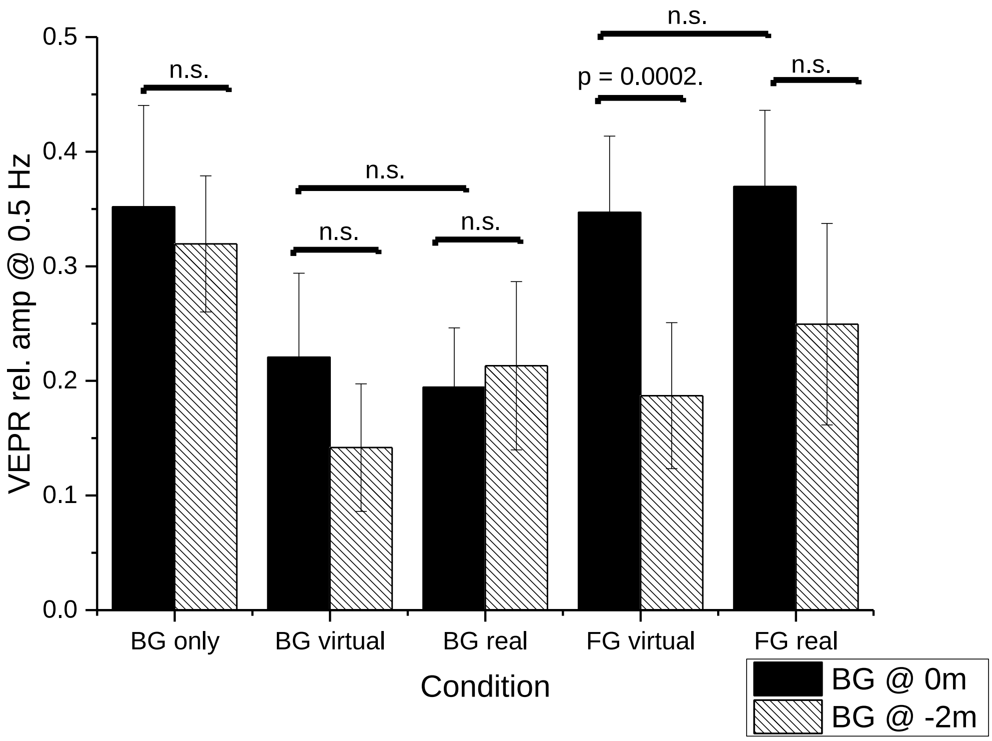


Figure 7. Relative amplitude of postural responses to visual stimulus in motion between 0m experimental block (black bars) and -2m experimental block (grey bars). The labels identify the fixation point and whether the foreground object was a virtual or real.

The results of planned comparisons are summarised on Table 3. Our experimental results show significant VEPRs between 0m and -2m blocks during the virtual foreground fixation only (p = 0.0002, d = 2.45). The VEPR component of postural sway was abolished in background fixation whilst real and virtual foreground objects were present in the environment. No significant differences in VEPRs were recorded in the background only condition and during the foreground fixation whilst the real object was present.

Table 3

Lateral sway energy at 0.5Hz during visual stimulation in 0m condition and in -2m conditions

Condition	% Energy at 0.5 Hz Visual Motion at 0m Mean (SD)	% Energy at 0.5 Hz Visual Motion at - 2m Mean (SD)	Difference	Effect size (Cohen's d)
BG only	35.19% (8.8%)	31.95% (5.9%)	$t_{(14)} = 0.8605$, $p = 0.4$	0.43
BG virtual	22.07% (7.3%)	14.18% (5.7%)	$t_{(14)} = 2.427$, $p = 0.02$	1.21
BG real	19.45% (5.2%)	21.22% (7.3%)	$t_{(14)} = 0.5887$, $p = 0.565$	0.29
FG virtual	34.70% (6.7%)	18.71% (6.4%)	$t_{(14)} = 4.9075$, $p = 0.0002$	2.45
FG real	36.95% (6.7%)	24.95% (8.8%)	$t_{(14)} = 3.077$, $p = 0.008$	1.53

Table 3: Lateral Sway energy at 0.5Hz during visual stimulation and rest and comparison statistics for the target and control conditions in Experiment 1, background at 0m. Comparisons where there are significant differences between the visual stimulation and rest condition (Bonferroni corrected alpha levels of $p < 0.007$, (0.05/7)) are highlighted in bold. The effect sizes (Cohen's d) indicate the magnitude of the observed differences in means.

3.3.1.7 Summary of the Experiment 1

All five experimental conditions were designed to replicate conditions from Meyer et al. (2013). The previous data were matched in the 0m experimental block; the VEPR component of postural sway was significant during background fixation when no foreground object was present (BG only) and during foreground fixation when real and virtual objects were presented and fixated (FG virtual and FG real). It was expected that in the -2m experimental block, the VEPR component of the postural sway will be decreased. The data show that significant VEPRs were observed only during the background fixation with no foreground object present (BG only). The VEPRs were unaffected in all the other conditions regardless of fixation point or foreground objects. In conclusion, the increased postural sway caused by visual motion was observed when visual motion was presented directly on the screen plane as opposed to being rendered as 2 meters behind the projection screen.

3.3.2 Experiment 2. The explicit assessment of the positional stability of environmental anchors and its effects on postural sway.

3.3.2.1 The main aim

The main aim of Experiment 2 was to investigate whether the perceived stability of real environmental anchors presented as foreground objects will modulate the VEPR component of postural sway. To allow comparison between studies the same design was chosen as the study in Experiment 1. The convergent plane was projected slightly behind the projection screen whilst the spatial position of the real objects, teapot and helium balloon, was 2 meters in front of the participants. The main reason for choosing helium balloon was in its different physical properties as compared to teapot i.e. teapot will not move freely, it is stable reference point, helium balloon can sway in the air, unstable reference point. Three conditions were completed in two experimental blocks (teapot and helium balloon). The VEPR component of postural sway between visual motion and pause was recorded, computed and compared as in Experiment 1.

All participants completed all conditions in both experimental blocks. Due to the virtual reality set up the order of the conditions could not be changed so all participants completed the conditions in the same order. The order of the experimental blocks was counterbalanced across participants.

3.3.2.2 Experimental conditions

1. BG only	- fixate a target on the background (BG), no foreground object present.
2. BG real	- fixate a target on the background (BG), a real foreground object (teapot or helium balloon) is present.
3. FG real	- fixate a target on the real foreground (FG), either a teapot or helium balloon.

3.3.2.3 Experimental results

The behavioural performance was recorded as a number of correct clicker responses whilst fixation target transiently disappeared. All participants correctly identified 99.57% of events (SD = 6.13) from which 99.48% (SD = 6.89) in the teapot block and 99.65% (SD = 5.42) in the balloon block. This means that the attention or fixation position as a cofounding variable can be rejected.

3.3.2.4 Experiment 2a. Postural responses to visual stimulation with teapot as a real foreground object

The postural sway was investigated in lateral direction in both experimental blocks. In the first part of the experiment a real teapot was presented as a real foreground object in two out of three conditions. Mean and standard deviations were computed after the bootstrapping procedure and are therefore based on the population averages rather than a sample. A repeated measures ANOVA with fixed factors 'condition' (BG only, BG real, FG real) and 'visual stimulus' (ON, OFF) and a ‘participant’ as a random factor showed a significant main effect of visual stimulus (F(1, 108) = 10.5, p = 0.0016, $\eta^2 = 0.53$) and a significant main effect of condition (F (4,108) = 5.06. p = 0.009, $\eta^2 = 0.51$). There was no significant interaction between condition and visual stimulus. The analysis of the VEPR component of postural sway at 0.5Hz in comparison to the same frequency component when no motion is present (pause) with adjusted Bonferroni correction (0.05/4) was conducted. Figure 8 shows that postural responses to the real foreground object differ in the point of fixation.

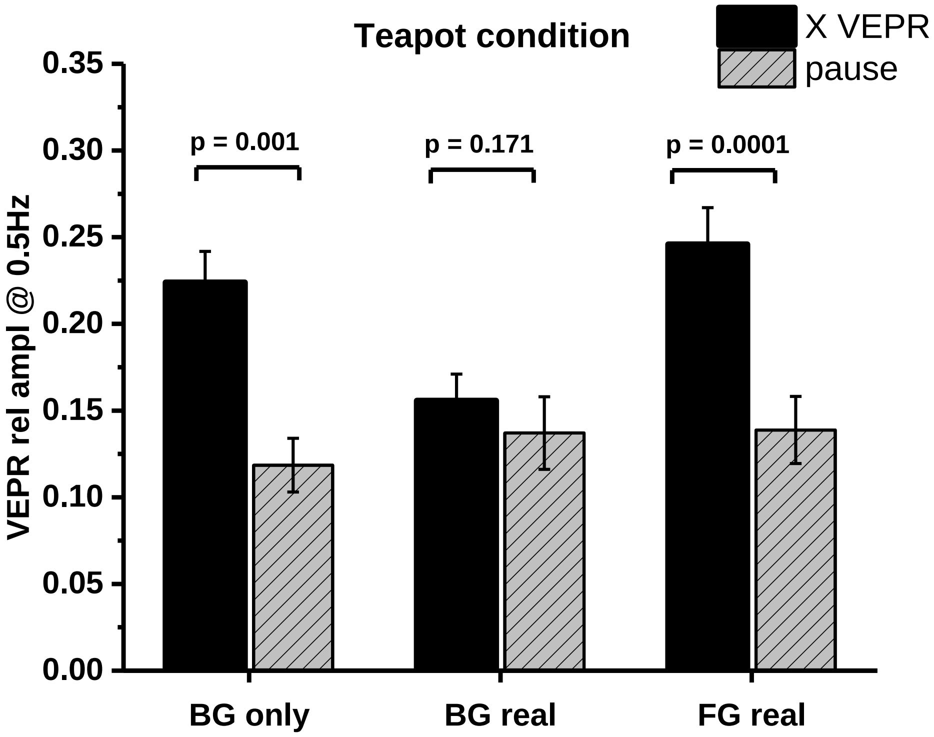


Figure 8. Relative amplitude of postural responses to visual stimulus in motion (black bars) and in pause (grey bars), when a teapot was presented in a foreground. The labels identify whether the fixation was on foreground or background.

The summary of experimental results can be seen in Table 4. When the background was fixated without any foreground object present in the VR scene, significant VEPRs were recorded between visual motion and pause (BG only). When the fixation on the background was maintained and the teapot was presented as a foreground object, the significant VEPRs were abolished. When the fixation moved to the teapot on the foreground the significant VEPR component of postural sway was recorded. These experimental results confirm the findings from the Experiment 1 (0m condition) and previously reported data with physical static reference points (Meyer et al., 2013; Bronstein & Buckwell, 1997). The VEPRs were recorded in the anti-phase to the BG only condition.

Table 4

Lateral sway energy at 0.5Hz during visual stimulation in teapot condition

Condition	% Energy at 0.5 Hz <u>Visual Motion</u> Mean (SD)	% Energy at 0.5 Hz <u>Rest</u> Mean (SD)	Difference	Effect size Cohen's d
BG only	0.22 (0.06)	0.12 (0.06)	$t_{(24)} = 3.9715$, $p = 0.0001$	1.66
BG real	0.15 (0.02)	0.13 (0.04)	$t_{(24)} = 5.1708$, $p = 0.1711$	0.63
FG real	0.25 (0.03)	0.13 (0.04)	$t_{(24)} = 5.0502$ $p = 0.0001$	3.39

Table 4: Lateral sway energy at 0.5Hz during visual stimulation and pause. A Bonferroni correction was applied (0.05/3). Comparisons where there are significant differences between the visual stimulation and rest condition. The significant VEPRs were recorded in BG only and FG real conditions. No significant VEPRs were recorded for BG real condition.

3.3.2.5 Experiment 2b. Postural responses to visual motion with a helium balloon as a foreground object.

In the second experimental block, a helium balloon was presented as the real foreground object instead of the teapot. During the task, a helium balloon was attached to the stand by a piece of string and secured to the top of the stand via double-sided tape, but the participants were not aware of this. Extra free floating balloons were positioned by the entrance to encourage participants to think that the balloon can indeed swing in the air, however they were not visible from where the participants were standing during the actual experimental task i.e. no balloons were visible in the participants' peripheral viewing during the task. Means and standard deviations were computed after the bootstrapping procedure and are therefore based on the population averages rather than a sample. A repeated measures ANOVA with fixed factors 'condition' (BG only, BG real, FG real) and 'visual stimulus' (ON, OFF) and a 'participant' as a random factor showed a significant main effect of visual stimulus (F(1, 108) = 12.93, p = 0.005, $\eta^2 = 0.48$) and a significant main effect of condition (F (4,108) = 4.59, p = 0.0018, $\eta^2 = 0.52$). No significant interaction between condition and visual stimulus was recorded. The analysis of VEPRs at 0.5Hz in comparison to the same frequency component when no motion is present (pause) was conducted with adjusted Bonferroni correction (0.5/4). Figure 9 shows that postural responses to the real foreground object differ in the point of fixation.

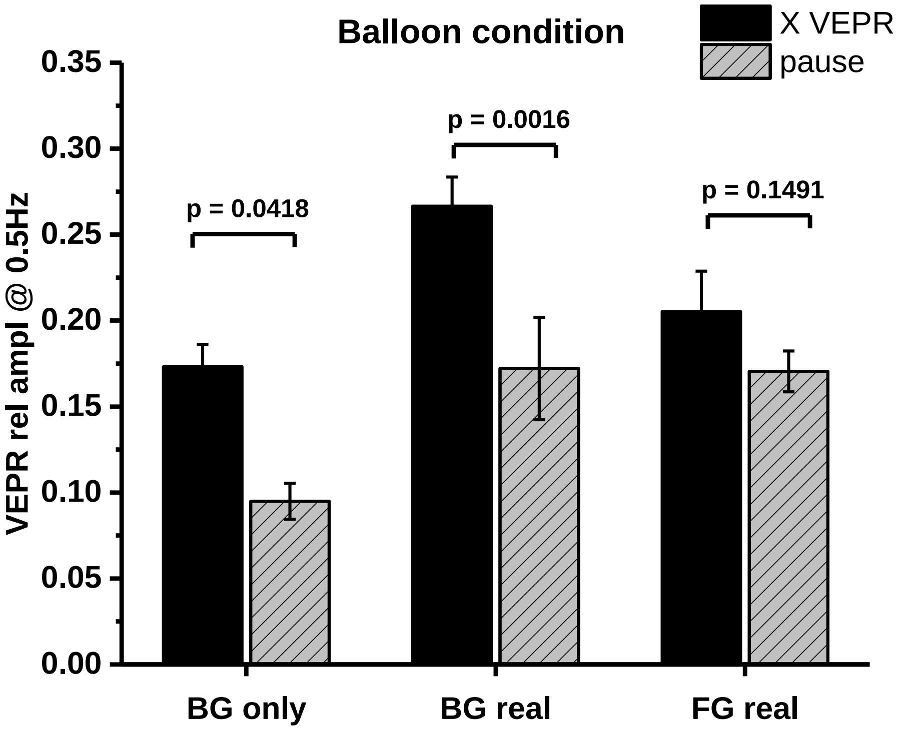


Figure 9. Relative amplitude of postural responses to visual stimulus in motion (black bars) and in pause (grey bars) when helium balloon was presented as a foreground object. The labels identify whether the fixation was on foreground or background.

The summary of experimental results can be seen in Table 5. Significant VEPRs were recorded during the background fixation whilst no foreground object was present. Significant VEPRs were also recorded when the helium balloon was presented as a foreground object during the background fixation. When fixation shifted to the helium balloon the significant VEPR component of postural sway was abolished. The amplitude of postural sway for the helium balloon as a foreground object did not differ significantly in regards to the point of fixation (p = 0.134).

Table 5

Lateral sway energy at 0.5Hz during visual stimulation in balloon condition

Condition	% Energy at 0.5 Hz Visual Motion Mean (SD)	% Energy at 0.5 Hz Rest Mean (SD)	Difference	Effect size Cohen's d
BG only	0.17 (0.09)	0.09 (0.03)	$t_{(24)}$ = 2.1534, p = 0.004	1.19
BG real	0.26 (0.04)	0.17 (0.04)	$t_{(24)}$ = 3.5607, p = 0.0016	2.25
FG real	0.21 (0.05)	0.17 (0.05)	$t_{(24)}$ = 4.078 p = 0.1491	0.79

Lateral sway energy at 0.5Hz during visual stimulation and pause. A Bonferroni correction was applied (0.05/4). Comparisons where there are significant differences between the visual stimulation and rest condition. Significant VEPRs were observed in the BG only condition and in the BG real condition. Significant VEPRs were not recorded in the BG real condition.

3.3.2.6 Comparison of the magnitude of postural sway between the teapot and helium balloon presented as foreground objects

Planned comparisons were conducted via t-test with Bonferroni correction (0.05/2) to directly compare postural sway responses between the teapot and the helium balloon experimental blocks. It was predicted that the helium balloon would be disregarded as a stable reference point due to its unstable physical properties and therefore induce larger VEPRs. Significantly larger VEPRs for the balloon condition as compared to the teapot condition were observed. Figure 10 shows that postural responses differ significantly between the two foreground objects.

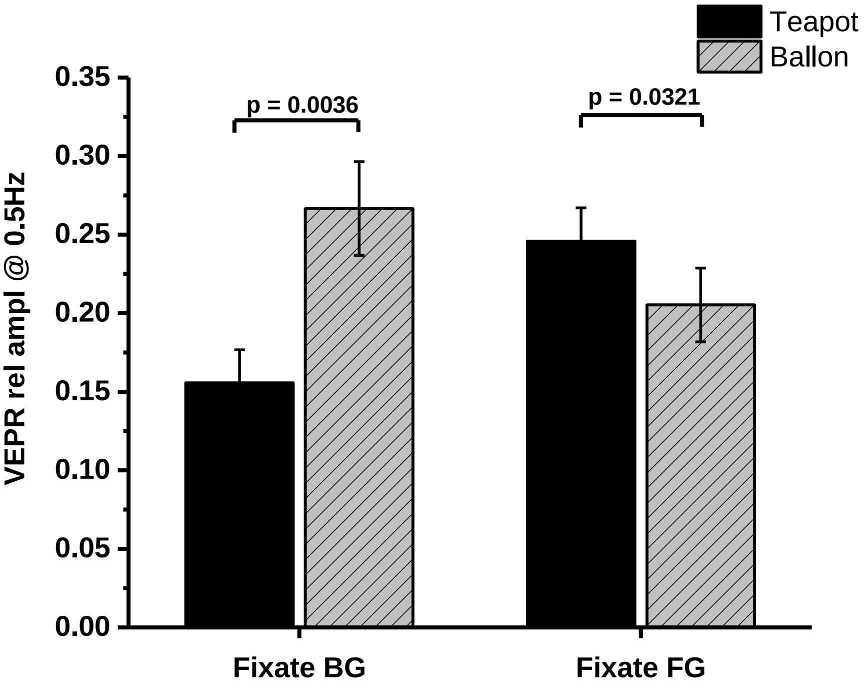


Figure 10. Relative amplitude of postural responses to visual stimulus in motion when the teapot (black bars) or helium balloon (grey bars) was presented in the foreground.

The results of planned comparisons can be seen in Table 6. During the background fixation, the helium balloon elicited significantly more postural sway as compared to the teapot when they were presented as foreground objects, but not attended. When the fixation was shifted to the foreground, larger VEPRs were recorded for the teapot condition, however after the correction was applied the difference between the magnitude of postural responses was not significant.

Table 6

Differences between teapot and helium balloon. Adjusted Bonferroni significance p < 0.025.

Condition	Teapot Mean (SD)	Balloon Mean (SD)	Difference	Effect size Cohen's d
BG real	0.15 % (0.02 %) 21.22% (7.3%)	0.26 % (0.04 %) 12.3% (4.9%)	$t_{(18)} = 3.22$, p = 0.0036	3.47
FG real	0.25 (0.03) 24.95% (8.8%)	0.21 (0.05) 14.8% (3.6%)	t(18) = 3.36, p = 0.035	0.97

Table 6: Lateral sway energy at 0.5Hz during visual stimulation in the teapot and helium balloon conditions. Comparisons where there are significant differences between the visual stimulation and rest condition. A Bonferroni correction was applied (0.05/2). Significant VEPRs were observed in the BG real condition but not in the FG real condition.

3.3.2.7 Summary of Experiment 2

The three experimental conditions (BG only, BG real, and FG real) in Experiment 2 were designed to replicate data from the Meyer et al. (2013) study. The observed postural responses matched previously reported data in the teapot condition in Experiment 2a as well as those in the Meyer et al. (2013) study. It was predicted that the helium balloon would be disregarded as a stable reference point due to its unstable physical properties. The recorded data of postural sway when the helium balloon was presented showed significant VEPRs when the helium balloon was presented but not attended. The VEPRs were abolished when the helium balloon was attended as a foreground fixation point. The direct comparison between the magnitudes of postural sway showed that the helium balloon elicited significantly higher VEPRs as compared to the teapot condition when they were presented as foreground objects but not fixated. The results show that the VEPRs can be influenced by the explicit assessments in physical stability of foreground objects.

3.4 Discussion

The primary aim of this study was to investigate how the explicit configuration of real and virtual environments modulates automatic visually evoked postural responses. Two experiments were designed and conducted to replicate and extend previous work showing differential responses for real and virtual foreground objects (Meyer et al., 2013). The results showed that VEPRs were modulated in lateral direction by the visual disparity cues provided by 3D stereoscopic displays. The evaluation of the positional stability of environmental anchors also modulates the magnitude of postural responses recorded during visual motion stimulation. The experiments confirmed previous research findings that VEPRs are modulated by the fixation point and by the stimulus configuration. The results also show the presence of significantly different behaviours for real and virtual foreground objects (Bronstein and Buckwell, 1997; Meyer et al., 2013).

Overall, participants successfully adjusted their postural responses to visual motion, which is in line with previous research (Bronstein and Buckwell, 1997; Meyer et al., 2013). The findings are also consistent with the postural instability theory where postural control was defined as coordinated stabilisation of all body segments. This theory proposes that when a person encounters a destabilising environment he will try to regain and maintain postural control (Ricco and Stofferegen, 1991). In all cases, participants adjusted their postural responses to visual stimulus; however the way the visual disparity cues were presented modulated this effect. The explanation for the observed effects is provided in the section bellow.

3.4.1 Visual disparity cues and vergence – accommodation conflict

The perception of depth in 3D environments is achieved by providing appropriate visual disparity cues that render virtual objects simultaneously to both eyes with slightly different views. Our brain then combines this information in order to allow us to see a single percept. Empirical research has shown that binocular disparity can affect perceptions of self-motion illusion, also called vection (Palmisano, 2002). In Experiment 1 the position of visual disparity cues was manipulated; background visual motion and real and virtual foreground objects were presented in different distances when the position of the convergent plane was projected as being a) right on the projection screen or b) 2m behind the projection screen. When visual motion was rendered as being directly on the projection screen, VEPRs were modulated in a similar way that was previously suggested (Meyer et al., 2013). When the fixation point was on the background, VEPRs were unaffected in the background only condition (BG only) and in the presence of the real foreground object (BGR), but abolished when the virtual object was present (BGV). By shifting the fixation point on the foreground objects, a significant postural sway was recorded when real and virtual objects were present and fixated (FGV and FGR). Some of these findings are in contrast with Meyer et al., (2013) study who found that VEPRs were abolished during foreground virtual fixation. I present a couple of explanations for the contrasting findings.

One of the explanations for reduced postural sway during background fixation may lay in the overall improvements of the VR set up. Since Meyer et al., (2013) study, the set-up of the virtual environment has been improved by adding extra motion tracking cameras, which improved the rendering of the scene as well as the improvement of the overall processing speed of the VR machine. It is possible that as a result of this technological improvement, virtual objects were perceived as more real and therefore used as stable environmental anchors in the same way as a real teapot during background fixation. When comparing the magnitude of postural sway between real and virtual foreground fixation no significant differences were observed, which suggests that both objects were treated in the same way. As postural responses are performed unconsciously, these findings support the idea that a high quality virtual reality system with improved motion tracking and graphical image rendering can increase the overall levels of immersion and presence during the visual motion stimulation.

The absence of postural sway could also be explained by the amount of information in the periphery, which is the information within the visual field that is not attended to, it is outside of the centre of gaze (Field, 2013). It has been suggested that a relative availability and saliency of peripheral visual cues is likely to modulate VEPRs in condition where a static foreground object is present (Meyer et al., 2013). In Experiment 1a participants were standing 4m from the projection screen. Even though the side blinkers were used to reduce the information from the peripheral cues, such as the edges of the projection screen, the visibility of the top and bottom edges of the projection screen could not have been prevented. It is therefore possible that during the background fixation the top and bottom edges of the screen, together with additional motion cameras that were positioned above the screen and were producing red light, could potentially reduce the levels of immersion, as the fixation point was relatively close to the top edge of the screen.

Meyer et al. (2013) suggested that visually evoked postural responses play a big part in the illusionary self-motion (vection) and subsequently in the perceived sense of presence that can be experienced in virtual environments (Meyer et al., 2013). Therefore the result from this study could be explained with reference to the perceived sense of presence. When the background was projected on the convergence plane a significant postural responses were recorded. This suggests that participants experienced sufficient level of presence that can be seen in the adjustment of their postural responses. When the background was rendered as being 2 meters behind the projection screen it evoked no significant postural responses. Given that the -2meter condition presented a consistent vergence-accommodation conflict, the visual disparity cues and the concomitant vergence-accommodation disparity is therefore a likely explanation for these findings.

As mentioned before, stereoscopic 3D displays, in contrast to the real world, generate artificial environments where coupling between accommodation and vergence is lost due to the differences in distance between virtual objects and display screens. While our eyes try to converge to see the 3D virtual objects, the displays require the eyes to focus (accommodate) on the screen where the image is the sharpest. Hence, the accommodation distance stays the same but vergence distance is varied as the eyes try to focus on virtual objects in different disparities (Lambooij, et al., 2010). Similarly, the distances for vergence and accommodation processes were also different during Experiment 1b where no postural responses were observed during background fixation. It may be possible that participants experienced visual fatigue and discomfort caused by the vergence–accommodation conflict, which in turn had negative effects on participant's perception and awareness of virtual environments in such way that they might start feeling uncomfortable and uneasy. The experienced discomfort lead to a break of immersion and subsequently to lower levels of presence as participants became more aware of their surroundings, and thus no postural sway was observed.

The visual attention perspective that brings to light the suppressive mechanisms of visual information processing, as recorded in Experiment 1b, can also explain the magnitude of postural responses. The critical role of visual processing is to highlight useful and relevant information while suppressing redundant and less informative signals, also called spatial suppression. This selective process is used to suppress background motion in order to free up resources for effective attention to foreground objects (Tadin, 2015). In this study, background motion might have been disregarded, as all attentional resources were used on a transiently disappearing red or blue dot on foreground objects. Furthermore, the overall improvement of the VR set up increased the quality of the rendering of virtual object and thus both, real and virtual teapots were treated as stable reference points whilst all background motion was disregarded. As a result no postural sway was observed. The findings from Experiment 1 supported the prediction that postural responses are modulated by the perceived fidelity of visual information presented in 3D environments.

3.4.2 Positional stability of real environmental anchors

The presence of real foreground objects can modulate the amount of postural sway recorded in response to visual motion (Meyer et al., 2013). The unattended real foreground object reduces the amount of postural sway however when fixation is shifted on the real foreground object the significant difference in postural sway is recorded in the opposite direction to the visual motion. The experimental data of postural responses recorded in Experiment 2a are in line with previous research (Meyer et al., 2013; Bronstein and Buckwell, 1997) that showed that postural sway is driven by the fixation point and visual motion. However, it was predicted that when the real object that has inherently different physical properties to the other object (helium balloon instead of teapot) is presented on the foreground, it will be disregarded as a stable environmental anchor. When the real foreground object was a helium balloon, whose physical properties are inheritably different to those of a teapot, i.e. a balloon can sway in the air and therefore might not be a static reference point, the prediction was confirmed during the background fixation. However, VEPRs were abolished when the helium balloon was fixated as a foreground object. This suggests that the helium balloon was perceived as a stable reference point around which the postural responses were anchored. Direct comparison of the magnitude of postural sway between teapot and helium balloon showed that VEPRs were significantly increased during background fixation and similar during the foreground fixation. This suggests that the unattended real teapot was used as an environmental anchor to stabilise the postural sway whilst the helium balloon was disregarded as a stable reference point in the same condition. These findings suggests that while postural control is perhaps accomplished without awareness it clearly draws on very high level of knowledge about the environment.

The experimental framework for evaluation of fidelity in VR that includes the assessment of postural responses has been proposed previously (Meyer et al., 2013). The findings from this study support this notion and further confirm that postural responses can be affected by explicit assessment of real and virtual environment. In particular, it was shown that visual disparity cues that are provided via 3D stereoscopic displays in virtual environments are important factors in modulating the levels of immersion and presence. The real and virtual environmental anchors presented in the environment were shown to modulate the levels of presence, however the findings suggests that this is also dependent on the evaluation of positional stability of the real environmental anchors. Therefore, the effective and appropriate presentation of visual information in VR should be considered as a main factor that can enhanced and modulate the levels of immersion. Additionally, the adverse effects of visual fatigue and discomfort that might be present during the illusionary self-motion could also be modulated by the appropriate visual information.

3.4.3 Limitations

The experimental design used in this study has been adopted from the previous study (Meyer et al., 2013), which is one the main limitation of this study. As the VR set up between the two studies was improved it made it difficult to directly compare the conditions to the previous findings. As no comparison was conducted, the findings from this study might be seen as less reliable. Another limitation of this study was the fact that the order of the conditions could not be randomised. This was mainly due to the VR set up and predesigned Matlab script from which the simulation run. However this meant that the results of the study could have been influenced by the order effect (Zeelenberg & Pecher, 2015). The sample size for both of the studies can be seen as another limitation of this study, however this was caused by the technological constrains as new VR set was being installed a thus only limited number of participants was recruited. This study also did not include any subjective opinions from the user, which is another limitation of this study. The future research investigating the postural responses to visual motion should therefore include larger sample size where subjective opinion will be collected and conditions of the visual stimulation and fixation will be counterbalanced to minimised order effects (Zeelenberg & Pecher, 2015).

Chapter 4

The effects of multisensory cuing on fidelity assessments in virtual environment

Chapter overview

This chapter covers findings from a study that was designed to investigate the role of augmented sensory feedback on performance and user perception of the virtual environment. The introduction will cover the role of the multisensory cues used in mediated environments, following with a background research into motion tracking available during VR interaction. The experimental aims and result will be described subsequently. The first part will cover the results concerning the effects of augmented sensory cues (visual, audio, tactile) on performance and subjective feeling of presence. The second part will focus on the effects of motion tracking accuracy on performance, presence and sickness. The chapter will conclude with a discussion of all findings.

4.1 Introduction

4.1.1 Multisensory cues

For a user to interact in a VR environment, virtual reality systems are configured to provide information in sensory modalities, particularly vision and audio. Other modalities, especially proprioception, touch or vestibular signals are sometimes more difficult to simulate; this is usually due to technological and financial constrains (Batfield & Hendrix, 1995). While it is, for example, quite easy to provide realistic visual and auditory cues representing a power-tool, it is not so easy to simulate haptic cues, like the weight of the tool or the torque that would be felt when the tool engages with the work piece.

It has been shown that the additional information represented in these cues may be expected to enhance performance and user experience (van Erp & van Veen 2004; Ramsamy, Haffegee, Jamieson & Jamieson, 2003; Burke et al., 2006). However, the fidelity of the VR environment has also been linked to task performance (Schuchard & Bowman, 2007) so presenting arbitrary cues, especially if they are not natural cues, may have a negative impact on the fidelity of the environment as well as on task performance. For detailed explanation about the effect of multisensory cues on performance please refer to Chapter 2.

4.1.2 Motion tracking

As well as the presentation of sensory cues, a high quality motion tracking is necessary for effective interaction as it records the position and the orientation of real objects in physical space and allows spatial consistency between the real and virtual objects (Lubeck et al., 2015). To achieve a high tracking accuracy, a three-dimensional high-quality motion tracking system needs to be implemented; accurate motion tracking is necessary to create and update the viewpoint of the user and allow natural interaction in VR (Greuter & Roberts, 2014). Motion tracking generally relies on expensive hardware; however recent availability of inexpensive commercial systems for the tracking of body movement enabled many researchers to effectively record positional movements of subjects during experimental studies. One of these systems is Microsoft Kinect, which was originally developed for the enhanced interaction in computer game environments. Due to its low cost range sensors it has become a widely used alternative to high range of motion tracking systems used in many applications (Khoshelham & Elbernik, 2012, Shin et al., 2013). However, when used for human tracking few drawbacks, such as error measurements, positional accuracy, limited tracking area and a resolution quality have been identified. Thus there is a need to investigate the effectiveness of these systems in order to reduce negative effects on user perception and performance. For more detailed information about motion tracking please refer back to Chapter 2.

The principal question driving this research was to investigate whether cues that are more difficult to generate in VR (weight and torque) can be replaced by simple arbitrary information-bearing cues in different modalities (tactile, visual and audio). At the same time, the effectiveness of the motion tracking system is investigated by introduction of positional inaccuracy that simulates the discrepancies of low-cost motion tracking systems reported previously (Khoshelham & Elbernik, 2012, Shin et al., 2013). The main aim was to investigate whether the presentation of additional augmented sensory cues, which carried task-relevant information, will modulate overall task performance and user acceptance of the VR environment.

4.1.3 Experimental approach and summary of the experimental aims

In this study, augmented sensory cues are provided during the virtual assembly task performed in the VR environment. A planar projection screen was chosen as the virtual reality platform, which enabled participants to move freely when interacting within the virtual environment and enabled the introduction of additional augmented sensory feedback. The task was to perform a tyre change on a virtual racing car whilst holding a physical tool that was also rendered within the virtual environment. Additionally, the VR environment was modulated in two experimental blocks: in one block accurate motion tracking was provided; in the other block a positional inaccuracy was introduced into the motion tracking to simulate positional discrepancies that are present in low cost motion tracking systems. The aim was to investigate whether this manipulation will influence user's behaviour and their acceptance of the VR environment. The overall time to complete the task was recorded as an objective measures. The users' acceptance of the virtual environment was assessed by a series of subjective questionnaires, in which participants rated their feelings of presence and discomfort. Throughout the study, objective and subjective measures were compared to examine any relationship between the measured variables. It was predicted that the presence of augmented sensory cues (visual, audio, tactile) would not interfere with task performance or users acceptability of the VR. It was also predicted that the multimodal condition would be perceived as the most favourable feedback, followed by other modes of feedback presentation in objective and subjective measures. In regards to the effects of tracking latency, it was predicted that positional inaccuracy would have negative effects on objective and subjective measures of user performance.

4.2 Methods

4.2.1 Participants

A power analysis (Faul et al., 2009) prior to the experiment was performed to determine a sample size required for this study and showed that for 0.05 level of significance, the effect size of 0.5 and power (1 – β err prob) of 0.7 total sample size required was 16. For this study, we recruited 17 participants via opportunity sampling; there were 12 males and 5 females, aged between 18 and 48 (M = 26.7, SD = 12.4). All participants signed a consent form and reported normal or corrected-to normal vision and hearing (Appendix A).

4.2.2 Virtual reality set up

The experiment was conducted at the Virtual Engineering Centre (VEC) facility located in the Science and Technologies Facilities Council (STFC) in Daresbury. The task was to interact within the virtual environment by holding a real pneumatic tool (impact wrench) and perform a wheel change on a virtual racing car whilst augmented cues, presented in the form of visual, tactile and audio sensory feedback provided additional task-relevant information.

The virtual reality set up consists of a planar display screen 6m in length and 2.1m in height behind which are two active stereo projectors that create 3390 x 1200 resolution images at a rate of 120Hz. 3D stereo images are produced by an NVIDIA Quadro K6000 GPU. Participants wear wireless LCD shutter glasses that are synchronized with the projectors to provide stereoscopic images. 16 high-spec infrared cameras (VICON Bonita B10, 250fps capture speed, motion resolution of 0.5mm of translation and 0.5 degrees of rotation in a 4m x 4m volume using 9mm markers) are used to track object motion in the VR environment. Position data, computed using VICON Tracker software, is broadcast in real-time across the internal network using a VRPN protocol at a rate of 200Hz and used to update the virtual environment. The following objects are tracked in order to provide the required interaction within the virtual immersive environment: LCD shutter glasses (for head tracking and POV adjustment), haptic gloves on subject’s hands (to enable tracking of subject’s virtual hands) and the impact wrench (PLC Prestige 1/2”, 1.3kg, 15.2cm long), the tool used to remove the bolts form the wheel.

4.2.3 Sensory stimuli

The wheel change simulation is designed in 3DVia system and runs at a constant speed of 75fps. The virtual simulation consists of a virtual racing car that is positioned on a stand in the middle of the screen. The virtual scene also contains two stands that are positioned 68cm in x direction, 14cm in y direction and 8cm in z direction from the centre point of the wheel on the virtual car. The stand on the right side of the car holds the spare tyre and the stand on the left side of the car is free for participants to put the car tyre on (Figure 11). In order to pick the wheel up participants had to stand directly in front of the wheel, thus every participant had to make postural adjustments to get in contact with the wheel. A faithful digital mock-up of the impact wrench is used to interact with the bolts.

Tactile feedback is provided by “tactile gloves” which have vibration motors attached to the palm of the hand in each of the VICON hand tracking kits (Figure 12 C). The motor (8.8mm in diameter, 25mm long) is actuated by PWM drivers receiving information on collision detection, level of vibration, etc. by a device wirelessly connected to the CPU running the immersive scenario. The vibration occurs with variable frequency, ranging from 15Hz (when participants hold the bolt) to 250Hz (the strongest vibration experienced when bolt is in). For example, the subject can feel an intermediate level of vibration when screwing a bolt out or back in place (around 85Hz), which increases to the maximum level (250Hz) as soon as the bolt is completely screwed in, or reduced to zero when it is completely removed. In this way we mimic the intensity of vibrations generated by the impact wrench, or torque, when performing the real task Figure 11.

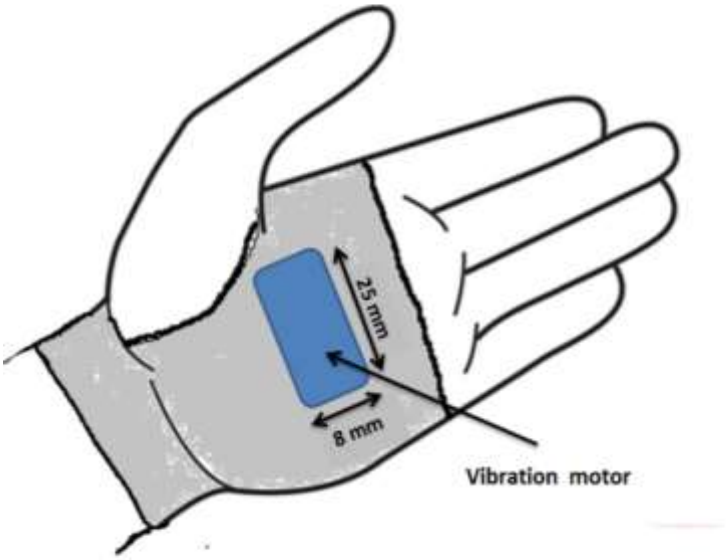


Figure 11. Diagram of a position of the vibration motor. Participants were able to feel vibration on the palm of their hand in various frequencies.

Audio feedback was presented via SONY headphones (RF811RK Wireless Headphones) with the frequency range between 20-20000Hz. When audio cues were enabled they provided sound of 10,000Hz. When audio cues were disabled, a white noise (800Hz) was played to mask the sound generated by the vibration motors in the haptic gloves.

Figure 12. Virtual tools and set up. Apparatus used in the experiment: 3D shutter glasses (A), impact wrench (B) and haptic gloves (C). Picture D shows the position of the wheel when the task is completed (first wheel on the stand, second wheel on the racing car).

4.2.4 Task procedure

The experiment was conducted in a room with closed doors, dim lights and no other distractions. Participants wore 3D shutter glasses, vibration gloves and headphones that played either audio cues (when they were on) or a continuous white noise to mask any vibration noise from the haptic gloves (Figure 11 C). The task was to change the wheel on the virtual racing car in the 3D environment as fast as possible. The augmented sensory cues were presented as unimodal, bimodal and multimodal feedback in a counterbalanced order. Every participant started with two practice trials.

The time started when the participants got in contact with the physical tool (impact wrench). First, they had to unscrew five bolts from the wheel on the virtual racing car. Then they had to put down the impact wrench, pick up the wheel and then put it on the stand located on the right side off the racing car (68cm, Figure 11, D). After this, they had to go and grab another wheel from the stand on the left side of the racing car, attach it on the racing car, grab hold of the impact wrench and screw the bolts back in. The overall recording stopped when the participants placed the tool back on the table, which was located on the right side, approximately 1.5m from the projection display. Between each condition, the subjective ratings of the sense of presence and involvement were recorded on a short PQ questionnaire (Appendix D). After every combination of cues was completed, participants were debriefed and thanked for their participation**.**

4.2.5 Objective and subjective performance measures

In this study we used overall task performance as an objective measure and the sets of questionnaire as a subjective measure.

Objective measures: Participants were instructed to perform the task as fast as they could. The mean times for each condition were used in the statistical analysis. As an incentive a leader board was displayed within the virtual scene where the participants' fastest times were shown and updated after each participant had completed the task.

Subjective measures: In order to investigate participants' acceptance of the virtual scenario, participants were required to fill in a questionnaire between each sensory condition. The selected dimension for the presence measure was a continuous scale. Before the task participant filled in ITQ and SSQ to provide baseline measures.

After performing the task in each sensory condition they were required to fill in a short questionnaire, which consisted of 7 items that were based on a 4-factor model of presence and adapted from PQ questionnaire (Witmer et al., 2005). After each experimental block participants had a break to rest and to fill in a long version of PQ and SSQ scale. When participants completed the second experimental bloc, during which they again completed a short version of PQ, they were asked to fill in another set of PQ and SSQ. For the analysis, the objective measures of performance and subjective ratings obtained from participants were compared to examine any possible relationship between the measured variables.

4.2.6 Multisensory feedback

A projection based VR system, where visual information is always present, was used in this study. Before each study a calibration process was performed with a calibration wand to assure accurate tracking of the VR tools. The augmented task-relevant cues (visual, tactile, audio) provided additional information to compensate for the cues that are not readily available in VR, such as weight and torque. The augmented visual cues consisted of: the bolts turning yellow when in contact with the tool and red when the bolts are completely in or out; the wheel turning yellow when in contact and red when in the right position; the virtual hands of the participant turning yellow when in contact with virtual objects (Figure 13). The tactile cues were presented as vibration sensation when: the tool was in contact with the bolt and a more intense vibration when the bolt was completely in or out; when the virtual hands were in contact with the wheel (carrying the wheel from the car to the stand and from the stand on to the car). The audio cues included a drilling noise when in contact with the bolt and a ‘snap’ sound when the wheel was placed on the stand and on the car.

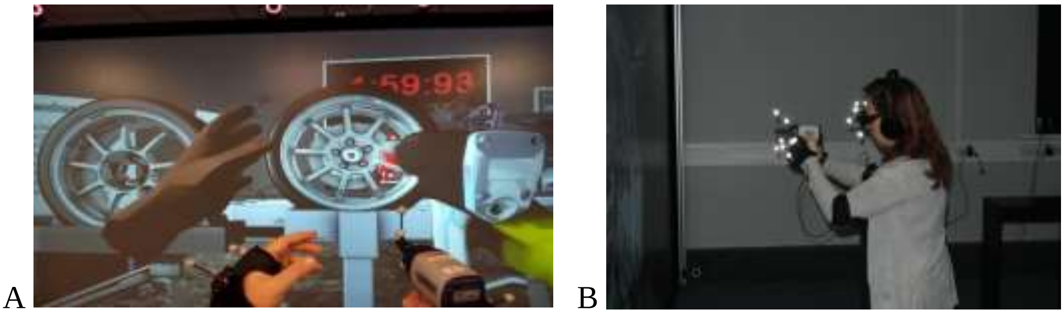


Figure 13. View during the task. The simulation view from a participants' perspective: (A) Participants were able to see their virtual hands and virtual too; (B) Participants wore 3D shutter glasses, headphones, vibration gloves and were holding an impact wrench whilst performing the task.

4.2.7 Experimental design

This study adopted a repeated measures within-subject design. There were eight possible combinations of sensory feedback provided: audio (A), visual (V), tactile (T), audio-visual (AV), audio-tactile (AT), tactile-visual (TV) and audio-visual-tactile (AVT). The condition where no cues were presented was also included (NONE). The 2 x 2 x 2 factorial design (audio x visual x tactile cues) was used where all possible combinations of sensory conditions were presented in quasi-random order. During the task, the order of the conditions was designed in such a way that each condition appeared equally often during the task. Additionally, the virtual reality environment was modulated in two experimental blocks. In one the experimental blocks, accurate motion tracking was provided during the virtual simulation. In the other experimental block, the positional inaccuracy was introduced via the motion tracking system to simulate positional discrepancies that are often present with low-cost motion tracking systems (Khoshelham & Elbernik, 2012). The manipulation corresponded to the 0.5Hz back-and-forth movement of the whole VR scene during the entire experimental block whilst participants were performing the experimental task. This manipulation was unknown to the participants.

The analysis of variance (ANOVA) was performed on mean completion times to investigate the effects of each sensory modality. Visual, audio and tactile cues were set as binary factors that were either present or absent. The Mauchly test of sphericity was applied and when significant, Greenhouse-Geisser corrections were adopted. Partial eta squared (η^2) is reported for effect sizes. Paired sample t-tests, with correction for multiple comparisons, were conducted to investigate whether the presence of sensory modality affected participant's behaviour and performance. When the direction of the relationship was predicted, one tailed test results were reported. The significance level for all statistical tests was set at 0.05. For paired sample t-tests, the Cohen's d (Cohen, 1992) was chosen as a measure of effect size, which is calculated as the difference between two means divided by pooled standard deviation. The equation is as follows:

(Equation 1)

$$d = \frac{(M1 - M2)}{SD\ pooled}$$

where *d* stands for Cohen's effect size, *M1* and *M2* are the means from two groups and *SD pooled* is calculated as follows:

(Equation 2)

$$SD\ pooled = \sqrt{SD1 + SD2}$$

where *SD pooled* is the standard pooled variance, *SD1* and *SD2* are the standard deviations from two groups. The accepted suggestions for the magnitude of effect sizes are: 0.2 = small effect, 0.3 = medium effect and 0.5 = large effect (Cohen, 1992). When the data did not follow the general assumptions for the parametric tests, non-parametric tests were used. This was during the SSQ data analysis as the collected data were ordinal thus did not have parametric distribution. The Freidman Anova was used for the analysis and Kendall's Coefficient of Concordance (W) was used as a measure of effect size (Field, 2009). When a-priory assumptions were made, planned comparisons using a Wilcoxon sign rank test were performed. The calculations of the effect sizes for non-parametric tests were performed by dividing Z by the square root of *N*, where *N* is the number of observations not the number of participants. The equation used for the calculation of effect sizes was as follows:

(Equation 3)

$$r = \frac{Z}{\sqrt{Nx + Ny}}$$

where r is an estimated effect size, Z is the test statistics, *Nx* and *Ny* refers to the number of observations in the study (Pallant, 2013).

4.3 Results

The results section is separated into two parts. The first part of the result section will focus on measures obtained during presentation of the additional augmented cues. The second part focuses on the effects of manipulating the positional tracking accuracy.

4.3.1 Part 1. The effects of augmented sensory cues on performance

4.3.1.1 Objective measures

In this study, the effects of additional cues in the conventional VR set up on task performance and user acceptability of VR were investigated. In order to assess user performance, mean task completion times were recorded in each condition and used for the analysis. Overall, it took participants an average of 50.3 seconds (SE = 1.9) to complete the virtual wheel change task. The mean task completion times in each sensory cue condition are shown on Figure 14.

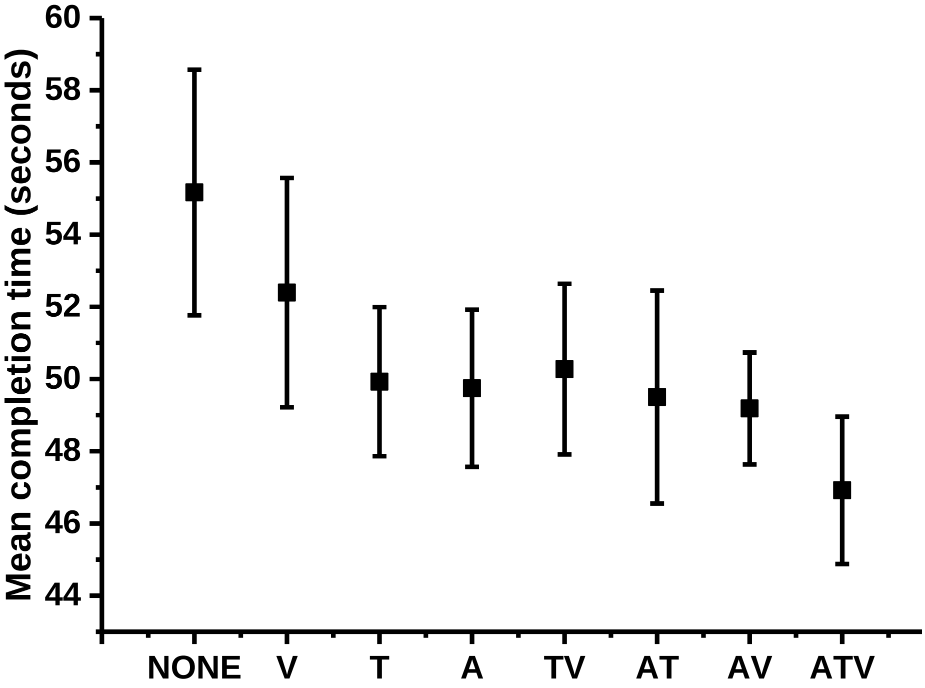


Figure 14. Objective data. Means and standard errors in each condition for mean completion times. Labels show what additional cues were presented during the simulation: NONE no additional cues, V visual, A audio, T tactile, TV tactile-visual, AT audio-tactile, AV audio-visual, ATV audio-tactile-visual. Participants were asked to perform the task as fast as they can; therefore shorter completion times indicate better performance. The data indicate that as the amount of sensory cues in the simulation increased, the task was completed faster. The order of conditions that facilitated performance the most is: ATV (M = 46.9, SE = 2.03); AV (M = 49.1, SE = 1.54); AT (M = 49.5, SE = 2.9); TV (M = 50.2, SE = 2.36); A (M = 49.7, SE = 2.17); T (M = 49.9, SE = 2.06); V (M = 52.3, SE = 3.17) and NONE (M = 55.1, SE = 3.41).

Repeated measures analysis of variance (ANOVA) within subject design was used in this study where visual, audio and tactile cues were set as binary factors that were either present or absent. The ANOVA revealed that there were significant main effects of audio (F(1,16) = 5.4, p = 0.034, η^2 = 0.25) and tactile modality (F(1,16) = 5.013, p = 0.04, η^2 = 0.24), but not visual modality (F(1,16) = 0.714, p = 0.411, η2 =

0.04). No significant interaction was observed between any other combinations of sensory modalities. Planned comparisons using Sidak adjustment were performed on data that were separated into groups where each binary condition was either present or absent. The analysis revealed that the task was performed much faster when audio cues were present (M = 48.8, SE = 1.61; t(16) = -2.324, p = 0.015, d = 0.36, one tailed) as when they were not provided (M = 51.9, SE = 2.42). Significant differences were also observed for tactile modality; the mean time to complete the task was significantly faster when the tactile cues were present (M = 49.6, SE = 1.85; t(16) = -2.044, p = 0.025, d = 0.29, one tailed) as oppose to when they were absent (M = 51.6, SE = 2.21). No significant difference were observed on mean completion times for visual modality when it was either present (M = 49.6 SE = 1.86; t(16) = -1.027, p = 0.15, d = 0.17, one tailed) or absent (M =51.1, SE = 2.23) (Figure 15).

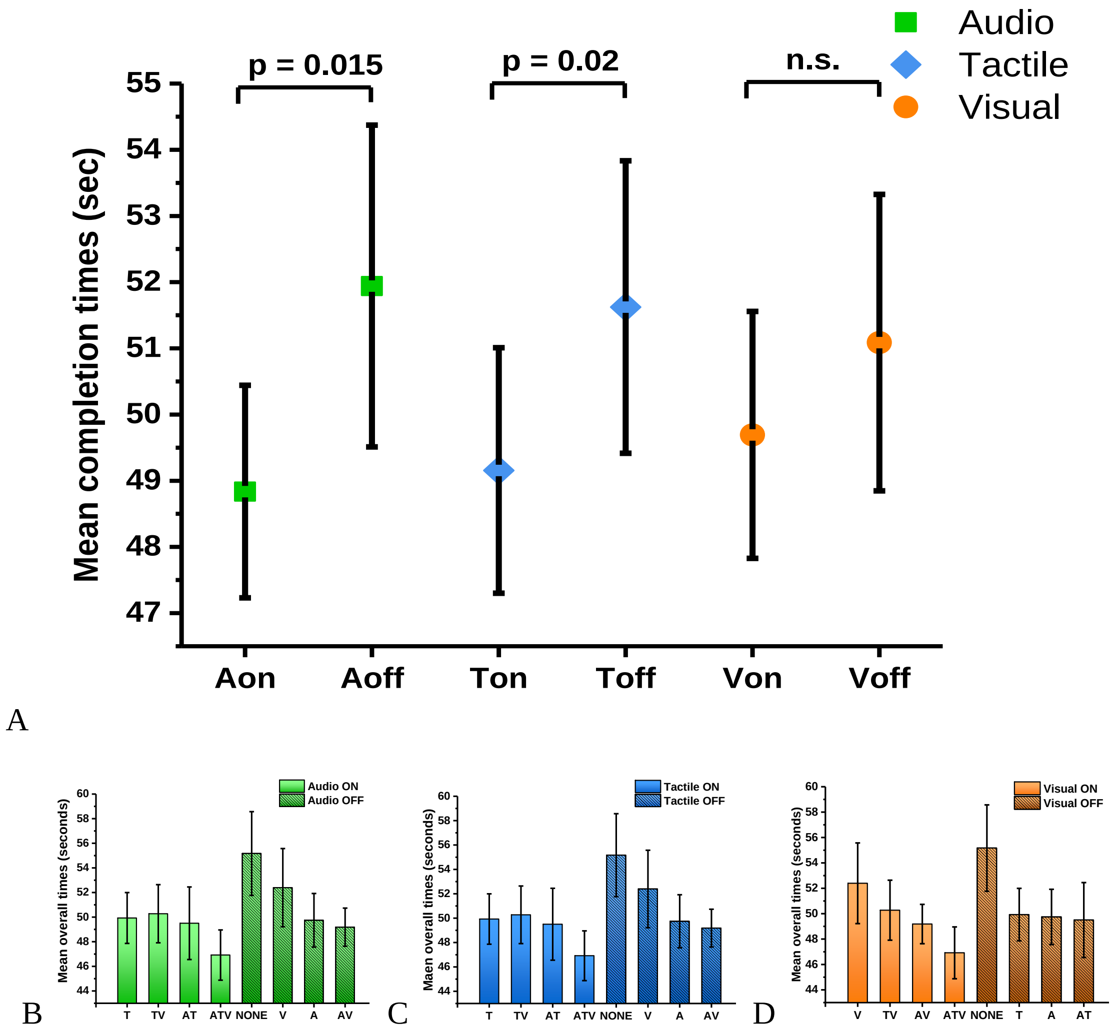


Figure 15. Main effects of performance for each sensory cue. Means and standard errors for (A) audio, (B) tactile and (C) visual cues. To visualise the data easily, all sensory cues were separated into groups when each of the cues were present or absent. Graphs B, C and D represent the same data but are displayed to show individual contribution to the main effect. Paired sample t-tests with Sidak correction revealed a significant effect of audio (p=0.015) and tactile cues (p=0.025). No significant differences were found for visual cues (p = 0.15).

To further investigate whether the modes of feedback presentation had an effect on mean completion times, the data were grouped into four modes of feedback and the mean across the groups was used for the analysis (Figure 16). As it was predicted that multimodal and bimodal sensory feedback will facilitate performance more that no additional feedback condition all pairwise tests are reported one tailed.

Planned comparisons using Sidak correction ($p < 0.01$) revealed that the task was performed significantly faster with multimodal feedback (M = 46.9, SE = 2.03) as oppose to no feedback (M = 55.1, SE = 3.4; t(16) = -2.884 , p = 0.005, d = 0.7), and there was a trend towards a significant difference with unimodal feedback (M = 50.6, SE = 2.25, t (16) = -2.099, p = 0.025, d = 0.5). Significant difference was also observed between no feedback (M = 55.1, SE = 3.4) and bimodal feedback (M = 50.2, SE = 1.9; t(16) = - 2.515, p = 0.015, d = 0.57). No significant difference was observed between bimodal (M = 50.2, SE = 1.9) and unimodal feedback (M = 50.6, SE = 2.25; p = 0.25, d = 0.16) which suggests that both modes influenced the task performance in a similar manner.

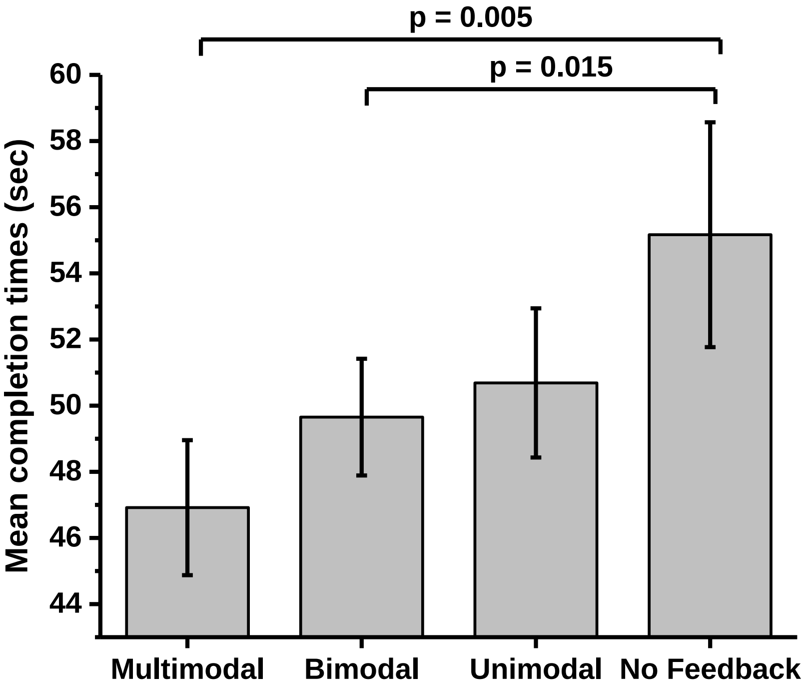


Figure 16. Main effects for modes of feedback for performance. Means and standard errors for overall completion times for each mode of feedback: Labels show what cues were presented during the simulation: Multimodal (audio-tactile-visual), Bimodal (audio-tactile, AV audio-visual, TV tactile-visual), Unimodal (A audio, V visual, T tactile), No feedback (NONE no additional feedback). Data for bimodal (AV, AT, TV) and unimodal (A, T, V) category were group together and the mean score is used for the analysis. Significant difference ($p < 0.01$)

4.3.1.2 Subjective measures

In order to assess user's acceptability of the virtual environment, participants were asked to rate their feeling of presence and involvement on a short questionnaire adapted from PQ (Witmer & Singer, 1998) after each experimental condition. The reliability of the rating scale was assessed by calculating the Cronbach coefficient -alpha (Nam et al., 2008). The standardised alpha of rating scales showed acceptable reliability (sense of presence = 0.97). Overall mean rating score was 6.65 (SE = 1.74), which indicated that the VR environment induced sufficient levels of presence and involvement. The mean subjective ratings in each condition can be seen on Figure 17.

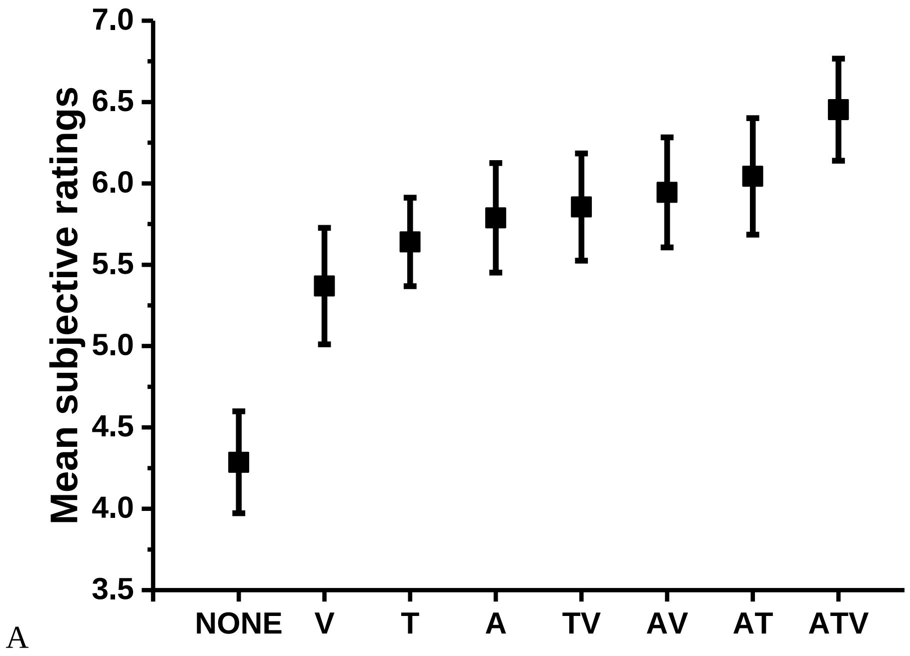


Figure 17. Subjective rating scores. Means and standard errors of presence ratings in all experimental conditions. Labels show what additional cues were presented during the simulation: NONE no additional cues, V visual, A audio, T tactile, TV tactile-visual, AT audio-tactile, AV audio-visual, ATV audio-tactile-visual. Participants were asked to rate on a continuous scale from 0-10 how strongly they agree or disagree with statements. All statements were formulated in a positive manner, thus higher scores indicate an enhanced sense of presence (A). The order of conditions that were perceived as the most compelling is as follows: ATV (M = 6.55, SE = 0.33); AT (M = 6.23, SE = 0.38); AV (M = 6.04, SE = 0.34); TV (M = 5.98, SE = 0.34); A (M = 5.92, SE = 0.35); T (M = 5.69, SE = 0.28); V (M = 5.47, SE = 0.35) and NONE (M = 4.3, SE = 0.31).

A 2 x 2 x 2 factorial, repeated measures analysis of variance (ANOVA) within subject design was used where visual, audio and tactile sensory cues were set as binary factors that were either present or absent. The analysis with Greenhouse-Greisser correction determined that the presence of sensory modalities significantly influenced subjective ratings: there was a main effect of audio modality (F(1,16) = 33.380, $p < 0.001$, $\eta^2 = 0.67$), the main effect of tactile modality (F(1,16) = 21.203, $p < 0.001$, $\eta^2 = 0.57$) and the main effect of visual modality (F(1,16) = 17.828, p = 0.001, $\eta^2 = 0.52$). Significant interactions between all sensory modalities were also recorded: audio and tactile (F(1,16) = 5.525, p = 0.032, $\eta^2 = 0.25$); audio and visual (F(1,16) = 6.470, p = 0.022, $\eta^2 = 0.28$); tactile and visual (F(1,16) = 8.490, p = 0.01, $\eta^2 = 0.34$) and audio, tactile and visual (F(1,16) = 15.199, p = 0.001, $\eta^2 = 0.48$). Planned comparisons using Sidak adjustment revealed that participants experienced a higher sense of presence and involvement when audio cues were present (M = 6.18, SE = 0.34) as oppose to absent (M = 5.38, SE = 0.29; t (16) = 5.784, $p < 0.001$, d = 1.39). The same effect was seen for the presence of tactile cues (M = 6.09, SE = 0.31) as oppose to their absence (M = 5.48, SE = 0.32; t(16) = 4.389, $p < 0.001$, d = 1.06), and as well as the presence of visual cues (M = 6.03, SE = 0.33) as oppose their absence (M = 5.53, SE = 0.29; t(16) = 4.282, p = 0.001 , d = 1.03) (Figure 18).

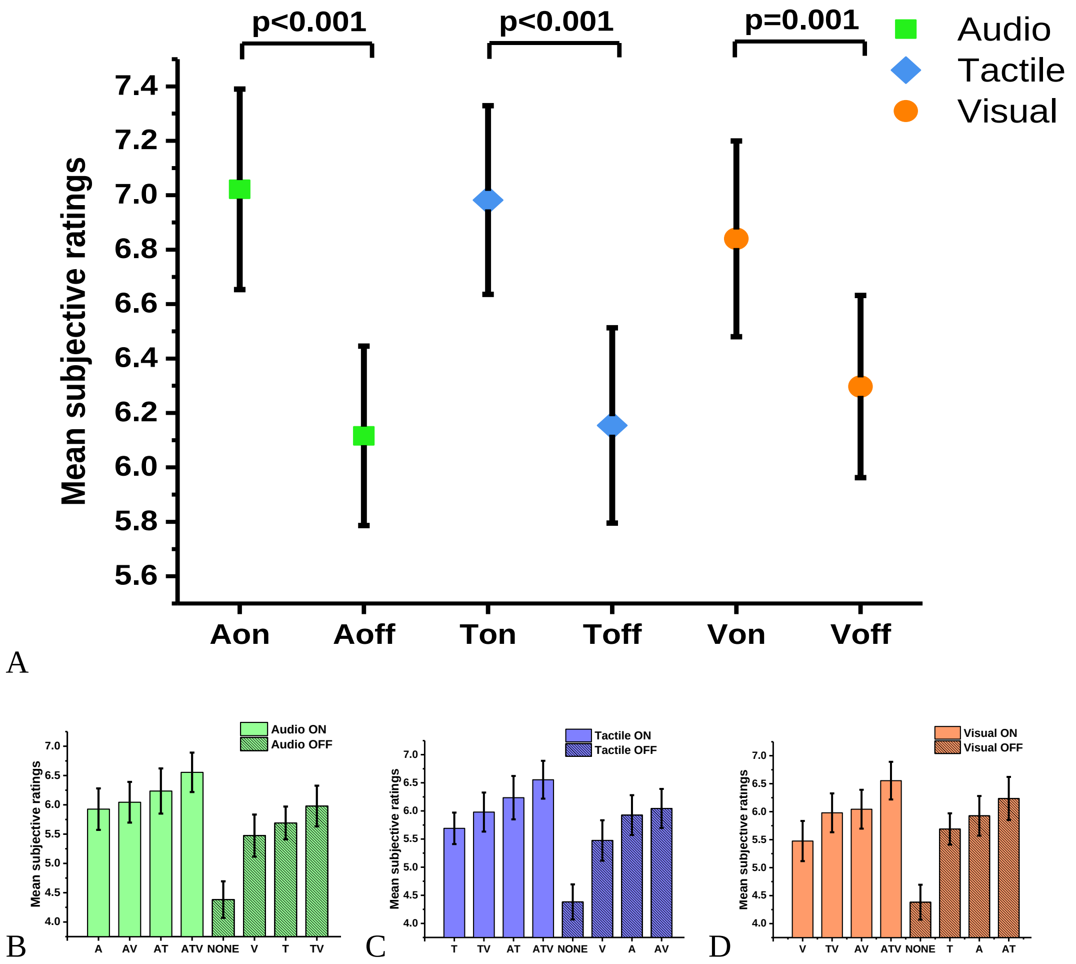


Figure 18. Subjective ratings on a short PQ scales for each sensory cue. (A) Means and standard errors for subjective PQ ratings of presence for audio, tactile and visual cues. Panel B, C and D represent the same data as panel A to aid visualisation and to show an individual contribution to the main effects.

The effects of different modes of feedback on subjective ratings were also examined. Planned comparisons with Sidak adjustment ($p < 0.01$) revealed that the subjective ratings of presence and involvement were significantly higher when multimodal feedback was presented ($M = 7.51$, $SE = 0.37$) as compared to bimodal ($M = 6.9$, $SE = 0.37$, $t(16) = -3.787$, $p = 0.002$, $d = 0.86$), unimodal ($M = 6.47$, $SE = 0.35$; $t(16) = -4.967$, $p < 0.001$, $d = 1.25$) and no feedback ($M = 4.89$, $SE = 0.36$; $t(16) = 7.210$, $p < 0.001$, $d = 1.79$). Subjective ratings also differed significantly between no feedback ($M = 4.89$, $SE = 0.36$) and unimodal feedback conditions ($M = 6.47$, SE

= 0.35; t(16) = 6.708, p < 0.001, d = 1.51) and between no feedback and bimodal feedback conditions (M = 6.9, SE = 0.37, t(16) = 5.807, p < 0.001, d = 1.44). No significant difference was observed between bimodal (M = 6.9, SE = 0.37) and unimodal feedback conditions (M = 6.47, SE = 0.35; t(16) = -2.654, p = 0.02, d = 0.55) (Figure 19).

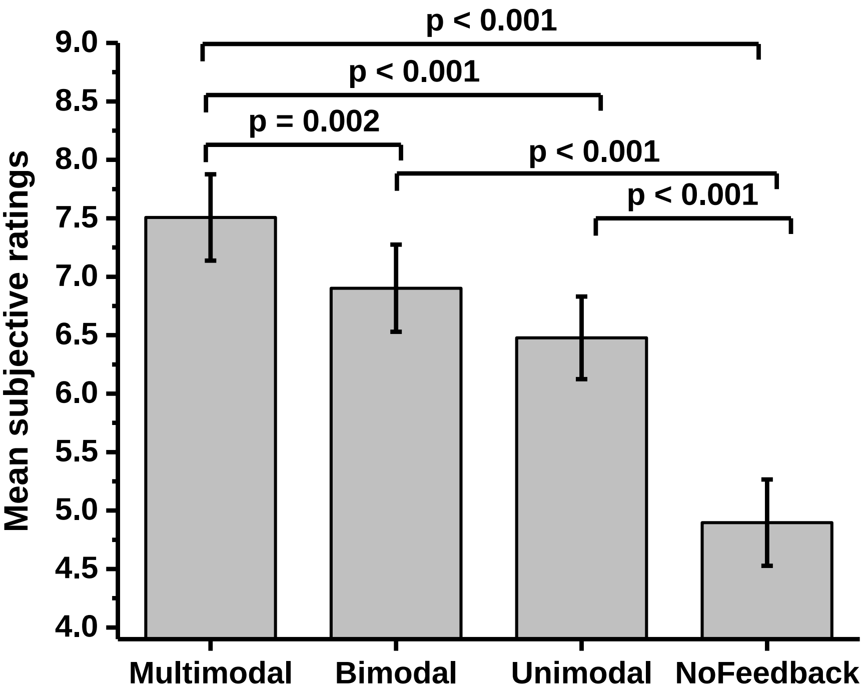


Figure 19. Main effects of modes of sensory feedback for subjective ratings. Means and standard errors of subjective ratings for each mode of feedback. Labels show what cues were presented during the simulation: Multimodal (audio-tactile-visual), Bimodal (audio-tactile, AV audio-visual, TV tactile-visual), Unimodal (A audio, V visual, T tactile), No feedback (NONE no additional feedback). Data for bimodal (AV, AT, TV) and unimodal (A, T, V) category were group together and the mean score is used for the analysis.

4.3.1.3 Correlations

The main interest of this study was to evaluate the contribution of sensory cues to overall performance and user behaviour. To investigate the relationship between the measured variables a correlation analysis was performed. However, it has been noted that this type of analysis can be misinterpreted due to the amount of variability within and across the participants (Bland & Altman, 1995). In fully factorial design, where the same subject is tested across a range of conditions, the variability of measurements between subjects is usually greater than the variability between measurements on the same subject (Bland & Altman, 1995). To account for the idiosyncratic mapping between participants on the objective and subjective measures, the mean data we used as a basis for our correlation analysis (Table 7).

Table 6

Means and SD for objective and subjective measures.

Condition	Objective data (seconds) Task performance Mean (SD)	Subjective data Short PQ questionnaire Mean (SD)
A	49.7 (8.9)	5.7 (1.4)
V	52.4 (13.1)	5.3 (1.5)
T	49.9 (8.5)	5.6 (1.2)
AV	49.1 (6.4)	5.9 (1.4)
AT	49.5 (12.1)	6 (1.5)
TV	50.2 (9.7)	5.8 (1.4)
ATV	46.9 (8.4)	6.4 (1.2)
NONE	55.1 (14.2)	4.2 (1.3)

Table 6. Objective data were recorded in milliseconds and then transformed into seconds. Lower times indicate better performance. Subjective measures were collected on a continuous scale from 0-10. Higher ratings indicate an increased sense of presence.

Firstly, correlation analysis was performed on the individual scores for each participant in objective and subjective measures. The results showed that the data did not correlate together mainly because they included variability between participants. Secondly, the correlational analysis was performed on the raw data and showed a significant relationship (r = -0.240, p = 0.005). However, as this data included an idiosyncratic variability within participants, we decided to pool the means across each sensory condition, as the main aim was to show the effects of sensory modalities on subjective and objective data rather than look at the variability across participants.

The correlation revealed a significant negative relationship between objective and subjective data on the level of sensory condition ($r = -0.978$, $p < 0.001$) (Figure 20). The negative correlation implies that shorter completion times are better in terms of objective performance. This means that when participants experienced an enhanced sense of presence whilst performing the task, the overall task completion times were shorter. This finding suggests that an enhanced sense of presence and involvement can facilitate task performance.

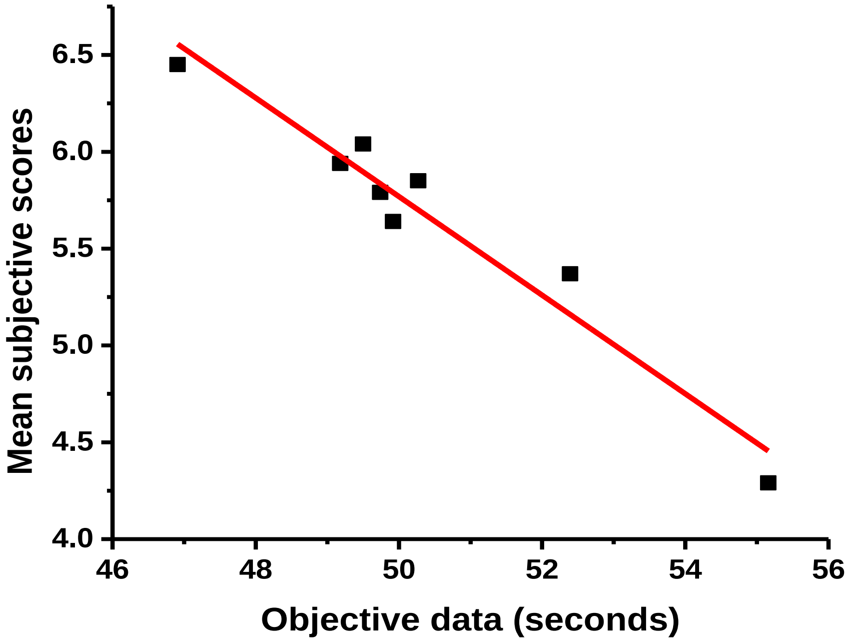


Figure 20. Correlation analysis. A significant negative relationship ($r = -0.978$, $p < 0.001$) was found between the subjective ratings of presence and the objective measures of overall completion time. Subjective measures were recorded on the 0-10 continuous scale; objective measures were recorded in milliseconds and then converted into seconds.

4.3.2 Part 2. The effects of positional inaccuracy of motion tracking system

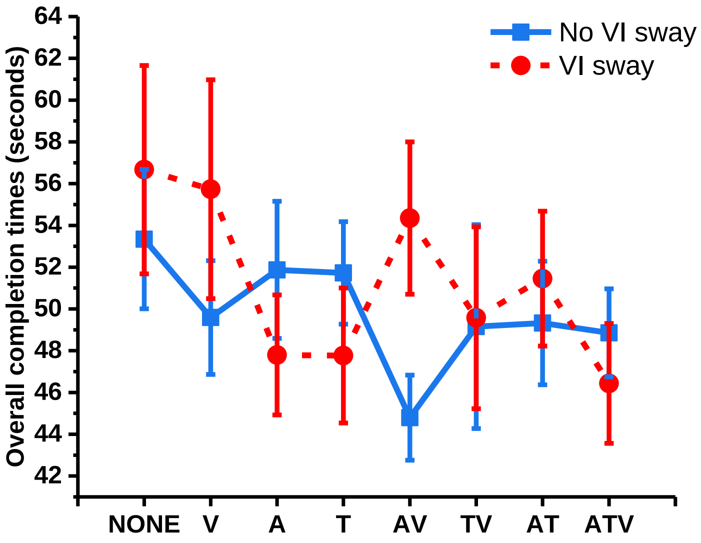
4.3.2.1 Methodology

To record positional changes in the virtual environment and allow comparison between accurate and inaccurate motion tracking, a VICON tracking system was used. This tracking system has been found very effective in measuring motion in VR, mainly due to its accuracy and possibility to attach markers to any device for calibration, allowing easy addition of extra input devices for minimal extra cost (Murray et al., 2003). In this study two experimental blocks were compared: one where accurate motion tracking was provided, no virtually induced (VI) sway was present; and one where virtually induced (VI) sway was introduced during the task. The VI sway consisted of 0.5Hz of continuous oscillations corresponding to 2cm backward and forward movements of the virtual scene during the task in one of the experimental blocks. As it was previously shown that the movements along the z-axis cause significant postural sway (Meyer et al., 2013), the reason behind this manipulation was to simulate positional discrepancies that are similar to discrepancy measurements previously reported with low cost tracking systems (Khoshelham and Elbernik, 2012). The main aim was to investigate whether the introduced VI sway that simulates the positional inaccuracy of the VR environment (sway movement) will have an effect on overall user experience during a virtual assembly task. It was predicted that the VI sway will have adverse effects on performance and user acceptability of the virtual environment. Furthermore, I wanted to examine whether the presence of the augmented sensory cues could alleviate the negative effects of inaccurate motion tracking.

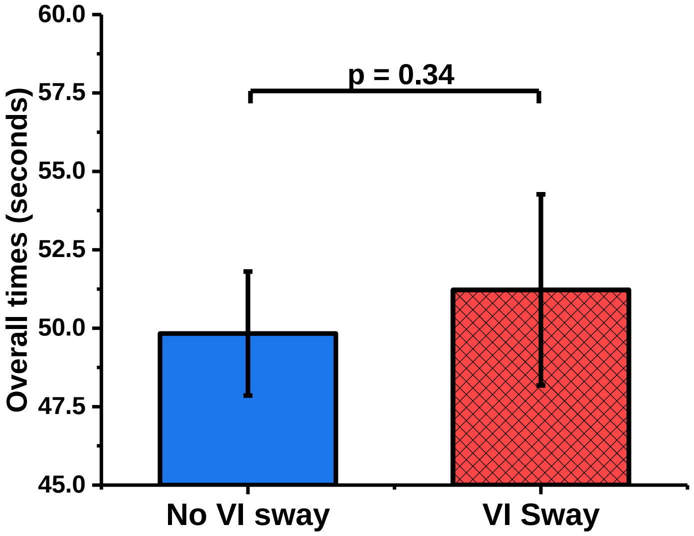
4.3.2.2 Task performance in each experimental block

Due to a technical fault, one participant did not perform the task in the modulated experimental block and therefore only 16 participants were included in the statistical analysis. All participants completed the task 16 times in counterbalanced order during two experimental blocks: one block with accurate motion tracking (No VI sway) and one with a tracking latency of 0.5Hz (VI sway). Overall, it took participants about 49.8 seconds (SE = 0.91) to complete the task in No VI sway condition and about 51.2 seconds (SE = 1.39) to complete the task in VI sway condition. During the accurate motion tracking block participants performed fastest in audio-visual condition (AV, M = 44.7, SE = 2.03); when scene sway was present the fastest time was recorded in multisensory condition (ATV, M = 46.4, SE = 2.41). All of the data were analysed using a repeated measures analysis of variance (ANOVA) (Figure 21). As it was predicted that the performance will decrease with inaccurate tracking, all pairwise tests are reported one tailed.

The results revealed a significant interaction between VR environment and sensory condition (F(7,105) = 1.798, p = 0.04, η^2 = 0.46) and there was a trend towards significant main effect of condition ((F(3,52) = 2.345, p = 0.07, η^2 = 0.18). No significant effect of environment was observed (F(3,52) = 1.045, p = 0.32, η^2 = 0.07). A Sidak post hoc test did not show any significant differences between modulated and normal VR environment. This finding suggests that although participants performed the task slightly faster when accurate motion tracking was provided, the difference in the mean completion times recorded in the condition with a tracking latency was not significantly different.



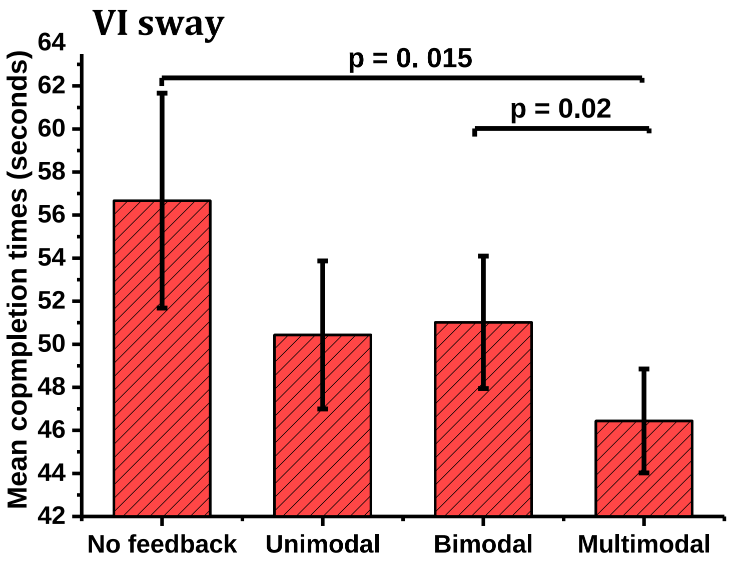
A



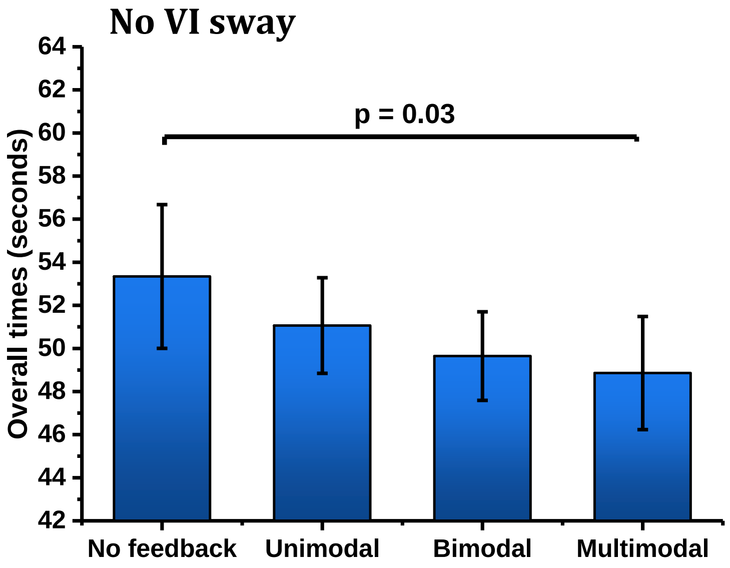
B

Figure 21. Overall mean times in each sensory condition (A) VI = virtually induced sway, and for each experimental block (B). Error bars represent standard error of the mean.

When examining which mode of feedback presentation was the most beneficial to overall task performance similar patterns in both experimental blocks could be seen (Figure 22). The multimodal condition facilitated the performance the most in both experimental blocks (No VI sway, M = 48.8, SD = 10.4; VI sway, M = 46.4, SE = 2.03), followed by bimodal, uni-modal and no additional feedback condition. Slight differences were observed in the Sway experimental block between unimodal and bimodal mode of feedback, where mean completion time for unimodal condition (M = 50.4, SE = 3.34) was faster than the mean completion time for bimodal condition (M = 51.1, SE = 3.07). In both experimental blocks the longest time to complete the task was recorded in the condition were no sensory feedback was provided (No VI sway, M = 53.3, SE = 3.33; VI sway, M = 56.6, SE = 4.91). To examine the beneficial effects of modes of feedback during the two experimental blocks planned comparisons with Sidak adjustment ($p < 0.02$) were conducted and revealed that, in the VI sway block, performance improved significantly when additional sensory cues were presented. The results showed that participants performed significantly worse with no feedback (M = 56.6, SE = 4.9) as compared to multimodal feedback (M = 46.4, SE = 2.41; t(15) = 2.296, p = 0.015, d = 2.6) and bimodal feedback (M = 51.01, SE = 3.07; t(15) = 2.211, p = 0.02, d = 1.36) and there was a trend towards significance with unimodal feedback (M = 50.4, SE = 3.43; p = 0.03, d = 1.46) (Figure 22, A). In No VI sway block no significant difference between the modes of feedback were recorded but there was a trend towards significance between multimodal feedback (M = 48.8, SE = 2.62) and no feedback (M = 53.3, SE = 3.33; p = 0.03, d = 1.5) (Figure 22, B).



A

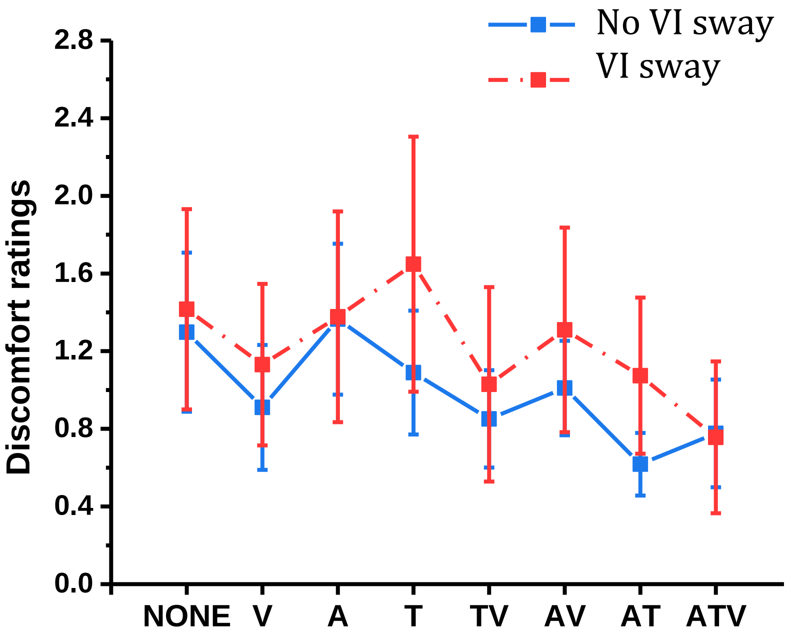


B

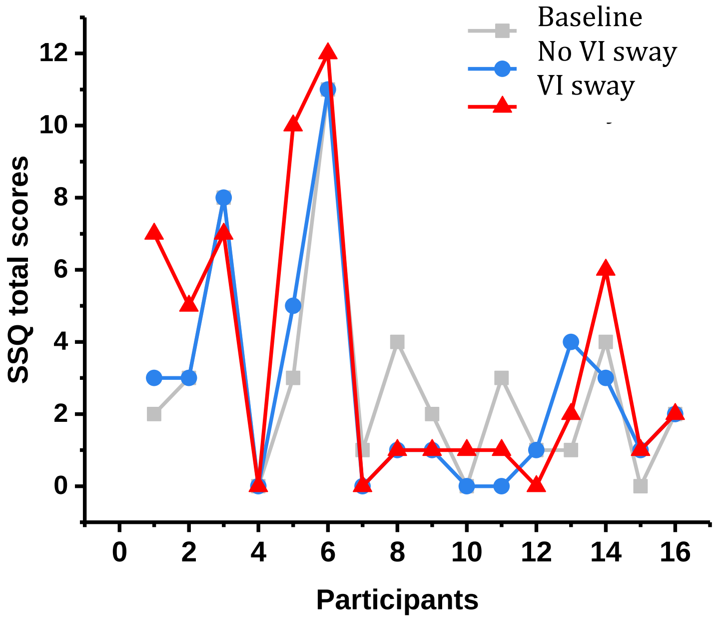
Figure 22. Means and Standard Error in VI sway and No VI sway conditions. (A) Data from the Sway condition, (B) Data from No Sway condition. Labels show what cues were presented during the simulation: Multimodal (audio-tactile-visual), Bimodal (audio-tactile, AV audio-visual, TV tactile-visual), Unimodal (A audio, V visual, T tactile), No feedback (NONE no additional feedback). Data for bimodal (AV, AT, TV) and unimodal (A, T, V) category were group together and the mean score is used for the analysis. Significant effects ($p < 0.02$) were observed only in Sway block; there was only a trend towards significance in No Sway block. Error bars represent standard error of the mean.

4.3.2.3 Simulation sickness questionnaire

The objective results across all trials are mirrored in the subjective ratings of general discomfort. During the task participants were asked to answer one question stating how much general discomfort they experienced. When investigating the subjective scores of general discomfort it was found that the increase in the amount of sensory cues presented decreased the overall ratings in discomfort (Figure 23, A). After each experimental block, the overall simulation sickness scores were recorded on Simulation Sickness Questionnaire (SSQ) (Kennedy et al., 1993). The SSQ consists of three subscales: nausea, oculomotor and disorientation subscale. Data recorded for disorientation subscale remained zero across all participants in all phases (baseline and after each experimental block); therefore this subscale was not included in further analysis. Figure 23 (B) shows overall scores for each participant in both experimental blocks. In the Sway condition, nine participants reported increased simulation sickness; three participants reported decreased simulation sickness and four participants did not report any changes. It should be noted that the four participants who did not reported any changes in sickness scores had previous experience with VR simulation.



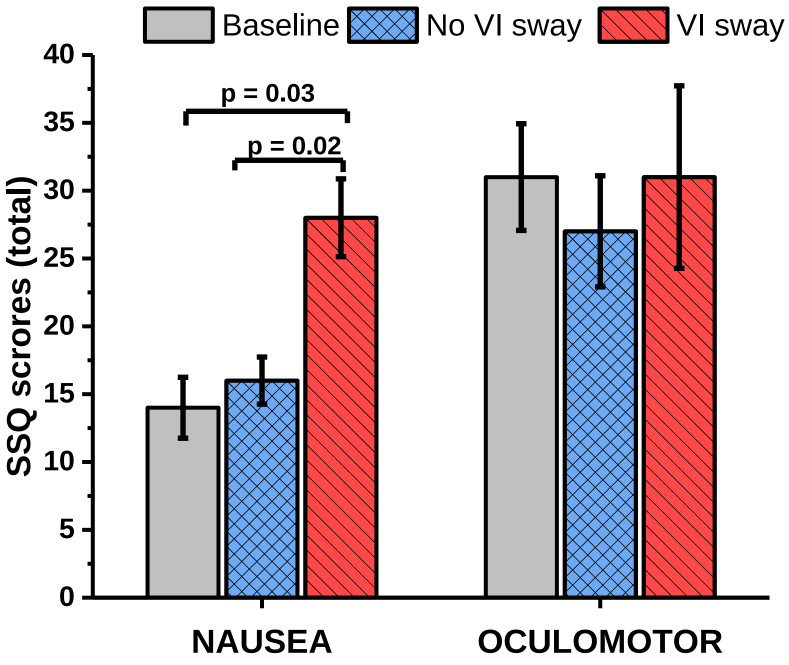
A



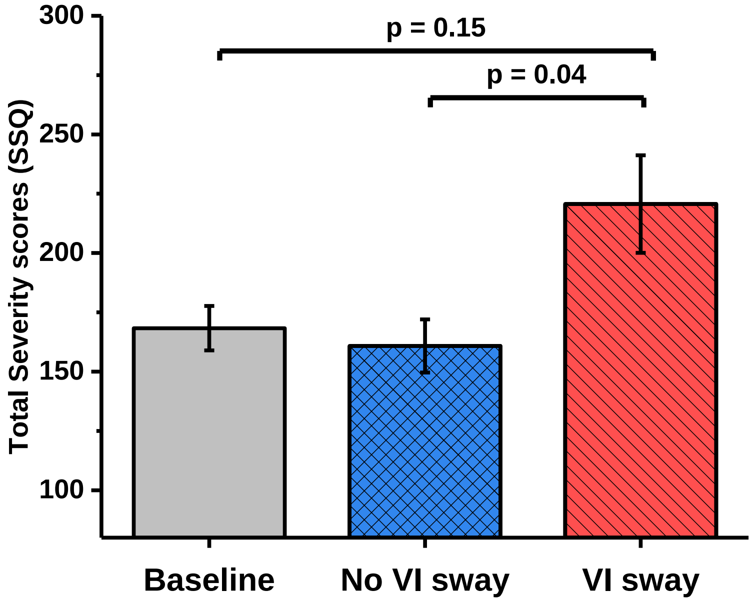
B

Figure 23. Simulation Sickness Questionnaire. VI = visually induced sway. (A) Discomfort ratings recorded after each sensory condition on short questionnaire. (B) Total SSQ scores across all participants: baseline (grey), No Sway (blue) and Sway (red) block.

As the data collected on SSQ scale were ordinal data, a nonparametric Friedman ANOVA was used to analyse whether the weighted SSQ scores were affected by the manipulation of the environment with three dependent variables: nausea, oculomotor and total SSQ scores. The analysis revealed a statistically significant difference in sickness scores across nausea and oculomotor subscale and total scores (x^2 (5) = 42.116, p = 0.0001, W = 0.72). As it was predicted that the VI sway block will produce increased levels of simulation sickness, one-tailed planned comparisons with Wilcoxon sign rank test were performed on Nausea and Oculomotor subscales and total SSQ scores. A significant increase in total sickness scores in VI sway block as compared to No VI sway block was observed (Z= -1.691, p = 0.045, r = 0.29). The scores on Nausea subscale differed significantly between baseline and VI sway block (Z = -1.806, p = 0.035, r = 0.31) and between VI sway and No VI sway experimental blocks (Z= -2.047, p = 0.02, r = 0.36). The possible explanation of high baseline score for Oculomotor subscale can be interpreted through demand characteristics, previously reported by Young, Adelstein and Ellis (2007). The differences on Oculomotor subscale during the experimental task were observed; however these were found not to be significantly different (Figure 24).



A



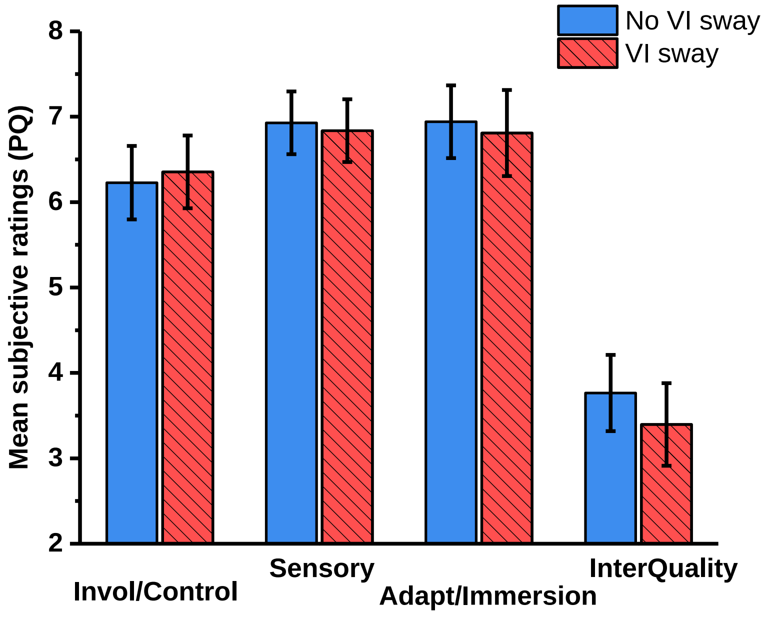
B

Figure 24. SSQ subscale analysis. VI = visually induced sway.
(A) Overall scores on the Nausea and Oculomotor subscales of the SSQ during the manipulation of the environment; (B) total severity scores. Non-parametric paired sample t-test identified a significant increase in nausea during the Sway block. Error bars represent variance.

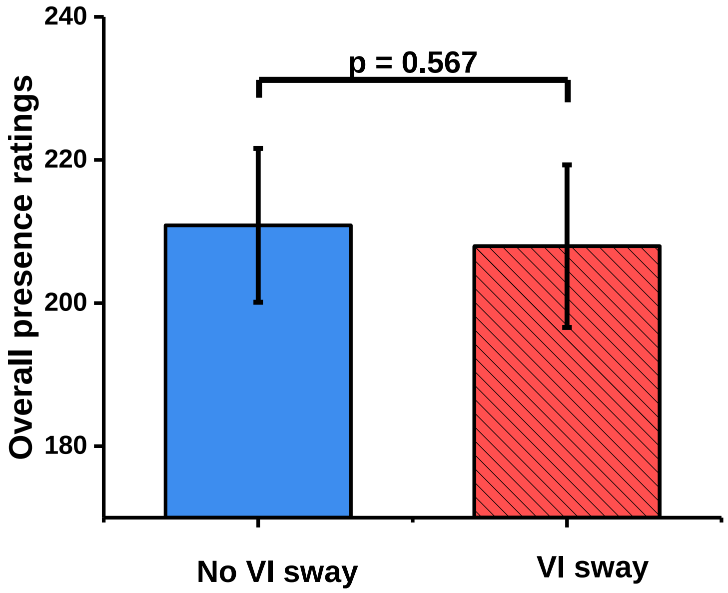
4.3.2.4 Presence questionnaire

The effects of positional inaccuracy implemented through the motion tracking system on perceived sense of presence were recorded on Presence Questionnaire (PQ) after each experimental block. The reliability of the questionnaire was assessed by calculating a coefficient Cronbach alpha. The standardised alpha of the PQ questionnaire (presence = 0.948) reports good reliability with coefficient greater than suggested value of 0.7 (Nam et al., 2008). The presence ratings after each experimental block were collected and the mean scores were compared. In No VI sway block the mean PQ score for 16 participants was 209.32 (SD = 45.2) with a maximum score of 279.8 and a minimum score of 120. The mean PQ score in the VI sway block was 207.4 (SD = 48.3) with a maximum score of 275.1 and a minimum score of 79.6. Overall, participants experienced a slightly higher sense of presence in No VI sway than in the VI sway block, however this difference was not significant (t(16) = 2.141, p = 0.56, d = 0.16). Faas et al. (2014) suggested that a score above 100 should be taken as a reference point for a high level sense of presence. All mean scores on PQ were above 100 and as the difference in the mean scores between two experimental blocks was minimal, it can be concluded that both experimental blocks evoke high levels of presence during virtual interaction.

Previous research that investigated the sense of presence in VR using the PQ questionnaire, conducted a factor analysis in order to examine the internal validity of the questionnaires. Witmer and colleagues (2005) analysed data from seven experiments (N=352) in order to extract a 4-factor model of presence that includes Involvement/Control subscale, Sensory subscale, Adaptation/Immersion subscale and Interface Quality subscale (Witmer et al., 2005). All of the factors extracted (except Adaptation/Immersion subscale) corresponded to those identified in a cluster analysis of data from an earlier version of the questionnaire (Witmer & Singer, 1998) and thus all the subscales were shown to provide a robust and reliable measure of presence (Witmer & Singer, 2005). In this study the data were sorted in accordance with Witmer and Singer's (2005) 4-factor model of presence with four subscale: Involvement/Control subscale (11 items, M = 6.2, SD = 1.71), Sensory subscale (6 items, M = 6.8, SD = 1.47), Adaptation/Immersion subscale (8 items, M = 6.8, SD = 1.85) and Interface Quality subscale (3 items, M = 6.3, SD = 1.95). Figure 25 shows the mean presence ratings on all 4 subscales during both experimental blocks (No VI sway and VI sway). Repeated measures ANOVA was performed on the PQ data, however no significant main effects or interaction were found between the two experimental blocks. The prediction that VI sway experimental block will decrease the overall feeling of presence was not supported. The results show that the sensory and Adaptation/Immersion subscales received the highest scores in both experimental blocks, which suggests that sensory presentation was the most dominant in influencing the feeling of presence (Witmer & Singer, 1998).



A

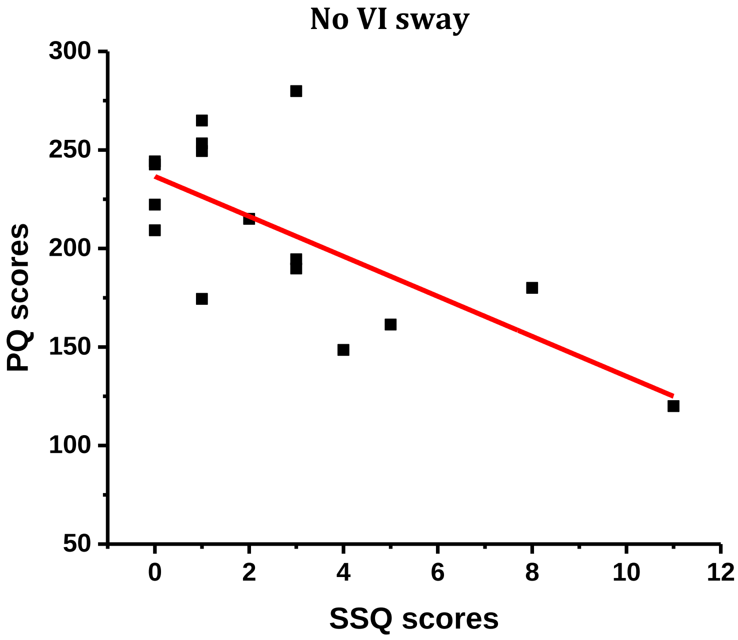


B

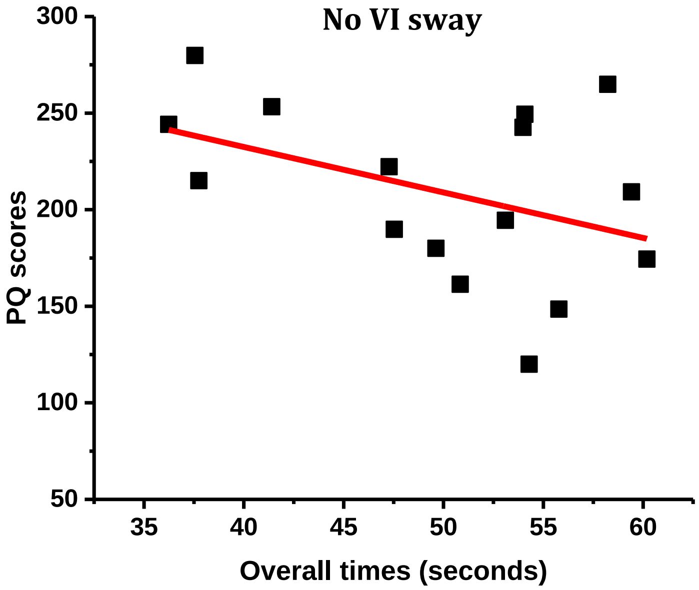
Figure 25. PQ subscales. VI = visually induced, (A) Means and standard errors for each subscale on PQ. (B) Overall score on PQ. Error bars indicate a standard error of the mean. Higher response indicates increased sense of presence. Sensory and Adaptation/Immersion subscales were scored the highest suggesting that presented sensory cues and the overall quality of the VR environment were dominant in influencing the sense of presence. Interface Quality was more evident in Sway condition than in No Sway as it received lower ratings.

4.3.2.5 Correlation

To investigate relationships between the perceived discomfort, total completion times and feeling of presence, correlational analyses in both experimental blocks were performed. In No Sway condition, a significant negative correlation between SSQ scores and PQ scores was found (r = - 0.585, p = 0.017). No significant correlation was found between SSQ and overall completion time (p = 0.94) or between PQ scores and overall completion time (p = 0.11) (Figure 26). In Sway condition significant negative correlations were found between SSQ scores and PQ scores (r = -0.639, p = 0.008) and between PQ and overall completion time (r = -0.666, p = 0.005) (Figure 26). No correlation was found between SSQ and overall completion time in Sway condition (p = 0.19) (Figure 27). Negative correlation implies an opposite direction of collected scores, which means that when a higher degree of sickness was experienced, the overall sense of presence decreased in both experimental blocks (No Sway and Sway). Additionally, when the scores on the presence questionnaires increased the task was performed much faster, but only in Sway condition thus our prediction was only partially supported.



A



B

Figure 26. Correlation between presence and sickness scores (A) but no correlation was found between overall time and presence scores (B) in the condition where no visually induced sway was present (No VI sway).

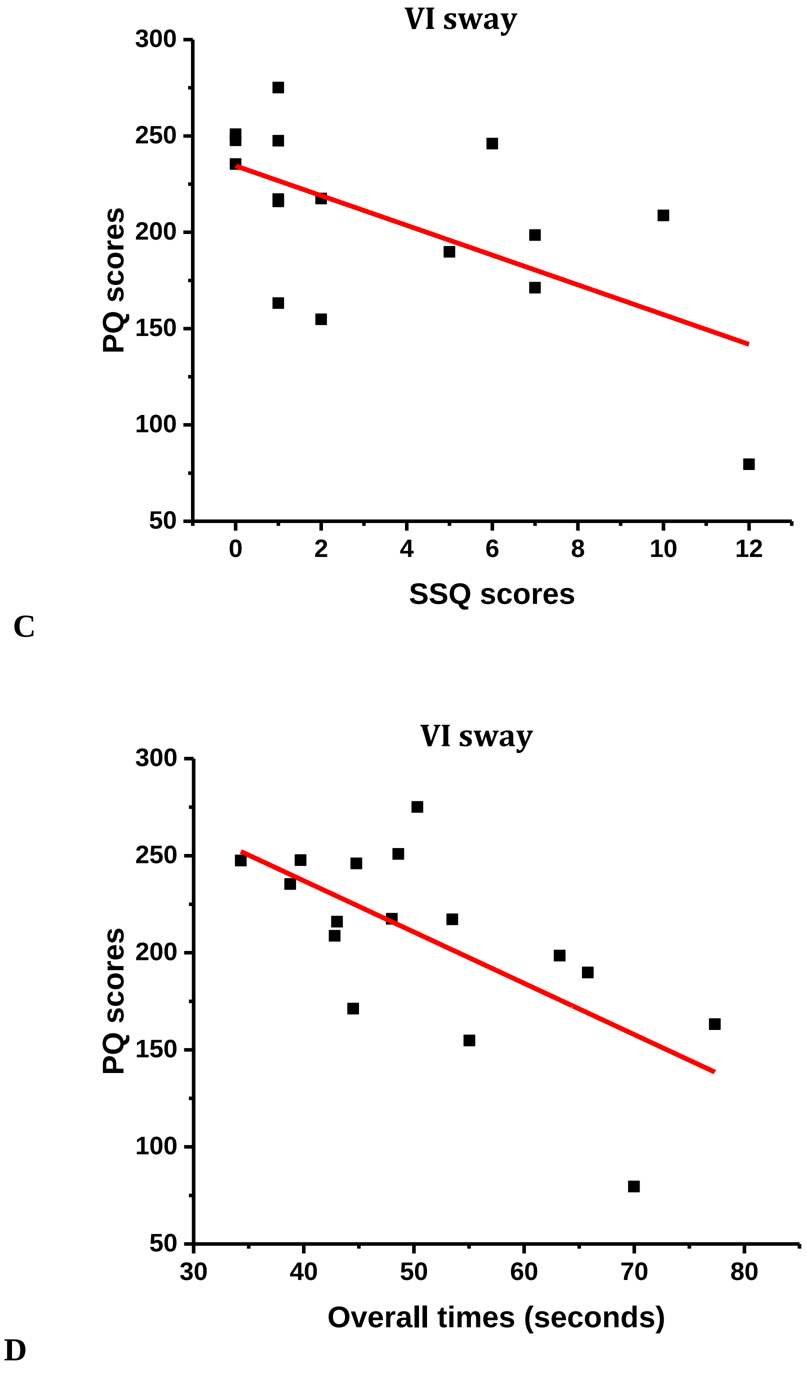


Figure 27. Correlational analyses in visually induced sway condition (VI sway). Significant negative correlations were found between SSQ and PQ scores and between overall times and PQ ratings.

4.4 Discussion

The main aim of this study was to investigate how additional augmented multisensory feedback affects human performance and subjective evaluation during a completion of the manual assembly task in the VR environment. Additionally, the effects of positional inaccuracy in the motion tracking system on performance, perceived sense of presence and simulation sickness were also investigated. Previous research has argued that the evaluation and comparison of VR environments needs to be based on usability metrics and measurements of perceived latencies (Lugrin et al., 2013). In this study, participants interacted with a VR environment: they performed a wheel change task on a virtual racing car whilst usability metrics were recorded. During the task, different combinations of augmented sensory cues were provided as unimodal, bimodal and multimodal feedback. The task was performed in two experimental blocks: one where an accurate motion tracking was available (NO visually induced sway, No VI sway) and one where a motion tracking latency was simulated in the VR in order to mimic the latency of low cost motion tracking systems (visually induced sway, VI sway), previously reported in other studies (Khoshelham & Elbernik, 2012). The main aim of this study was twofold: a) to investigate the effect of augmented sensory cues and b) to examine the effects of positional inaccuracy in the motion tracking system. In both cases, the user’s performance and behaviour was recorded and measured through the set of objective and subjective measures and statistical analysis was performed to understand the effects of the stimulus manipulations.

4.4.1 The effects of augmented cues on performance

The main finding in this study is the fact that when augmented cues are presented in the VR environment participant's task performance improves. In particular, it was found that augmented audio and tactile cues provide additional information in an efficient manner without being distractive. The advantages of audio and tactile feedback have been noted in previous research; audio feedback improved task performance in terms of accuracy and spatial attention without affective perceptual workload (Jacko et al., 2004; Meyer et al., 2012; Santangelo & Spence, 2007; Martinez et al., 2011; Schoonmaker et al., 2006; Jai et al., 2011; Lee & Bilinghurts, 2013; Shi et al., 2007; Oviat et al., 2004). Similarly, the presence of haptic feedback improved interaction, spatial guidance and learning in VR (Weber et al., 2011; Spelmezan et al., 2009; Van Der Linden et al., 2011; Van Erp et al., 2004; Ramsamy et al., 2006, Burke et al., 2006). In this study, the analysis of the objective measures revealed medium significant effects of audio and tactile modalities on task performance; however no significance effect of visual modality was obtained.

This finding could be explained by the visual information overload (Jeon, Davison, Nees, Wilson & Walker, 2009): during the visual condition no audio or tactile feedback was provided thus all information was taken through the visual modality. A conventional projection-based VR display used in this study presented complex visual information and thus the additional visual information might not have sufficiently influenced the participant's performance.

This finding is also consistent with the Multiple Resource theory proposed by Wickens (1986). The main proposition in this psychological theory relates to the fact that there is a limited amount of attentional resources available for the perception of sensory information. This theory suggests that there is a greater interference in task performance when most of the information is perceived through single modality than when it is perceived through multiple modalities. In this study the task presented main visual information on the screen and additional visual augmented information that was relevant. However as both were presented in one modality (visual) the additional visual cues might of exceeded attentional resources for visual information and as such the significant beneficial effects of additional visual cuing was not observed on the overall completion times. However participants reported that additional visual cues facilitated their sense of involvement in VR; therefore further research is needed to fully explore the potential benefits of additional visual cues.

Previous research has shown the benefits of multimodal interfaces for spatial understanding (Schudart & Bowman, 2007) placement accuracy performance (Richard et al., 2006; Jacko et al., 2004) gaming environment (Nam et al., 2008) and enhanced performance in dynamic threat scenarios (Oskarsson et al., 2012). Similarly, in this study it was shown that multimodal feedback was most helpful for overall task performance as well as for enhancing the levels of presence and involvement. This provide support to the findings from previous studies that suggested that appropriate sensory cues can improve task performance (Schudart & Bowman, 2007; Oskarsson et al., 2012; Richard et al., 2006; Jacko et al., 2004; Nam et al., 2008). One surprising finding arising from this research was the effectiveness of audio and tactile feedback presented unimodally. It has been noted that audio and tactile feedback can be perceived as distracting and annoying and can even decrease the accuracy of performance (Vitense et al., 2003; Oakley et al., 2009). However, in this study unimodal audio and tactile feedback influenced performance in a positive manner. Individually, the mean time for audio condition was 49.7 seconds and for tactile it was 49.9 seconds. In comparison to bimodal conditions, such as AV (49.1s), AT (49.5s) and TV (50.2s) it can be seen that audio and tactile feedback facilitated performance in the same manner as bimodal feedbacks. This is confirmed by the result of the analysis that showed no significant difference between unimodal and bimodal modes of feedback in the mean completion times ($p = 0.25$, $d = 0.16$). This suggests that the presentation of audio and tactile feedback alone was as effective in supporting task performance as bimodal feedback presentations.

These findings can be also explained by the Multiple Resource Theory proposed by Wickens (1986). As mentioned before, when information is coming through one modality some interface will occur as when information is received through multiple modalities. In this case, audio and tactile cues were provided in addition to basic visual information. As attentional resources for each modality were not exhausted these cues were perceived as useful and facilitated overall performance.

The analysis of the subjective ratings of user's perception in VR environments showed similarities to the objective measures of performance. All modalities had a large significant effect on perceived sense of presence, which can be seen when looking at the magnitudes of the effect sizes. Similarly, the significant interactions between modalities suggest that all of them were effective in producing significant levels of presence. Participants rated the multimodal condition as most effective in enhancing their sense of presence, followed by bimodal and then by unimodal feedback. The magnitude of these effects is substantial and is confirmed by the effect sizes in the analysis, thus confirming that availability of sensory cues during the VR interaction can influence users perception in VR (Bowman & McMahan, 2007; Laha et al., 2012; Meyer et al., 2012). When each sensory modality is compared in terms of their presence, the difference between bimodal and unimodal feedbacks showed no significant differences. When looking at the scores for individual conditions it can be seen that audio and tactile feedback presented unimodally (A = 5.92, T = 5.69, V = 5.47), were scored similarly to bimodal feedback (TV = 5.98, AV = 6.04, AT = 6.23) (see Figure 7 and Figure 8). Thus, it can be concluded that all modes of feedback presentation influenced the levels of presence during the VR interaction.

The main intention of this study was to present additional augmented sensory information to aid performance in VR rather than improve the fidelity of the environment. Nonetheless, the participants reported an increased sense of presence when augmented cues were present. The increased sense of presence was recorded even when overall fidelity of the VR environment was decreased. This finding could be explained by the type of the task performed as this task had a high level of interactivity. Empirical research has suggested that the degree of presence experienced in the environment depends not only on the fidelity of sensory information but also on interactivity and involvement (Witmer and Singer, 1998; Witmer et al., 2005). Keyson (2007) and many other have suggested that interactivity is a main factor that can contribute to the sense of presence. The task in this study involved a high level of interactivity as participants had to directly interact with the virtual objects and perform that task against the clock. Participant felt much more involved and immersed and as a result they experienced and reported an increased sense of presence. The fact that the interactivity levels were high might of contributed to the increased levels of the sense of presence even when the fidelity of sensory information was decreased. These findings therefore provide further support the notion that interactivity and involvement are important for perceiving sense of presence (Witmer & Singer, 1998; Witmer et al., 2005; Keyson, 2007).

In addition, these finding also supports the notion that the overall fidelity of the environment may not be such an important factor as it was previously suggested (Bowman & McMahan, 2007; Laha et al., 2012). Previous studies have shown that increasing the fidelity of various components of VR to increase immersion, rather than overall VR set up, could be sufficient to observe the beneficial effects on performance (Ma & Kaber, 2006; Bowman & McMahan, 2007; Laha et al., 2012; Faas et al., 2014; Jai et al., 2011). Findings from this study support previous research and further imply that some tasks in VR can be successfully performed with less immersive and less costly displays with no loss on efficiency (Bowman & McMahan, 2007; Laha et al., 2012, Ma & Kaber, 2006).

In this study, lower overall completion times indicate better performance, i.e. faster is better. The opposite is true for subjective ratings, where higher scores indicate an enhanced sense of presence. To account for idiosyncratic mapping between participants in objective and subjective data, correlation analysis was performed on the level of sensory conditions. When investigating the possible relationship between mean task completion times and subjective ratings of presence a significant negative correlation was found. This suggests that when participants experienced an enhanced sense of presence, their mean completion times were much faster. As previous studies have suggested (Spelmezan et al., 2009; Faas et al., 2014; Bowman & McMahan, 2007; Laha et al., 2012; Ma & Kaber , 2006) these results reflect the fact that multimodal feedback can maximise human physical abilities as well as enhance the user’s sense of presence in the VR environment.

4.4.2 The motion tracking inaccuracy and performance

Murray et al. (2003) suggested that after the display device, tracking is the most important component within large screen projection VR systems. Thus accurate motion tracking of users and objects in VR is necessary for effective interaction (Lubeck et al, 2015). Participants in this study interacted with virtual objects whilst performing the manual assembly task in two experimental blocks. Faster mean completion times were recorded when accurate motion tracking (NO VI sway) was provided as compared to the experimental block with motion tacking inaccuracy (VI sway); however a significant difference between the mean times was not obtained. This result could be explained by Sprague et al. (2006) who measured the effects of error in correct and incorrect head tracking and suggested that sensorimotor adaptation allows users to compensate for incorrect tracking. The sensorimotor adaptation paradigm could explain why no significant differences were observed between the two VR environments in this study. The scores recorded on Sensory and Adaptation/Immersion subscales in PQ suggest that presented cues helped participants mitigate the negative effects of tracking inaccuracy in VR. This finding is important as it implies that the presentation of augmented multisensory cuing can decrease the perception of evident and sometimes unavoidable motion tracking inaccuracy of the VR system.

4.4.3 The motion tracking inaccuracy and simulation sickness

Negative effects of the motion tracking inaccuracy during the interaction process in VR were observed not only on the mean performance times but also on user’s subjective feeling of discomfort (Murray et al., 2003). The scores for sickness symptoms identified in SSQ were significantly lower during the accurate motion-tracking block (No VI sway) as compared to the experimental block where motion tracking inaccuracy was introduced (VI sway). When positional inaccuracy was present, increased symptoms were recorded on Nausea subscale, such as sweating, nausea and dizziness. Oculomotor symptoms, such as headache, eyestrain and fatigue decreased during the accurate motion tracking condition; however they increased again when participants performed the task in the environment with inaccurate motion tracking. Positional inaccuracy is often observed with low cost motion tracking systems in terms of error measurements. Previous studies have reported some advantages of these systems (Zollhoffer et al, 2011, Shin et al, 2013), however, when used for human motion tracking in virtual settings the use of these systems is very limited in terms of tracking ability and accuracy (Khoslejham & Elbbernik, 2012, Schaffer et al., 2015, Greuter & Roberts, 2014). The findings from this study confirms previous research and further suggests that the use of low cost motion tracking systems should be avoided in human motion tracking as the positional inaccuracy associated with them can have negative effects on user's acceptability of the VR.

The reported sickness symptoms were found to be not related to the overall exposure time in VR, as longer the participant spend in the VR the fewer symptoms they reported. This finding is in line with previous studies showing that longer exposure is associated with decreased feelings of discomfort (Kennedy et al., 2010). The reason for a decrease in simulation sickness during longer exposure might be explained by the adaptive abilities of the nervous system (Novak, Mihejl & Munich, 2012). The nervous system adapts to an environment gradually, which was shown in the studies where repeated VR exposure reduced symptoms of simulation sickness (Kennedy et al., 2000). However, this might cause a problem especially in the development process of VR systems, as developers who are used to virtual environments might underestimate the severity and frequency of the aftereffects. A continuous investigation of the factors modulating sickness symptoms is therefore encouraged in future research in order to inform researchers and developers how to design tasks and VR set ups to effectively manage unpleasant aftereffects of VR exposure.

4.4.4 The effects of motion tracking inaccuracy on sense of presence

The effectiveness of VR has been also linked to the sense of presence, which can be defined as the subjective experience of being in one place or environment even when one is physically situated in another (Witmer & Singer, 1998). The relationship between the task performance and subjective feeling of presence in VR has been investigated in many studies (Der Streathen, 2000, Ma & Kaber, 2006). For example, Faas et al., (2014) suggested that the level of presence might serve as an indicator of performance and learning during the task in VR. Inaccurate motion tracking, such as visual latency between user's movements and the update of the VR set up, can have negative effects on the feeling of presence and involvement and as a result, decrease task performance (Murray et al., 2003). The motion tracking inaccuracy in this study was presented as 0.5Hz forward-and-backwards movement that corresponds to the error measurements previously found with low cost motion tracking systems (Khoshelham & Elbernik, 2012). It was predicted that the lower levels of presence will be recorded during the inaccurate motion tracking (VI sway) as compared to accurate motion tracking (No VI sway). However, the results shown that both environments were equally efficient in producing the feeling of presence. This finding could be due to the presentation of augmented multisensory cues during the task. Augmented cues were presented in a highly unrealistic fashion, and as such decreased the fidelity of the environment. However, the decreased physical fidelity of the sensory cues did not alter the levels of perceived presence in users, mostly due to the fact that the augmented cues carried information that was directly relevant to the task, which increased interaction and involvement (Cooper et al., in press). Moreover, as participants became fully engaged in the tasks, the positional inaccuracy of the motion tracking system became unnoticeable. This could explain why no significant difference between the two experimental blocks was recorded in overall completion times. This finding therefore provides further support for the use of augmented multisensory cues during tasks in virtual reality environments (Cooper et al., in press, Lee & Billinghurst, 2013).

4.4.5 Relationship between variables

The relationship between task performance, sense of presence and simulation sickness was also examined. The negative correlation between sickness and presence scores suggests that when participants experienced an increased sense of presence their sickness symptoms decreased. This finding is in line with Millevele-Pennel and Charron (2015) who reported lower sickness ratings with higher presence scores. Similarly, the correlation between the presence scores and performance implies that a relationship exists between these variables as previously reported (Millevel-Pennel & Charron, 2015, Lin et al., 2002, Sinclair & Kincaid, 2015). This study has shown that augmented multisensory feedback can be used to mitigate negative effects of positional inaccuracy of the motion tracking systems. During the manipulation of the environment, participants also reported lower levels of discomfort when augmented multisensory feedback (visual, audio and tactile) was present. This supports previous findings that reported the beneficial effects of the increased amount of sensory modalities on performance and sense of presence (Dihn et al., 1999). The main implication arising from this research study suggests that the discrepancies in positional accuracy of the motion tracking system can have adverse effects on performance and user experience during the manual assembly task in VR, yet this effect can be mediated by the presentation of augmented sensory cues especially when they present information relevant to the task in hand. Furthermore, in order to avoid negative effects of motion sickness on user' experience a high quality motion tracking system is necessary to allow for natural interaction with human-computer interfaces.

4.4.6 Limitation of the study

As with any other empirical research study some limitations were observed during this study. One of these limitations relates to the individual difference and variability among the participants that could have influenced the findings of this study. A fundamental problem when investigating the relationship between (subjective) self-evaluation and (objective) performance data is in the intrinsic inter-individual variability that is present in the subjective data (Dahlstrom et al., 2012). While (objective) performance is measured relative to an *external* and *common* standard (the time to complete the task), the subjective measure (for example presence ratings) relies on internal scales that are unique to each observer. This means that individual subjective ratings can only reflect the *relative* changes between conditions that are experienced by the users and that there is no common subjective standard or range. While individuals can reliably judge relative subjective changes in their experience across the conditions, there is no common standard for the absolute numerical values. Each participant therefore rated their sense of presence on their internal scale, which may have differed to others. This suggests that the presented findings may have been influenced by this factor.

Despite using a high quality state-of-the-art virtual reality set up, some inferences during the task were observed which could be suggested as another limitation of this study. For example, when performing the task, the tracking of the impact wrench become slightly cumbersome, i.e. the virtual representations froze on the screen, which may have contributed to the loss of concentration and caused the break in immersion and presence. The subjective ratings might of been influenced by this interference. Although this limitation cannot be easily addressed, future research could concentrate on providing a high quality object tracking in VR to eliminate these drawbacks when investigating presence and immersion.

Chapter 5

The effects of augmented sensory cues on training transfer

Chapter overview

This chapter covers findings from an experimental study that investigated the training transfer from virtual to real environments. The introduction will cover a background review of training methods used within VR environments and present techniques used to assess the task performance. The experimental aims and results of this study will be described in the following sections and the chapter will conclude with a discussion of the study findings.

5.1 Introduction

5.1.1 Virtual training

The most important goal of any training method is to increase the level of skill that can be transferred into real life situations. Virtual reality (VR) training has a potential to train an individual to a high level of objectively measured skill before they can transform the knowledge into real life situations (Seymour, Gallagher, Roman, O'Brien … & Satava, 2002). The effectiveness of the training transfer from virtual to real environments has been shown in previous research. For example, aviation training research has long recognised the beneficial effect of flight simulators for pilot training (Hays, Jacobs, Prince & Salas, 1992), military combat training and navigational skill training that translate well into real situations (Loftin, Scerbo, McKenzie, Catanzaro … & Perry, 2004). Medical research has demonstrated that skills learnt in VR significantly reduce error rates and successfully transfer into a real clinical setting (Ahlberg, Enochsson, Gallagher, Hedman … & Arvidsson, 2007; Park, MacRae, Musselman, Rossos … & Reznick, 2007; Gallagher, Seymour, Jordan-Black, Bunting, McGlade & Satava, 2013). Virtual reality applications were found to be a great asset in architectural and design training (Richard, Chamaret, Inglese, Lucidarme & Ferrier, 2006), in sports training (Miles, Pop, Watt, Laurence & John, 2012) and in industrial assembly training (Yunus, Baser, Masran, Razali & Rahim, 2011) Additionally, it was also found that in industrial maintenance and assembly task it is not only virtual reality but also augmented reality training that can be very effective (Gavish, Gutiérrez, Webel, Rodríguez… & Tecchia, 2015). However, some studies have also reported that the beneficial effects of training are not always transferable into real life situations (Farrell, Arnold, Pettifer, Admas, Graham & MacManamon, 2003; Wierinck, Puttemans, Swinnen & Steenberghe, 2005). A further investigation of the factors that can aid and support effective training and learning transfer is therefore necessary.

5.1.2 Factors influencing the effectiveness of virtual training

There are many factors that can affect learning in VR and multisensory feedback was shown to be one of these. Virtual reality scenarios relay on visual modality as the main source of sensory stimulation, however, audio and haptic feedback has also shown to facilitate performance in virtual environments. For example, multisensory feedback has been very effective in dual task and manipulation task studies (Jeon et al., 2009; Adams et al., 2001), in surgical simulations (Cao et al., 2007), in rehabilitation (Feintuch et al., 2006) and in training (Meyer et al., 2012). The sense of presence was identified as another factor that can increase the likelihood of learning in VR and the general assumption suggests that when more sensory signals are stimulated the feeling of presence will be enhanced (Ma & Kaber, 2006). Studies have confirmed this when investigating the effects of multimodal feedback (Dinh et al., 1999) and the effects of design features in virtual environments (Ma & Kaber, 2006). Furthermore, research studies in military and driving simulator training have shown that there might be a relationship between the perceived sense of presence and performance (Steven & Kincaid, 2015; Deniaud et al., 2014). Similarly, cognitive load associated with a given task can also influence training effectiveness. Improved performance and decreased workload was observed in 3D viewing studies (Klein et al., 2009), flight simulation studies (Singh et al., 2005) and in the industrial studies (Stone et al., 2011). Leung et al. (2010) argued that investigating the conditions under which mental workload can increase or decrease can help reduce human errors as well as improve the system safety and increase productivity. However, others have opposed these findings when they reported no significant differences in workload scores between various training conditions (Balaji et al., 2015).

5.1.3 Assessments of training

In order to assess performance, learning curve techniques are used as an efficient tool to monitor performance in new and repetitive tasks in many industries (Sanchez, Matt, Goh & Case, 2013; Anzanello & Fogliatto, 2011, Guimarães, Anzanello & Renner, 2012; Ramsay, Grant, Wallace, Garthwaite. Monk & Rusell, 2001). Learning curves were first introduced in the 1930’s by T.P. Wright and were used initially in the aircraft production industry. Wright found that there was a pattern in the way that people learned. The central proposition of learning curve theory is that performance will improve with increasing cumulative task experience, at a decreasing rate (Wright, 1936). The proposed shape of the learning curve has been evidenced across a diverse range of jobs and industries, from airplane manufacturing to pizza making, and across different levels, from the employee through the organizational to the industry level (Lapré & Nembhard, 2010, Krausert, 2015). Learning curve has been identified as an efficient tool to monitor performance in new and repetitive tasks in automotive, aerospace, construction and electronic industries (Sanchez et al., 2013, Anzanello & Fogliatto, 2011, Guimaraes et al., 2012) in the assessment of health technologies (Ramsay et al., 2001). While several methods exist, no one learning curve model is generally accepted as the best. The first formal LC model that was proposed by Wright (1936) had the following mathematical representations:

(Equation 4)

$$Y = \mathrm{C}^1 \, x \, b$$

where y is the average time (or cost) per unit demanded to produce x units, and C_1 is the time (cost) to produce the first unit. Parameter b ($-1 < b < 0$) is the slope of the LC, which describes the workers' learning rate. Values of b close to -1 denote high learning rate and fast adaptation to task execution (Anzanello & Fogliatto, 2011). Further mathematical modification on Wright's model enabled estimating the total time to produce x unit and the time required to produce a specific unit I by means of another equation were suggested by many researchers. The learning curve technique was also used in this study in order to evaluate the performance of users in each experimental group. The comprehensive background on learning curve is beyond the scope of this thesis however; interested readers are referred to Anzanello and Fogliatto (2011).

5.1.4 Summary of experimental aims

In the previous chapter (Chapter 4) I have showed that additional augmented sensory cues used within virtual environments can improve performance and enhance user's acceptability of VR. The main focus of this study was to investigate whether virtual training with augmented cues translates into performance improvements in real life environments. Similar to previous experiments, the projection based virtual reality was used as a VR training platform. However, this time the virtual system has been changed; the projection screen is smaller than in previous studies and additional features were added into the virtual simulation. The main aim of this study was to train participants in the virtual wheel change task and then let them perform the wheel change on a real racing car. Three groups of participants were recruited: two groups were trained in the virtual environment, one group perform the real wheel change task only. It was predicted that the group that received virtual training would perform the real wheel change more efficiently than the group that did not receive virtual training. Throughout the study, objective and subjective measures were recorded and compared to examine the relationship between the measured variables. It was predicted that the type of virtual training (with or without augmented sensory cues) would influence user's performance and subjective perception of VR. It was also predicted that the type of virtual training would influence task performance and workload levels on the real wheel change task.

5.2 Methods

5.2.1 Participants

A power analysis (Faul et al, 2009) was performed to determine a sample size required for this study and showed that for 0.05 level of significance, the effect size of

0.5 and power (1 – β err prob) of 0.9, the total sample size required was 34. In total, 42 participants were recruited via opportunity sampling. Participants were randomly allocated into three training groups but were matched across age, gender and experience level. Each group consisted of 8 females and 6 males with the age ranging between 17 and 30. All participants signed a consent form and reported normal or corrected-to-normal vision and hearing.

5.2.2 Virtual reality set up

The virtual training was conducted in the Virtual Engineering Centre (VEC) laboratory that is located in the School of Engineering at the University of Liverpool. The main task was to perform a wheel change, either with or without additional augmented sensory cues. After the virtual training, participants were required to perform the wheel change on a real racing car.

The VR setup consists of one Active Mode display screen, 2.74m in length and 1.72m in height, behind which is one active stereo projector that creates 1920 x 1200 resolution images at a rate of 120Hz. When performing the task participants were required to wear wireless LCD shutter glasses that are synchronized with the projectors to provide stereoscopic images. 6 high-spec infrared cameras (4 Bonita 10 and 2 Bonita 3) are used to record and track object positions: LCD shutter glasses (for head tracking and POV adjustment), which enabled 3D stereo view; haptic gloves and impact wrench (1.94kg, 15.2cm) (Figure 26). A faithful digital mock-up of the impact wrench is used to interact with the bolts. Position data, computed using VICON Tracker software, is broadcast in real-time across the internal network using a VRPN protocol at a rate of 200Hz and is used to update the virtual environment.

5.2.3 Sensory stimuli in VR

The wheel change simulation is implemented in Unity and runs at an average frame rate of 75fps. Two stereo speakers positioned next to the projector behind the screen provide audio feedback. Two tactile gloves, with a vibration motor attached to the palm, provide tactile stimulus when hands collide with or grab objects. The vibration occurs with variable frequency ranging from 15 to 250Hz and variable vibration amplitude (up to 7g). The change of vibration intensity helps to reinforce learning by providing more detailed information in specific situations. For example, the fastening of the bolt starts with an intermediate vibration level and as soon as the bolt is fully inserted in place, the tactile stimulus (vibration) on the hand holding the impact wrench increases to the maximum value, mimicking what happens in real life with a real tool. Apart from obvious visual feedback, participants are also provided with additional visual cues that are presented as a colour change (red and yellow) of simulated hands, wheels and bolts to signal contact with the virtual objects. The main difference between the two wheel change scenarios (Chapter 2 and Chapter 4) was that this time participants had to remove the bolts from the impact wrench with their hands and place them into a tray located 68cm to the left of the virtual racing car. They were also required to press a button to change the direction of screwing when the bolts needed to be unscrewed and screwed back in. The main reason for this was to simulate the required steps that need to be performed during the real wheel change task as realistically as possible in terms of functional fidelity.

5.2.4 Real task stimuli

For the real wheel change task, the frame of a real racing car was used that was provided by the Formula One team of the University of Liverpool (Figure 28).

The wheels used in the task were attached to the car by four bolts. For the real task, a Clark Air ½” impact wrench (2.1kg) that was attached to the compressed air line via an expandable recoil hose was used and had a 19mm socket attached to manipulate the bolts. During the experiment participants were required to wear protective overalls, steel-toe capped boots, gloves, eye protection and ear defenders.

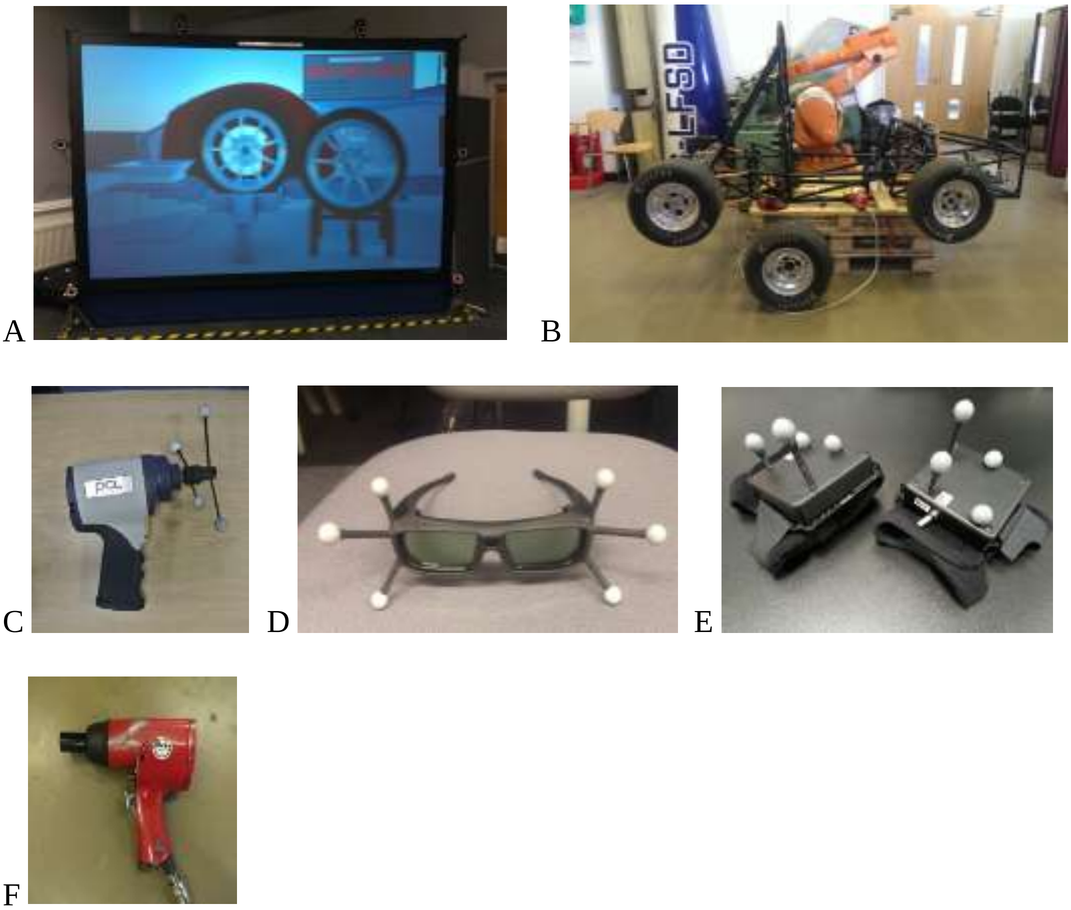


Figure 28. Virtual (A) and Real (B) set up for the wheel change. The equipment used in the virtual task included: impact wrench (C), LCD shutter glasses (D) and haptic gloves (E). The equipment used in the real task included: impact wrench (F).

5.2.5 Task procedure

All of the participants self-declared no previous experience with the task or projection screen VR environment. Participants were randomly allocated into three experimental groups. One group performed and trained on the real wheel change task only (RO group). The next group was first trained in the virtual environment with minimalistic cues, where only basic visual information was provided (NC group). The final group was trained in the virtual environment with additional augmented sensory cuing during the whole simulation (ATV group). Immediately after virtual training, both groups performed the wheel change task in the real environment.

5.2.5.1 Real task procedure

Before the real task, participants provided informed consent and filled in pre-test questionnaires. These questionnaires consisted of participants' demographics: name, age, gender, occupation and any previous experience with wheel changes. Before starting the task participants were asked to complete a simple dexterity test to assess their manual motor skills and colour vision in order to match the groups in terms of their ability to perform the task. After this, the experimenter explained the task procedure and showed a full demonstration of the task. Participants were instructed to perform the task as fast as they can without any errors, i.e. no loose bolts. If, after completing the task, the experimenter was able to unscrew the bolt by hand, a five second penalty for each unfastened bolt was added to the overall completion time. After the first run, participants were required to fill in a workload questionnaire and then they were instructed to perform the task four more times as fast as they can. At the end, participants filled in another workload questionnaire. To maintain motivation, a leader board was displayed where the participants' fastest times were shown and updated after each participant had completed the task. After completing the whole task, participants were debriefed and thanked for their participation.

5.2.5.2 Virtual task procedure

The participants that were assigned to the virtual training groups started the task at the VEC laboratory. Before the task, they filled in pre-test questionnaires to obtain baseline measures and performed a dexterity test to assess their manual motor skills and colour vision. Then the experimenter performed a demonstration of the whole task whilst participants were wearing 3D shutter glasses. After the demonstration, participants put on haptic gloves and were instructed to perform the task themselves. The time started at the moment when the subject grabbed the wrench for the first time, corresponding to the first collision event between hand and tool in the virtual environment. After the first trial run, they were asked to fill in sets of questionnaires and then they performed the simulated wheel change task four more times. Two groups of participants performed the virtual wheel change task: NC group received training with minimalistic cues (no vibration, no change of colour, no audio); ATV group received virtual training whilst additional augmented sensory cues were present during the task (additional visual information, vibration, audio cues). After the task was finished participants filled in sets of questionnaires and then they were taken into the assembly laboratory where the real wheel change task was performed. The procedure for the real task was the same as described in section 5.2.4.1 with the exception of the dexterity test.

5.2.6 Performance measures

The overall time to complete the task was recorded as an objective measure of performance. Additionally, the amount of errors performed during the real wheel change was also recorded. One error corresponded to one loose bolt that the experimenter was able to unscrew with her own hand, after the task has finished. A penalty of five seconds was added to participants' overall times for each loose bolt. The subjective measures were obtained from the sets of questionnaires that participants filled in before, during and after the task. These included Immersive Tendencies Questionnaire (ITQ), Presence Questionnaire (PQ) (Witmer & Singer, 1998), the National Aeronautics and Space Administration Task Load Index, in short called NASA TXL workload questionnaire (Hart & Staveland, 1988) and Simulation Sickness Questionnaire (SSQ) (Kennedy et al., 1993) (Appendix B, D, E and F).

5.2.7 Experimental design

This study was a two-way between subject factorial design. Analysis of variance (ANOVA) was performed on the mean completion times in order to investigate the effects of the trial runs and the virtual training group on the overall task performance and subjective perception of the environments. The Mauchly test of sphericity was applied and when significant, Greenhouse-Geisser corrections were adopted. Partial eta squared (η^2) is reported for effect sizes. The significance level for all statistical tests was set at 0.05. Normality of the data was assessed with Shapiro-Wilk test statistics and when the assumption was violated the logarithmic transformation of the data was performed. Planned comparisons were performed using independent sample t-tests to investigate the differences between groups during the first and last trail run. For paired sample t-tests, the Cohen's d was chosen as a measure of effect size (Equation 1 and 2) (Cohen, 1992).

5.3 Results

5.3.1 Dexterity test

Before starting the task participants were asked to complete a simple dexterity test, which consisted of a manual task to assess their motor skills and colour vision. Participants were required to fix together screws and bolts of various sizes as well as arrange the colour blocks into two colour towers (Appendix D). Participants were instructed to perform the task as fast as they can. Performance measures were recorded on a stopwatch. On average it took participants around 336 seconds (SE = 14.2) which is about 5.6 minutes, to complete the dexterity test. The mean score for RO group was 310.7s (SE = 23.6), for NC group 330.71s (SE = 18.3) and for ATV group the mean time was 366.9s (SE = 29.9). Independent sample t-tests were performed to investigate whether there was any difference between the experimental groups in their manual motor skills. The analysis revealed that no significant differences existed between the mean dexterity group scores suggesting that the groups were matched on their abilities to perform simple manual tasks.

5.3.2 Objective measures

In order to investigate training effectiveness, task completion times were recorded after each trial. Each group consisted of 14 participants, 8 females and 6 males: RO group (M = 21.4, SD = 3.82); NC group (M = 20.1, SD = 0.87); ATV group (M = 20.6, SD = 2.27). No significant differences were observed between the groups for age or previous experiences. Descriptive statistics for overall mean completion times in all five trials and across all three experimental groups can be seen in Figure 29.

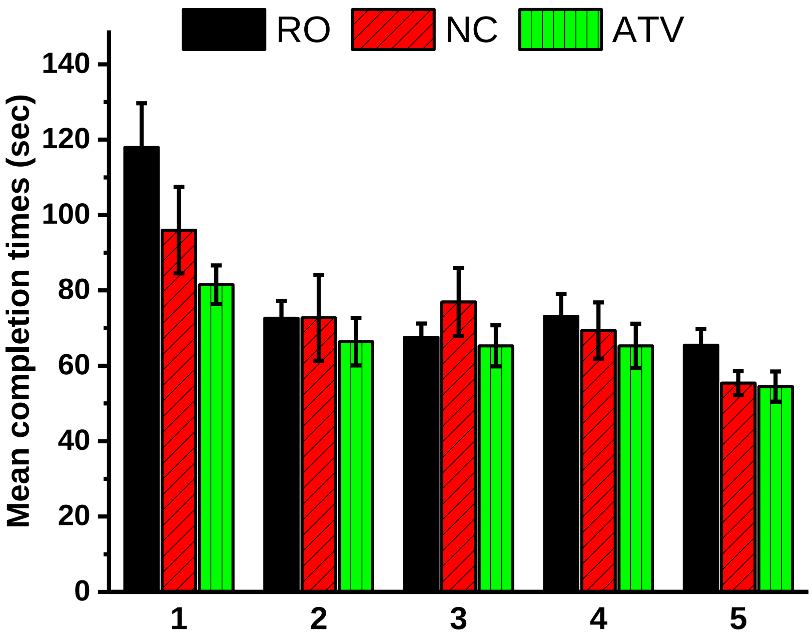


Figure 29. Mean and standard errors in all five trial runs for real only group (RO black), minimal cue group (NC red) and multisensory group (ATV green).

To assess the normality of the data, the Shapiro-Wilk test statistics was performed. The results revealed that the assumption for the normality of the data was violated. Therefore a logarithmic transformation of the data was performed. Mixed design ANOVA was adopted to investigate the effect of the trial runs and training groups on the mean completion times. The results showed a significant main effect of trial run (F (4,156) = 29.609, $p < 0.001$, $\eta^2 = 0.43$); no significant interaction between group and trail run (F (8,156) = 1.328, p = 0.0.24, $\eta^2 = 0.064$) was observed. Independent sample t-tests were conducted to compare the differences between the mean completion times on the first and last trail run. The analysis revealed that there was a significant difference between the groups on the 1st run and on the 5th run which suggest that the training occurred. When investigating the differences between the groups the analysis revealed that on the 1st trail run there was a significant difference between the RO group (M = 117.9, SE = 11.7) and ATV group (M = 81.5, SE = 5.1), (t(26) = 2.845, p = 0.009, d = 1.07). No significant differences were recorded between the RO (M = 117.9, SE = 11.7) and the NC (M = 96, SE = 11.45) groups (t 26) = 2.028, p = 0.193, d = 0.51); or between the NC (M = 96, SE = 11.45) and ATV groups (M = 81.5, SE = 5.1); (t(26) = 1.155, p = 0.259, d = 0.42). On the 5th run, a significant difference was recorded between the RO group (M = 65.5, SE = 4.2) and the NC group (M = 55.7, SE = 3.2), (t (26) = 1.338, p = 0.05, d = 0.72), and between the RO (M = 65.5, SE = 4.2) and ATV (M = 54.2, SE = 4.1) groups (t(26) = 1.763, p = 0.09, d = 0.77). No significant difference on the last trail was recorded between the NC (M = 55.7, SE = 3.2) and ATV (M = 54.2, SE = 4.1) groups (t(26) = -0.098, p = 0.963, d = 0.09).

In order to compare performance between groups an exponential learning curve was fitted for each participant in all three groups as well as for each training group for overall performance across all five trial runs. As all participants performed the same task, it was assumed that they all share the same characteristics of learning pattern and learning rate. I therefore fitted a single exponential decay (learning) curve to all three data sets, with the x-offset (x_0) as the only free parameter. Fitting the same curve for each group and finding a shift between the data from different training groups means that a change in training is equivalent to faster training thus a shift in this parameter represents gains in the VR training (Figure 30). The fits show that VR training is equivalent to 0.4 real tyre change training sessions while the VR training with augmented cues provide gains that are equivalent to 0.75 real tyre change training sessions. These results suggest that learning occurred during the virtual training and was evidenced in the performance improvements on the real task. Thus it can be concluded that virtual reality systems can be used as a viable training platform for tasks that need to be performed in the real environment.

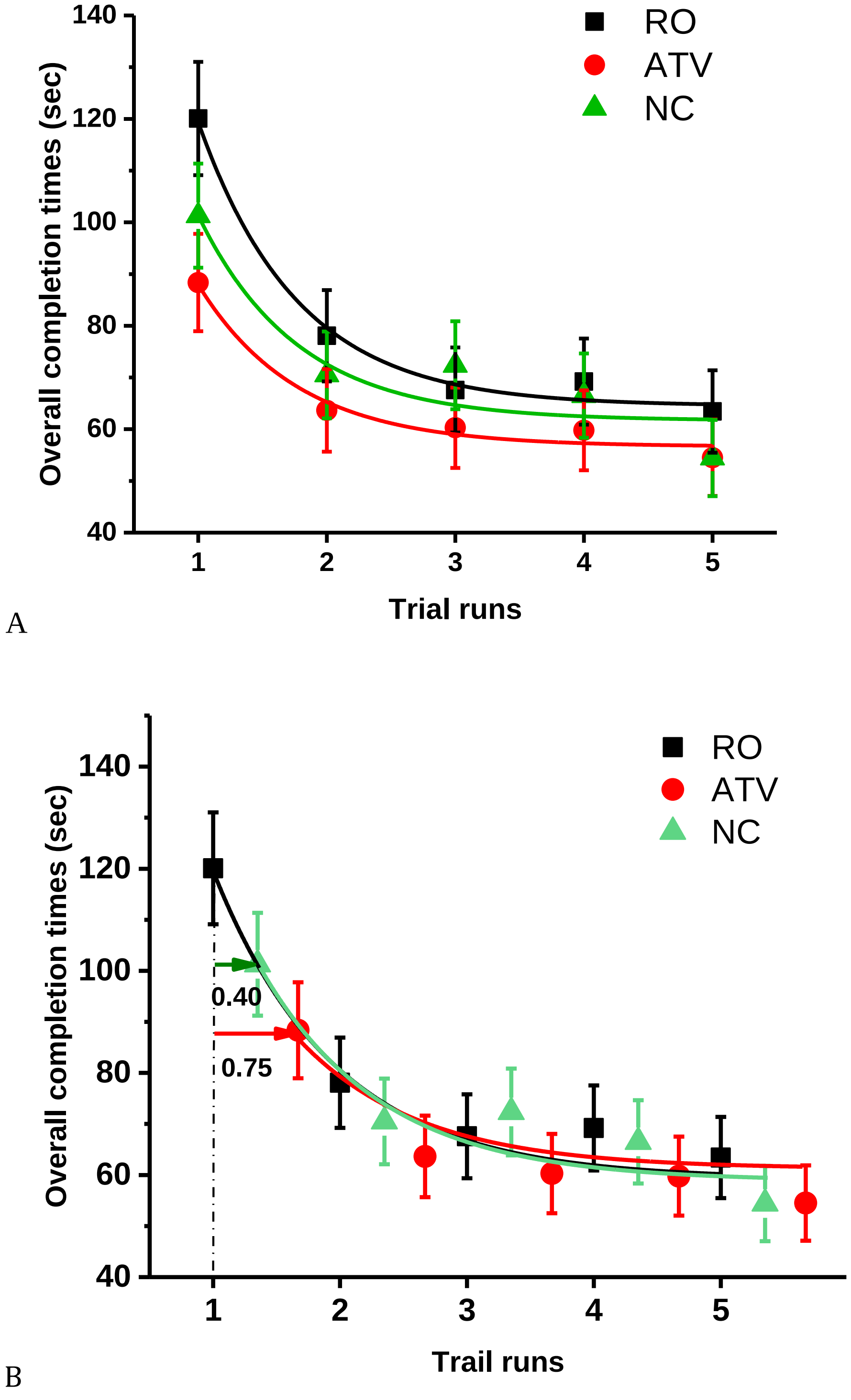


Figure 30. Mean and standard errors for each training group. (A) Learning curves with fixed parameters were fitted for each training group. (B) The X offset indicates a head start for two virtual training groups (RO = real only, ATV = VR training with additional sensory cues, NC = no cue VR training).

5.3.2.1 Task performance during virtual training

Task completion times after each trail were also collected during virtual training for two training groups. Both groups performed similarly during virtual training: mean completion time for ATV group was 87.4 seconds (SD = 9.7) and for NC group it was 90.3 seconds (SD = 18.8). Correlational analyses were performed to investigate relationship between subjective and objective measures of performance. Significant relationship between task performance and presence ratings was found in ATV training group (r = 0.446; p = 0.04) and there was a trend towards a negative significant effect between task performance and cognitive workload measure (r = -0.393; p = 0.08). No other correlations were found in the ATV group. No correlations between objective and subjective measures were found in the NC group. Correlation analyses during the real task were also performed between task performance and workload levels. Significant relationship was found in the real only group (r = 0.476; p = 0.04). No significant relationship in objective and subjective measures during the real task was found between two virtual training groups.

5.3.3 Subjective measures

Before, during and after the task participants were required to fill in a set of questionnaires to obtain subjective measures of performance. The questionnaires that were used included the NASA TXL workload scale to assess cognitive load during the task, the SSQ questionnaire to assess the levels of discomfort and sickness, and the PQ questionnaire to record participants' perceived sense of presence and immersion within the VR environment. The analysis of each subjective measure is presented in the sections below.

5.3.3.1 NASA TXL workload

The levels of workload were assessed by the NASA TXL workload questionnaire that contains five subscales. For the detailed description of the subscale please refer to Chapter 2. Participants reported their levels of workload after completing the virtual training and after completing the real task. The Real Only group reported their workload scores only during the real task. The overall mean score for workload are reported in Table 8.

Table 8

Mean NASA TXL workload scores

NASA workload scores	Real Only Mean (SE)	NC group Mean (SE)	ATV group Mean (SE)
Real task	5.3 (0.5)	6.9 (0.26)	6.4 (0.31)
Virtual training	-	4.9 (0.32)	5.6 (0.31)

Nasa TXL workload rating scale scores across all three training groups

Figure 31 shows mean workload scores in all three groups in each dimension of the NASA TXL workload scale during the real task. Overall, the RO group reported lowest mental workload (M = 3.7, SE = 0.52) and lowest frustration (M = 4.47, SE = 0.7) with the task and they also rated their own performance most favourably (M = 5.9, SE = 0.69) as compared to other two groups. The NC group reported highest mental (M = 6.2, SE = 0.63), physical (M = 8.04, SE = 0.39) and temporal (M = 8.2, SE = 0.34) workload as compared to other groups. The ATV group reported highest frustration with the task (M = 6.85, SE = 0.53) as compared to the other two groups.

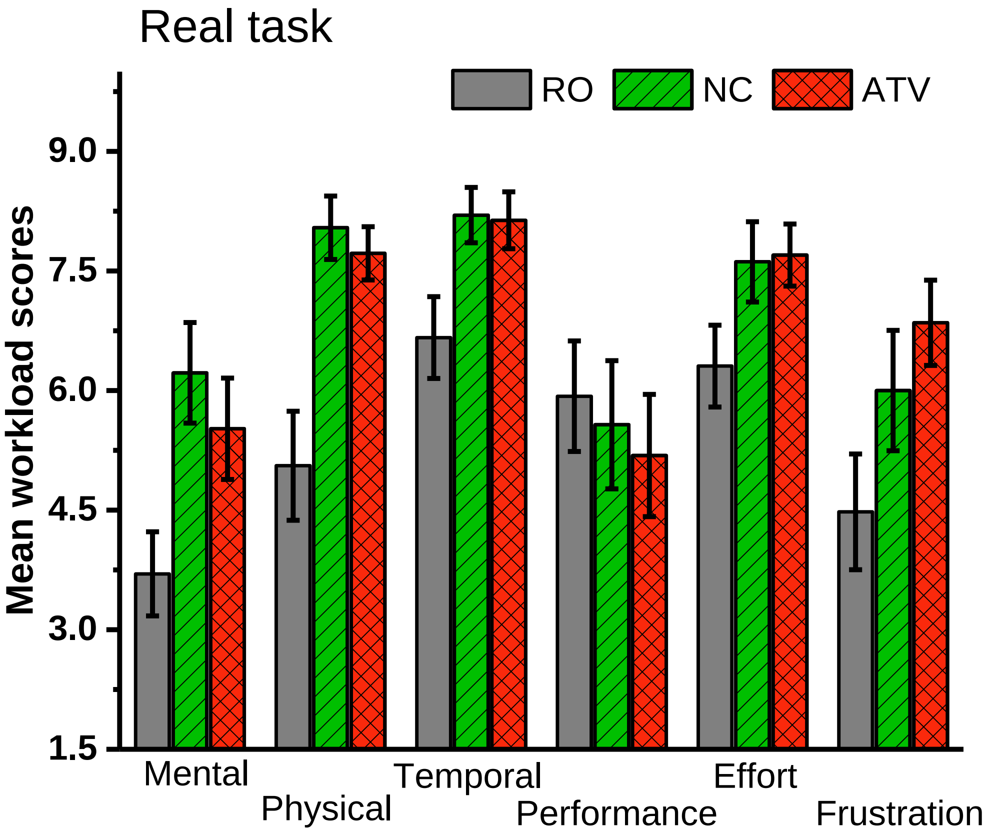


Figure 31. Mean scores on each dimension of NASA TXL workload scale for the RO group (real only), NC group (VR training with no additional sensory cues) and ATV group (VR training with augmented sensory cues).

During the virtual training the differences in workload score were also observed (Figure 32). When looking at each dimension of the workload scale, the ATV group scored higher on temporal workload (M = 7.4, SE = 10.4) and they also experienced higher frustration (M = 5.6, SE = 0.58) as compared to the NC group.

B

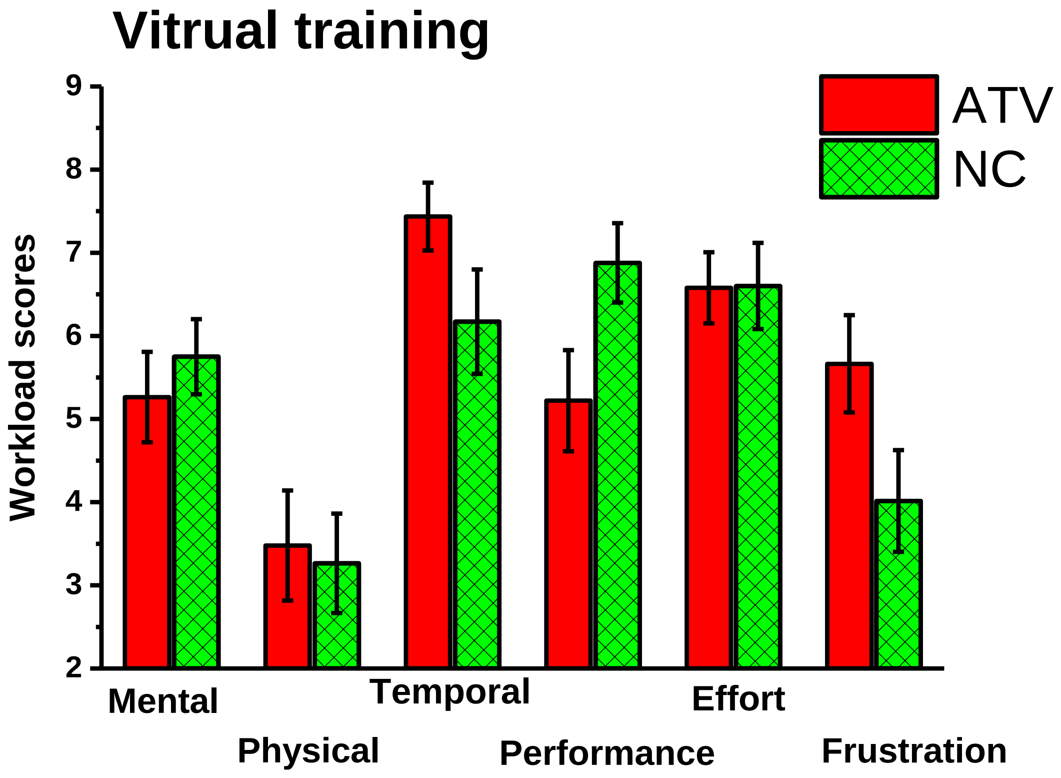


Figure 32. Mean scores on NASA TXL workload scale during virtual training for ATV and NC groups.

Independent sample t-tests were conducted on overall workload scores between the two virtual training groups (ATV and NC) to investigate whether their levels of workload differed significantly during virtual training and during the real task. During the virtual training, the overall mean score for the cognitive workload for ATV group (M = 5.6, SE = 0.31) was similar to the mean score of the NC group (M = 4.9, SE = 0.32); no significance was obtained. On the real task, the mean score for workload for the NC group (M = 6.9, SE = 0.26) was similar to the overall score reported by the ATV group (M = 6.4, SE = 0.31) and as such no significant differences between the groups and their workload levels were obtained (Figure 33).

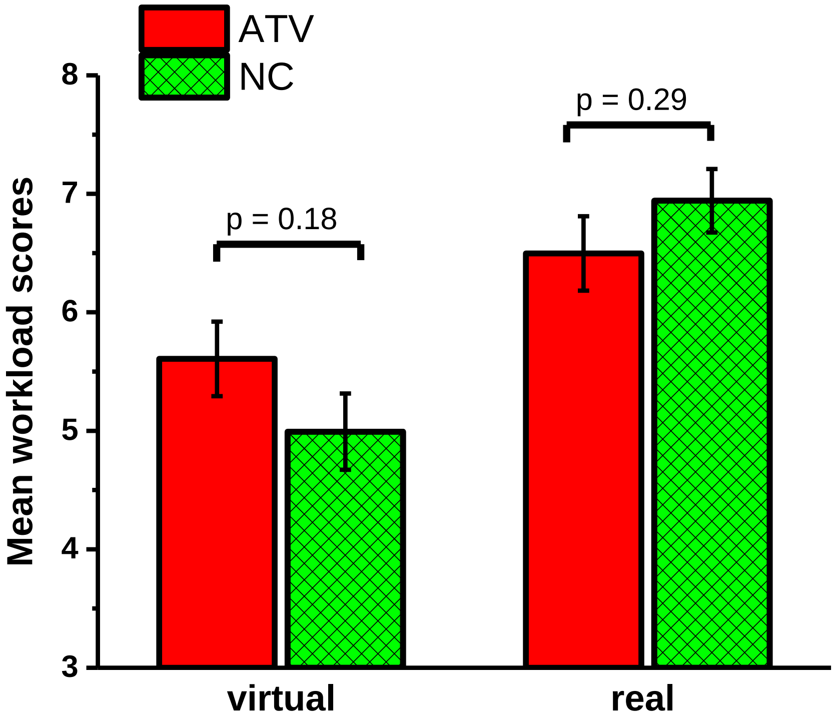


Figure 33. Mean overall workload scores for both groups on virtual and real training. No significant differences in the overall levels of workload were recorded in the virtual or real tasks.

5.3.3.2 Simulation Sickness Questionnaire

The level of discomfort and sickness were obtained from participants after they finished the wheel change task during the virtual training. No significant differences between the VR training groups existed on the pre-exposure SSQ total scores or on any of the weighted subscales, Nausea, Oculomotor and Disorientation (Figure 34). A non-parametric test Mann-Whitney U was performed to investigate whether there are significant differences between the two virtual training groups on Nausea, Oculomotor and Disorientation subscales and on total SSQ score. The analysis revealed that there was a trend towards a significant difference between the two groups on Oculomotor subscale (Z = -1.736, p = 0.083, r = 0.32). No significant difference was obtained for Nausea subscale (Z = -0.52, p = 0.958, r = 0.09) or on the total SSQ scores (Z = -1.151, p = 0.250, r = -0.21). The closer examination of the SSQ data collected before the task, after the 1st run, and after the 5th run showed that the longer the participants spend in VR, the fewer symptoms they reported. This is in line with previous research that suggested that simulation sickness decreases during long VR exposure (Faas et al., 2014).

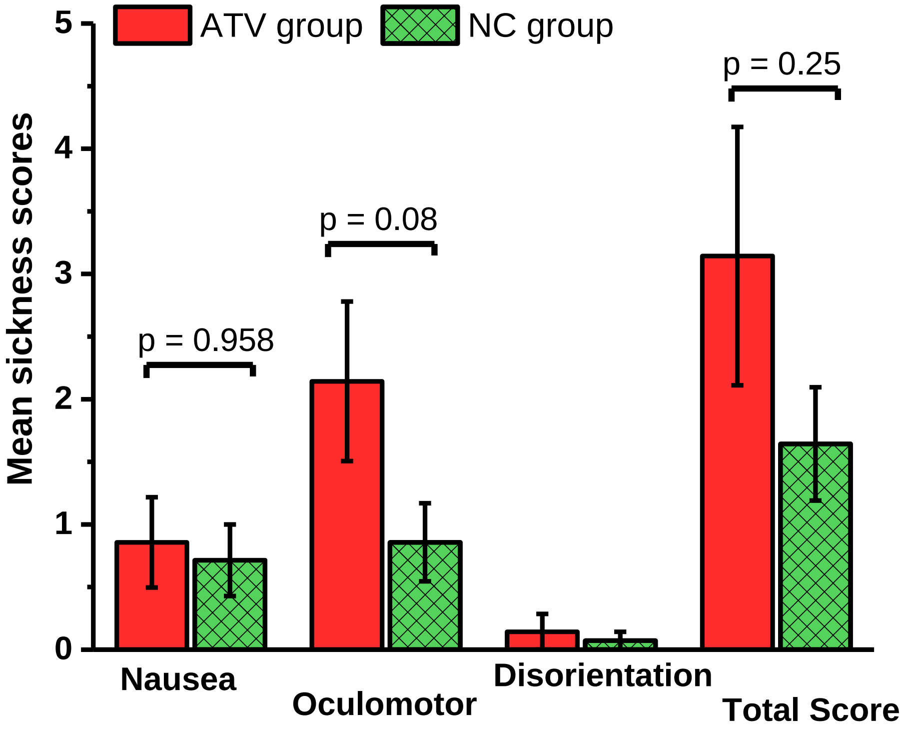


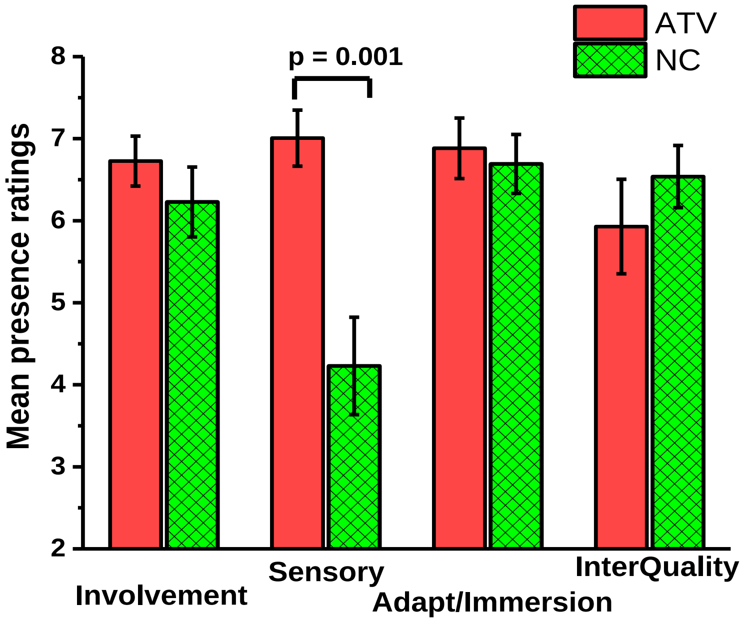
Figure 34. SSQ questionnaire subscales. Nausea, Oculomotor, Disorientation and Total SSQ score for two virtual training groups. There was a trend towards a significant difference on Oculomotor subscale, but no other differences between the two groups were statistically significant.

5.3.3.3 Presence questionnaire data

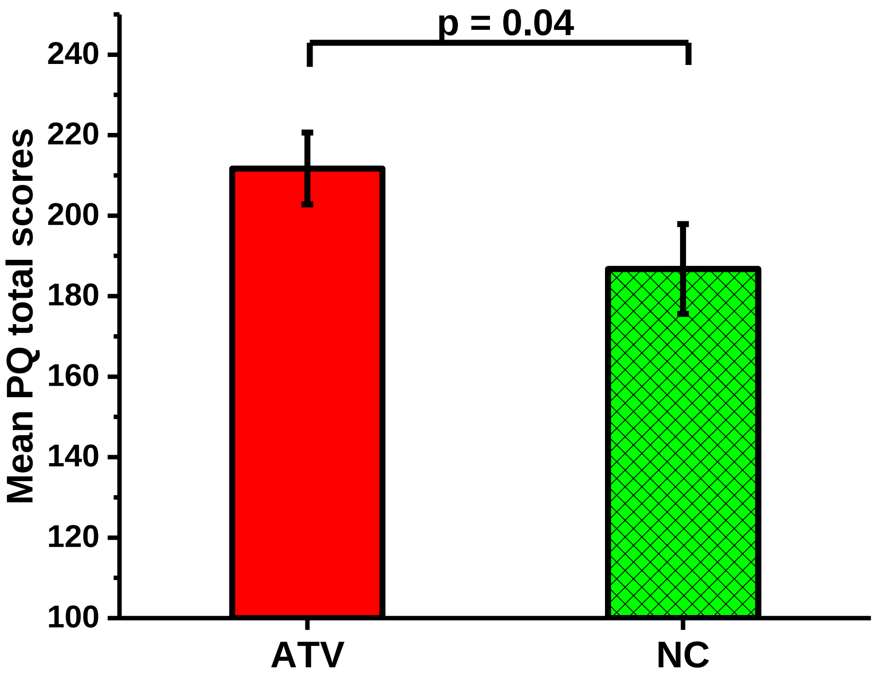
The perceived sense of presence was assessed through the PQ questionnaire that participants filled in after they performed the task in the virtual environment. The principal component analysis with direct oblimin rotation was the method used in the factor analysis. The 32 PQ items, (Witmer & Singer, 1998, Version 3.1), was subjected to principal component analysis using all PQ data collected in this study. Factors with eigenvalues greater than 1 were retained. A Keiser-Meyer Olkin (KMO) measure of 0.59 indicated moderate sampling adequacy (Field, 2013).

The detailed analysis of PQ showed that factor 1 consisted of 13 PQ items: 1, 2, 3, 4, 6, 7, 8, 9, 10, 14, 17, 18, and 29. As previously identified this factor is called Involvement/Control subscale and in this study accounted for 32.3% of variance. Factor 2 consisted of 6 items: 5, 11, 12, 13, 15 and 16. This factor was identified as Sensory subscale and accounted for 9.93% of the variance. Factor 3 consisted of 8 items: 20, 21, 24, 25, 28, 30, 31 and 32. This factor relates to immersion and therefore is called Adaptation/Immersion subscale that accounts for 8.5% of the variance. Factor 4 consisted of 3 items: 19, 22 and 23 which refers to Interface Quality subscale and accounted for 5.8% of the variance. Items 26 and 27 had negative loading and therefore were excluded from further analysis. All other items corresponded almost exactly with the items extracted in the previous work of factor analysis performed by Witmer and Singer, (1998, 2005). Item 28, 29 and 32 were not extracted in the previous work; item 28 refers to involvement in the experimental task; item 29 refers to physical interaction with the objects in VR; items 32 refers to information provided through different senses (vision, hearing, touch). As all three items relate directly to the experimental questions regarding the interaction in VR with multisensory feedback, all were retained and included in the further analysis.

A mixed design ANOVA was performed on the subscales of the PQ questionnaire data and indicated that there was a significant difference between the subscales ratings ($F(3,78) = 4.891$, $p = 0.012$, $\eta^2 = 0.158$) and there was also a significant interaction between the PQ subscales and VR training groups ($F(3,78) = 10.509$, $p < 0.001$, $\eta^2 = 0.29$). Planned comparisons with corrected significance value (Sidak) showed that there was a significant difference on sensory subscale ($t(26) = 4.053$, $p < 0.001$, $d = 1.55$). No other significant differences were observed. Independent sample t-test was performed on the overall PQ scores and showed that there was a significant difference in the overall levels of presence between two virtual training groups ($t(26) = 1.747$, $p = 0.045$, $d = 0.66$) (Figure 35)This finding indicates that the group that received virtual training with augmented multisensory feedback (ATV) experienced a significantly higher sense of presence than the group that received the training in conventional VR (NC).



A



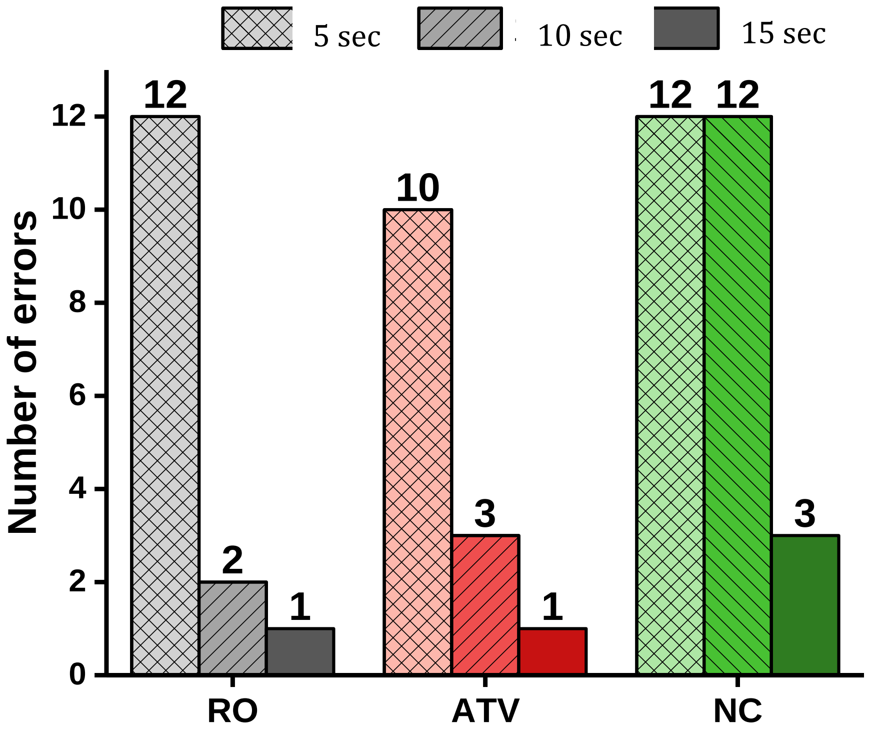
B

Figure 35. Mean scores of PQ subscales (A) and PQ total scores (B). Independent sample t-test showed a significant difference in PQ total scores (p = 0.04, one tailed) and on Sensory subscale (p = 0.001) between two virtual training groups. Error bars represent standard error of the mean.

5.3.4 Analysis of errors

During the virtual task the augmented cues signalled when the bolts were fully inserted by either a change of colour, a higher intensity vibrations or a drilling sound. Although the vibration and the drilling sound were also present during the real task there was no direct way how to check whether the bolts were screwed in correctly. In a real life scenario, if a bolt is correctly screwed in, it cannot be removed by hand only by an impact wrench. When a bolt is not screwed in properly i.e. the bolt is not lined up properly on to the nut, the bolt can either be easily unscrewed by hand or gets stuck on the ridges. During the real task in this study, when a bolt was not fasten properly to the wheel, i.e. the bolt could be unscrewed by hand, a five second penalty was added to the overall time. The number of participants who received a penalty and the amount of penalties given to participants were recorded for each experimental group (Figure 36).

Chi-square analysis (Chi^2) on the total number of errors in all three groups was carried out and found to be highly significant ($x^2(2) = 14.776$, $p = 0.001$). The result means that there was a significant difference between the amounts of penalties given between the three experimental groups. Participants in the ATV training group received the least amount of penalties. Participants in the NC virtual training group that received the VR training without any additional augmented cues, performed significantly more errors during the real task than the other two groups. The number of participants that were given the penalties was also highest in the NC group. The ATV training group received the least amount of penalties, which suggests that the information that was available to the participants during the virtual training aided their attention to detail, which is necessary in order to perform the task as fast and as effectively as possible.



	Overall number of participant in the group	Number of participant that made errors	Number of error penalties during the real task
RO group	14	9	15
ATV group	14	10	14
NC group	14	12	27

Figure 36. Penalty distribution on the real task across all three groups. The table shows how many subjects in each group, how many got erro penalties and the number of penalties given in each group. The Chi square (Chi^2) analysis was used to test for significant differences between the groups in the total number of error penalties given. Overall, the NC group performed significantly more errors as compared to the other two groups (RO and ATV).

5.4 Discussion

The main aim of this study was to investigate whether the learning transfer occurs from the virtual training task with augmented sensory cues to the same task in the real environment. Three groups of participants performed the wheel change task against the clock on the real racing car. Two groups received virtual training on the same task: one group was trained using a conventional VR set up whilst the other group was trained using the same VR set up with the addition of augmented sensory cues that were present during the task and that carried task-relevant information. The performance between the three groups was compared through objective and subjective measures of performance.

5.4.1 Virtual training and task performance

The findings obtained by the examination of the objective task performance data showed that participants' performance improved across the trials, which suggest that VR training improved performance on the real task; both VR training groups performed better on the real task than the real only training group. In the previous study, Loftin et al. (2004) trained individuals to perform a military checkpoint duty in fully immersive VR and compared the performance to the desktop-based scenarios. They found that overall levels of performance were better in fully immersive VR. In other studies, VR training has also improved the level of efficiency during surgical training (Ahlberg, Enochson, Gallagher, Hedman … & Arvidsson, 2007); the positive effects of VR training were evident even after two weeks (Carlson, Peters, Gilbert, Vance & Luse, 2015). The findings from this study confirm previous research and suggest that virtual reality can be used as an effective training platform for tasks that can then be performed in real life situations.

While no significant difference was observed between the two training groups on mean completion times during the virtual or real task, the group that perform training with minimalistic cues (NC group) received significantly more error penalties on the real task than the group that received training with augmented sensory cues (ATV group). The analysis of error rates revealed that there were significant differences in the amount of error penalties between the groups during the real task; in total 85% of participants in the NC group made at least one error compared to 63% in the ATV group and 70% in the RO group. During the real task, the NC group performed on average 56% more errors than the RO and ATV groups. This finding might be explained by the amount of acquired penalties, which could influence their levels of cognitive load as they tried even harder to perform the task without any errors. The participants in the ATV group also acquired penalties during the real task, however the amount of penalties was much lower (as a group they collected 15 error penalties) when compared to NC group (who collected 27 error penalties). This suggests that virtual training with additional sensory cues can provide a more efficient preparation, in terms of a higher attention to detail, for real life scenarios as oppose to virtual training without the additional cues. This finding is of particular importance as previous research has highlighted that human mistakes during manual assembly operations can cause serious and often costly consequences for the company and its environment (Kern & Refflinghaus, 2013). Therefore, virtual training with augmented sensory feedback is recommended as it can serve as an optimal approach for early detection and subsequent reduction of human errors. A link between simulation fidelity and training outcome has been identified in previous research (Hays et al., 1992); here I show that additional cues, that might reduce the overall fidelity but carry task-relevant information, can, in fact, improve performance.

5.4.2 User experience in virtual and real environment

In the analysis of user’s acceptance of the virtual environment, a significant difference in the perceived sense of presence between VR training groups was found. Although the presence of the augmented cues decreased the overall fidelity of the environment they still enhanced the sense of being within the virtual world. Previous research has suggested a possible relationship between the sense of presence and performance. For example, Faas et al. (2014) showed that higher levels of presence correlated with higher levels of performance during a manual assembly task. Steven and Kincaid (2015) investigated whether higher presence in virtual simulation training was associated with a higher sense of presence; their findings showed that higher visual immersion had positive effects on the sense of presence and on performance. In this study, the visual cues were highly unrealistic but although they did not significantly influence the mean task performance, they significantly increased the user's sense of presence; this is most likely due to the fact that they carried task-relevant information. Laha et al. (2012) showed that not all components of a virtual set up that influence immersion are necessary for efficient performance. Users in their study showed mixed preferences for different levels of immersion; their results revealed that search tasks benefited the most from the high levels of immersion, but not general descriptive tasks. The results from this study shown that even when the fidelity of the environment is decreased by the presentation of additional low-fidelity sensory cues, it can still enhance not only the overall task performance but also user perception and acceptability of the virtual environment.

5.4.2.1 Presence and immersion

To investigate the user's acceptability of the virtual environment, subjective measures of performance were collected before, during and after the virtual and real tasks. The levels of involvement, immersion and presence were recorded on the Presence Questionnaire; the results showed that the ATV group experienced a higher sense of presence despite the fact that the presented cues decreased the overall fidelity of the virtual environment. When investigating the individual subscales of the PQ, the ATV group rated Involvement, Sensory and Adaptation/Immersion subscales higher that the NC group; a significant difference between the two VR training groups was recorded on the sensory subscales which was expected. The NC group also experienced increased Interface quality as compared to the ATV group. This finding suggests that presented augmented sensory cues facilitated the levels of involvement and presence during the virtual task, which is in line with previous research (Dinh et al., 1999; Ma & Kaber, 2006).

5.4.2.2 Cognitive workload

The levels of cognitive load during the exposure to the VR environment were recorded on the NASA TXL workload scale. Previous research has found that within the virtual environment, the task complexity and the experience level can significantly influenced the levels of cognitive workload (Leung et al., 2010). Similarly, the length of the virtual training was also shown to influence the levels of cognitive workload when higher workload was reported during the short training as oppose to the long training (Singh et al., 2005). On the other hand, Stone et al. (2011) evaluated the VR training in terms of physical and cognitive impact and found that the levels of cognitive load did not differ significantly between the virtual training tasks or groups.

In this study, the virtual training groups reported higher workload levels during the real task than the control group (RO group) however the differences failed to reach significance which supports previous research (Stone et al., 2011). The main reason for this could be explained by the fatigues as both groups had already performed the virtual training earlier (Kennedy et al., 1993). When comparing the workload levels between the two virtual training groups during the virtual training, again no significant differences were recorded. The ATV group reported a slightly increased temporal and physical workload and a higher level of frustration during the virtual training but during the real task their levels of cognitive workload was consistently lower as compared to VR training group with no additional sensory cues. The NC group reported lower levels of workload during the virtual training, but systematically reported increased mental, temporal and physical workload levels. One possible explanation for the increased levels of workload in NC group during the real task could be due to the amount of error penalties that they received. The task was performed against the clock, participants tried to do the task faster and as they were not cued to direct their attention to possible errors they performed more mistakes, such as failing to tighten the bolt properly, which made them experience higher levels of frustration and thus increased their levels of cognitive workload (Hart & Staveland, 1988). It seems that the presence of multisensory feedback during VR training increased cognitive workload, however the beneficial effects of sensory cuing were seen during the real task because participants learned to direct their attention to error prone events.

This finding is in a slight disagreement with cognitive load theory (Sweller, 1994). The cognitive load theory proposes that the design of instructions during the task may impose cognitive load on participants' working memory which is limited during high workload and as a results the cognitive load might affect learning outcomes (Cerniak, Scheiter & Gerjets, 2012). In this study, the availability of augmented sensory cues caused a high workload situation during virtual training as participants reported higher levels of workload. However, during the real task this cue-based learning facilitated overall performance as participants paid more attention to learned cues and as a result performed fewer errors during the real task. This suggests that a certain amount of cognitive workload or alertness during the training task could be necessary for participants to accurately attain and later recall the learned information i.e. paying increased attention to the tightness of the bolt in order to perform the task correctly. Due to these findings, it is recommended that future research concentrates on the further evaluation and investigation of the effects of high and low cognitive workload and multisensory cuing on user acceptability of virtual environments.

5.4.3 Relationship between objective and subjective measures

Correlational analyses revealed that the relationship exists between objective measures of performance and users' subjective perceptions of VR environments. The task performance improved when participants experienced increased sense of presence. At the same time, there was a trend towards negative significant relationship between task performance and cognitive workload, which suggest that when participants experience increased levels of cognitive workload their task performance also improved. This suggests that medium levels of workload are required to keep user's attention and focus on the task in order to perform the task efficiently (Jackson, Kleitman & Aidman, 2014).

5.4.4 Limitations of the study

Despite using a state-of-the-art virtual reality set up some interference still occurred. During the task virtual tracking of the real objects became problematic which could have contributed to the increased levels of workload and frustration from the participants' point of view. The interferences could also contribute to the break in immersion and thus the rating on the subjective metrics could have been affected.

Future research could therefore ensure the use of an effective, fully calibrated motion tracking system to avoid these interferences.

Another limitation of this study is the reliability of the error penalties. The performance on the real task has been assessed by the recording of the overall time and error penalties, which were collected by the experimenter. However, this assessment lacks accuracy. Future research should therefore make use of more robust technology that can be used to assess the accuracy of the procedure (Lin, Pang, Zhang & Feng, 2014).

Chapter 6

The effects of motion cuing on fidelity assessments in simulated flight task.

Chapter overview

Although this chapter moves away from investigating primary sensory cues (visual, audio, tactile) when investigating the role of multisensory cues, motion cuing also needs to be considered. The empirical research studies that are conducted to investigate the effects of motion cuing in great detail emphasise the importance of this sensory cue on fidelity assessments. As the assessment of fidelity is the main theme of this thesis, the study investigating the role of motion cuing on fidelity assessments was also conducted. This chapter covers findings from a study that was designed to investigate the effect of motion cuing on user behaviour in terms of performance and subjective perception using 6 degrees of freedom (DOF) helicopter simulator. The introduction will cover the role of the motion platform in aviation and discuss techniques used during the fidelity assessments of VR environments. The experimental aims and results will be described subsequently and the chapter will conclude with the discussion of findings.

6.1 Introduction

6.1.1 Motion platform in aviation

Flight training is the most frequent application in which motion platforms are used primarily to reduce cost and increase safety (Schroeder, 1999). In order to provide a realistic flight experience, many flight simulators provide motion cues to the pilot. A considerable amount of research has been conducted to investigate whether the motion platforms, given their high cost and maintenance, are necessary for training effectiveness (McCauley, 2006). Despite the complexity and the use of state of the art components in modern simulators, it has been noted that motion platforms are not yet able to provide a fully coherent representation of reality (Perfect et al., 2010, White et al., 2013). The empirical research investigating the use of motion platforms has produced mixed results in terms of attainable benefits obtained from motion cuing. The presence of motion feedback was shown to influence the levels of simulation sickness (Aykent et al., 2011; Stein & Robinsky, 2012; Kluver et al., 2016), cognitive workload (Bell & Grant, 2011; Pasma et al., 2011) as well as the sense of presence (Steven & Kincaid, 2015; Milleville-Pennel & Charron, 2015). For the detailed description of the research studies that investigate the effects of motion cuing on user behaviour and performance please see Chapter 2.

6.1.2 High and low cognitive workload

Whilst within aviation most of the research on workload concentrated mostly only on high demand conditions, others have noted that the cost associated with low demand conditions has not been fully explored yet (Hancock, Williams & Manning, 1995). It was the work of Yeng and Wickens (1988) who first suggested that workload and performance dissociate from each other. This means that under certain conditions, performance could improve as the workload increases and vice versa. This assumption prompted a considerable amount of empirical research conducted into the effects of workload levels on performance. Hancock et al. (1995) supported this notion: participants were asked to perform a simulated flight task on a computer screen during three workload conditions. Their results showed that performance was better in high demand conditions as compared to the low demand conditions. Similarly, Dunn and Wiliamson (2012) investigated the role of task demand in monotony related tasks with train drivers and found a superior performance with far less errors during the high demand condition as compared to the low demand condition. Informed by the abovementioned studies, I have decided to investigate three different levels of perceived task demand conditions in order to determine whether the motion cuing and cognitive workload will influence overall task performance and user perception of VR.

6.1.3 Summary of experimental aims

Previous research has produced conflicting results in terms of the effectiveness of motion platform for training. Therefore further investigation into the effects of motion cuing is necessary. The main aim of this study was to investigate the effects of motion cuing on objective task performance and the subjective measures of the user's psychological state. The psychological measures of perceived feeling of presence, cognitive workload and simulation sickness were collected before, during and after the various helicopter manoeuvres. Two experimental blocks were compared during the simulated slalom task where motion feedback was either present or absent and the perceived task demand was manipulated. Qualitative and quantitative measures of user experience, including presence, sickness and workload were collected during the experiment. It was predicted that the presence of motion feedback will enhance the accuracy of performance and subjective feeling of presence and sickness. It was also predicted that motion cuing together with increased task demand will modulate user task strategies and negatively influence the levels of cognitive workload.

6.2 Methods

6.2.1 Participants

A power analysis (Faul et al., 2009) was performed to determine a sample size required for this study and showed that for 0.05 level of significance, the effect size of 0.5 and power (1 – β err prob) of 0.5 total sample size required was 10. For this study, 11 participants were recruited via opportunity sampling thus satisfying the required sample size. There were 10 males and 1 female with the age ranging from 21 and 70 (M = 33.8, SD = 17.7). All participants had previous experience with flying either in a fixed-wing or rotorcraft simulator. The participants' cohort consisted of 3 professional test pilots and 5 undergraduate students and 3 members of staff from the School of Engineering at the University of Liverpool. All participants signed a consent form and reported normal or corrected-to-normal vision and hearing.

6.2.2 Virtual reality set up

The task was performed in the helicopter simulator HELIFLIGHT-R situated at the University of Liverpool. This facility was installed in 2011 in order to encourage continued expansion in research and teaching capabilities. The HELIFLIGHT–R is a 12-foot tall visual dome mounted on the six-degrees-of-freedom motion platform (Figure 37). The system utilizes general-purpose Linux-based computers to drive the simulator from a central Instructor–Operator Station (IOS) PC. The IOS PC is connected to a local network that allows communication with each of the other elements of the system – three image generation (IG) machines that produce the visual environment, one machine to run the reconfigurable instrument panel displays (left and right primary flight displays, backup analogue displays and Head

Up Display), and a machine for the Instructor Station within the dome, which serves a dual purpose by creating the audio environment. In addition, the network is connected to the control interfaces for the control loading and motion systems (White et al., 2012).

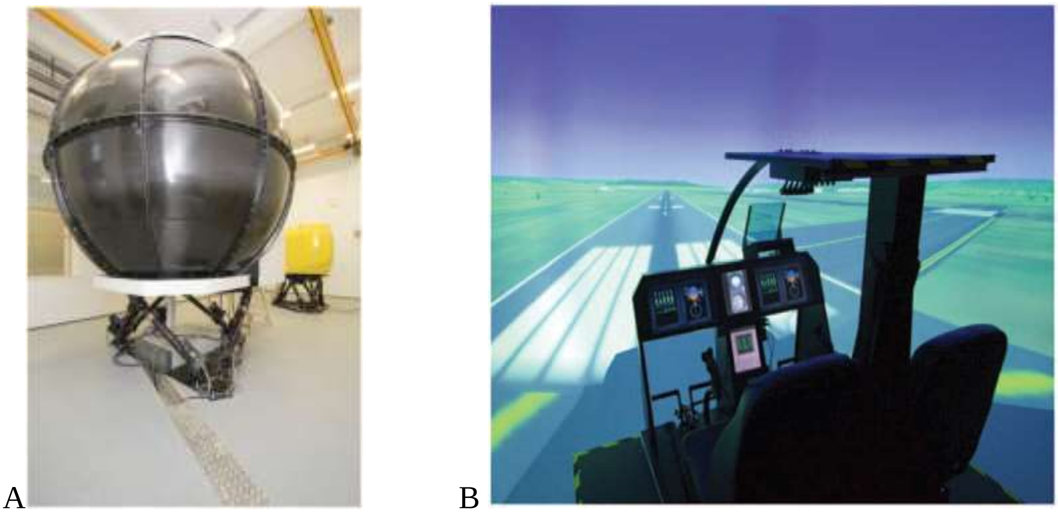


Figure 37. HELIFLIGHT-R simulator (A) and crew station inside the simulator (B)

Table 9.

HELIFLIGHT –R motion envelope

	Displacement	Velocity	Acceleration
Pitch	-23.3 /25.6	-+34/s	300/s^2
Roll	-23.2	-+35/s	300/s^2
Yaw	-+24.3	-+36/s	500/s^2
Heave	-+0.39 m	-+0.7 m/s	-+1.02 g
Surge	-0.46/+0.57 m	-+0.7 m/s	-+0.71 g
Sway	-+0.47 m	-+0.7 m/s	-+0.71 g

Inside the simulator: The crew station inside the simulator is equipped with two wide-screen 1900 LCDs to represent the primary flight information, engine information and navigation information. The crew station uses a four-axis (longitudinal and lateral cyclic, collective and pedals) Moog FCS ECoL 8000 Q&C-Line electric control loading system that back-drives the pilots' controls and allows fully programmable force-feel characteristics. The conventional rotorcraft controls can be replaced with an F-16-style side stick and throttle for fixed-wing operations. The dome is also equipped with an Instructor Station, which can, for fixed-base operation, control all simulator functionality. A head-up-display unit is available and uses a 1000 4:3 LCD screen with a beam splitter and is mounted on the glare shield on the right-hand side of the cockpit. An in-cockpit camera is installed on the left of the rear tower to provide a live display of the pilot and co-pilot to the IOS (Figure 37 B) (White et al., 2012).

Motion: Motion cueing is provided by a Moog MB/E/6dof/ 24/1800 kg electric motion system, consisting of six Moog electric actuators arranged in a hexapod structure to provide full six-degree-of-freedom motion. Each actuator has a 600mm stroke, giving peak accelerations of > 300_ /s2 in each rotational axis, 0.71g in surge and sway, and 1.02g in heave (see Table 9, values are given for single axis motions). The platform has a 1800kg loading capacity with the estimated weight of the cockpit being 900kg (White et al., 2012).

Simulator visuals: Three Canon SX60 projectors, with a 1400 x 1050 resolution, equipped with wide-angle lenses provide a wide field of view of 210, horizontally by . 30_ /_40_vertically. This results a 3.43 arc-min/pixel resolution, which is very close to the JAR STD-1H level-D visual requirement of 3 arc-min/pixel. The field of view represents a significant increase in capability compared with that available on the original HELIFLIGHT system. The Liquid Crystal on Silicon technology used in the projectors allows a high-quality visual display without the pixel gridding seen with LCD projectors. A set of Silicon-Optix-Image-AnyPlace Video Scalers is used to warp and edge blend the three Out The-Window images into one scene. The IG is provided using Boeing’s Multi-Purpose Viewer, an Open Scene

Graph based tool that supports rendering of any Open-Flight terrain or object database (White et al., 2012).

Control room: ART's FLIGHTLAB modeling and simulation software 22 is at the centre of operation at the new facility. FLIGHTLAB provides a modular approach to developing flight dynamics models, producing a complete vehicle systems model from a library of predefined components. A number of GUIs: Xanalysis, CSGE/GSCOPE and FLIGHTLAB Model Editor are available to aid in the generation and analysis of flight models (White et al., 2012).

6.2.3 Task procedure

Before the task, participants were briefed on how to fill in questionnaires. After they filled in first sets of questionnaires to provide a baseline measures, thy were taken inside the flight simulator and asked to sit down, fasten the seatbelts, put on the headphone in order to communicate with the control room and wait for the instruction to start the task. Each participant had two practice runs to familiarise them with the simulation. After this the experimental task began by flying each task demand condition three times, which was followed by filling in their workload levels. After each condition was flown in one of the experimental blocks (either Motion On or Off), they were taken out of the simulator to have a break and fill in set of questionnaires. After 30 minutes they went back into the simulator to complete another set of task demand conditions in other experimental block (Motion On or Off) and filling in their workload scores after flying each task demand condition three times. When the experimental block was finished, participants were taken out of the simulator back into the control room where they filled in the last set of the questionnaires, were debriefed and thanked for their participation.

The increase and decrease in perceived task demand was achieved by manipulating the pole width separation (PWS) in three conditions:

a) poles were 10 meters apart, which corresponded to high task demand condition,

b) poles were 20 meters apart, which corresponded to medium task demand condition and

c) poles were 25 meters apart, which corresponded to low task demand condition (Figure 36).

Each task demand condition was flown three times and after this pilots recorded their workload scores.

In order to perform the task to the assigned standards (Baskett, ADS33, 2000) participants had to follow a predesigned trajectory pattern (grey line on the Figure 38) that was shown as a yellow centre line going in between the poles in the virtual scene. The ADS33 (2000) states that a series of smooth turns must be performed at least 50ft from the runway centreline, at 500ft intervals and at least twice to each side of the course. To ensure the stability of the aircraft, which could affect performance of the novice pilots, only the horizontal movements were made available: the airspeed of the helicopter was fixed at 40kt and so was the altitude of the aircraft, which was at the height of 40ft. The movements were controlled by the collective stick only, which meant that the pedals did not produce any effects on aircraft dynamics. The slalom task manoeuvre was extended to consist of seven sets of poles in order to allow participants to adjust to the task. All together 20 turns were performed, 10 on each side of the runway, to optimise the reliability of the data set gathered.

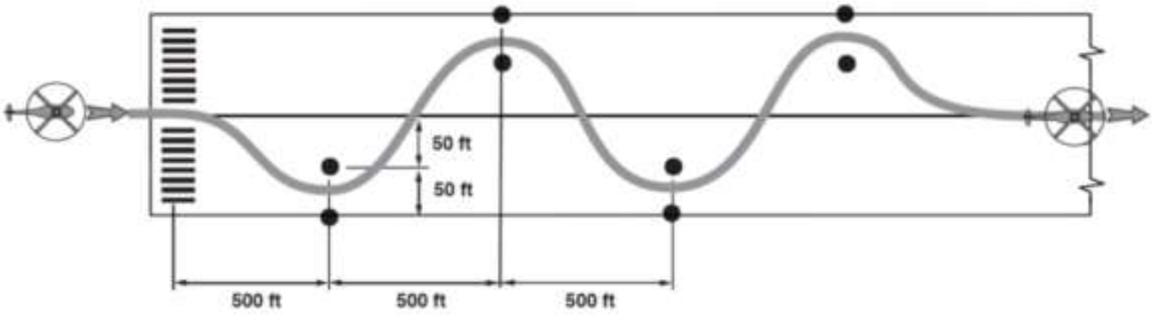


Figure 38. Slalom task diagram. The grey line represents the desired trajectory pattern. The slalom task manoeuvre was extended to seven sets of poles in order to gather enough data and give participants enough time to adjust to the task.

The motion cuing conditions and the task demand conditions were randomised across the participants. The example of the randomised order can be seen in the Table 10. Each experimental block lasted about 30 minutes and participants had at least 30 minutes break between the experimental blocks.

Table 10

Example of randomised order for Slalom task

	Motion	Pole separation (m)		
Participant 1	ON	10	25	15
	OFF	25	15	10
Participant 2	OFF	15	10	25
	ON	25	15	10
Participant 3	ON	15	10	25
	OFF	10	25	15
Participant 4	OFF	25	10	15
…	ON	15	25	10

6.2.4 Methodology

This experiment is based on the Fitts law that predicts that the time, T, to point to a target of width W at a distance A is logarithmically related to the inverse of the spatial relative error A/W. This is represented by equation where a and b are empirical constants:

(Equation 5)

$$T = a + blog_2\ (2A/w)$$

In accordance with Fitts law (Fitts, 1954) it is predicted that by reducing pole separation, x, (width of target) the reduction in the task performance will be more evident even though the actual difficulty of the task will not change. Based on the hypothesis that the pilots tune their level of arousal according to how they perceived the difficulty of the task, it is to be expected that the participants should become more aroused in the activity, which in turn will result in the higher levels of immersion. It has been found before that the level of arousal depends largely on training and experience (Roscoe & Ellis, 1990). The experiment will be performed in two experimental blocks: one where motion cuing is provided and one where no motion cuing is present. Previous studies have shown that the use of motion feedback can contribute to the increase in the task performance (Lu & Jump, 2014). I predict the participants will feel a larger sense of presence and immersion with motion on as oppose to off.

6.2.5 Statistical analysis

For the analysis of the performance data a factorial repeated measures design (ANOVA) with experimental factors ‘motion’ (2) and ‘task demand’ (3) was computed for the mean error data set. Partial eta squared is reported as an effect size (Field, 2013). As it was predicted that performance will be better when motion feedback is provided, planned comparisons were carried out in each task demand condition, across two motion conditions. Cohen’s d was chosen as a measure of effect size (see Equation 1 and 2) (Cohen, 1992). When multiple comparisons were performed an adjusted Bonferroni correction was calculated to provide an adjusted significance value (p value).

When the data did not follow the general assumptions for the parametric test, non-parametric tests were used. These were non-parametric Freidman ANOVA test where Kendall’s Coefficient of Concordance (W) was used as a measure of effect size (Field, 2013). When a-priory assumptions were made, planned comparisons using a Wilcoxon sign rank test were performed with the calculation of an appropriate effect size (see Equation 3) (Pallant, 2013).

6.3 Results

6.3.1 Objective data

In this study participants were required to perform a slalom task manoeuvre in the helicopter flight simulator whilst the perceived task demand was manipulated. The task was flown in two experimental blocks where motion cuing was either present or absent. As two participants did not complete both experimental conditions, data form 9 participants were included in the analysis. Participants were required to follow the centre line on the ground between the poles in order to successfully pass between the poles to perform an efficient slalom task manoeuvre (Baskett, ADS33, 2000). Deviations from the desired target route were recorded as errors. Mean errors in each task demand condition for each experimental block were computed and used for the analysis. A factorial repeated measures ANOVA was used for the analysis and showed a significant interaction between the motion and task demand (F(2,15)=

4.066, $p = 0.03$, $\eta^2 = 0.34$). No significant effect of motion ($p = 0.619$, $\eta^2 = 0.03$) or task demand ($p = 0.472$, $\eta^2 = 0.09$) was observed.

As it was predicted that the absence of the motion cues with decrease performance, pairwise comparisons between each task demand condition (10m, 20m, 25m) were conducted in both experimental blocks (Motion On and Motion Off). The results revealed a significant effect of motion cuing in the 25m condition ($t_{(8)} = 2.02$, $p = 0.035$, $d = 0.84$), but not in the 10m ($t_{(8)} = -0.467$, $p = 0.653$, $d = 0.11$) or 20m conditions ($t_{(8)} = -0.959$, $p = 0.366$, $d = 0.23$). Figure 39 shows that when perceived task demand was increased, the task accuracy was decreased. The motion cues facilitated the amount of error in the 10m and 20m conditions, but not in the 25m condition.

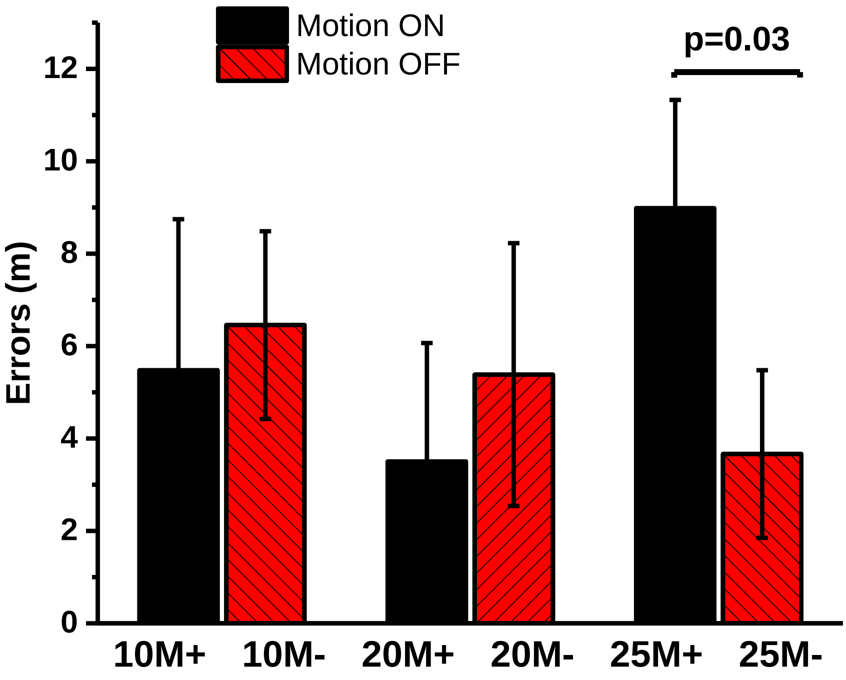


Figure 39. Means and standard errors of deviations from desired trajectory in all conditions with and without motion cuing. Error bars represent standard error.

To assess the magnitude of the deviation from the trajectory pattern across the whole data set, error data for each gate (7) were computed and subsequent analysis was conducted. Repeated measures ANOVA revealed a significant effect of gate (F(6,120) = 6.509, p = 0.0001, η^2 = 0.246), but no significant effect of motion (F(1,20) = 0.033, p = 0.85, η^2 = 0.002) (Figure 40). No significant interaction was recorded between the number of gates and motion cuing.

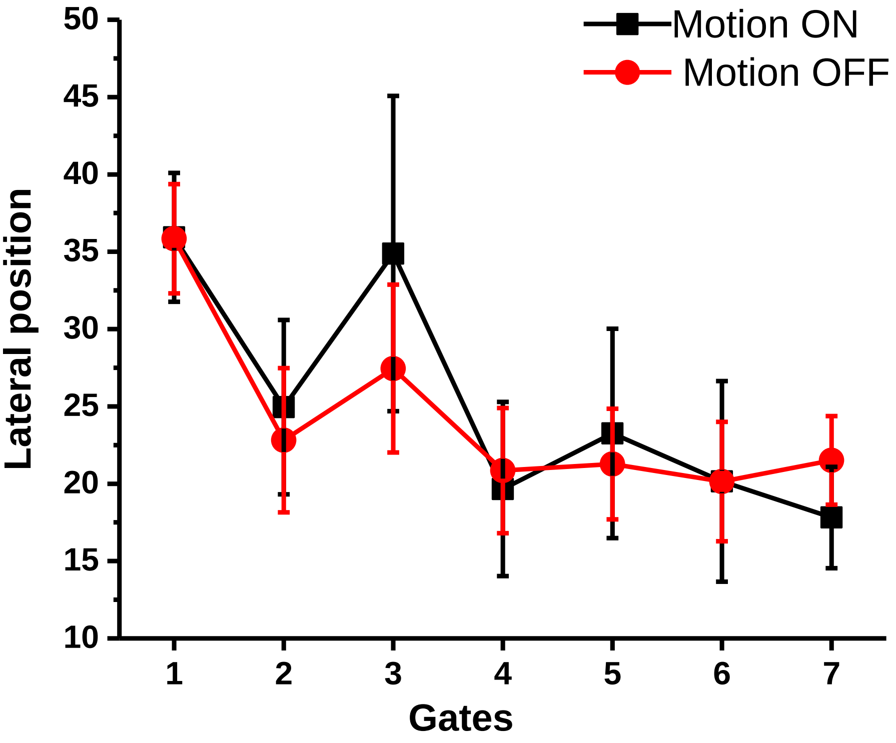
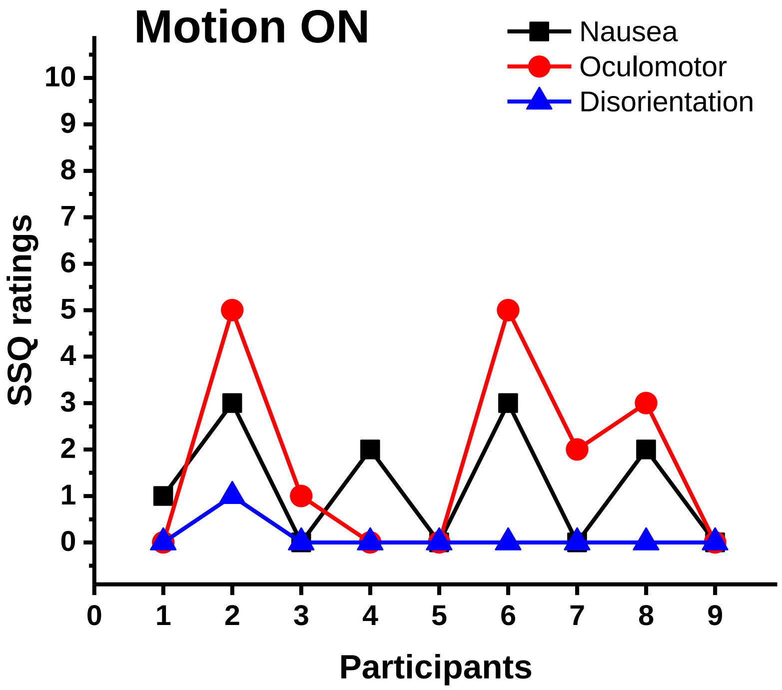


Figure 40. Means and standard errors of deviations from the trajectory pattern across all seven gates with and without the motion feedback.

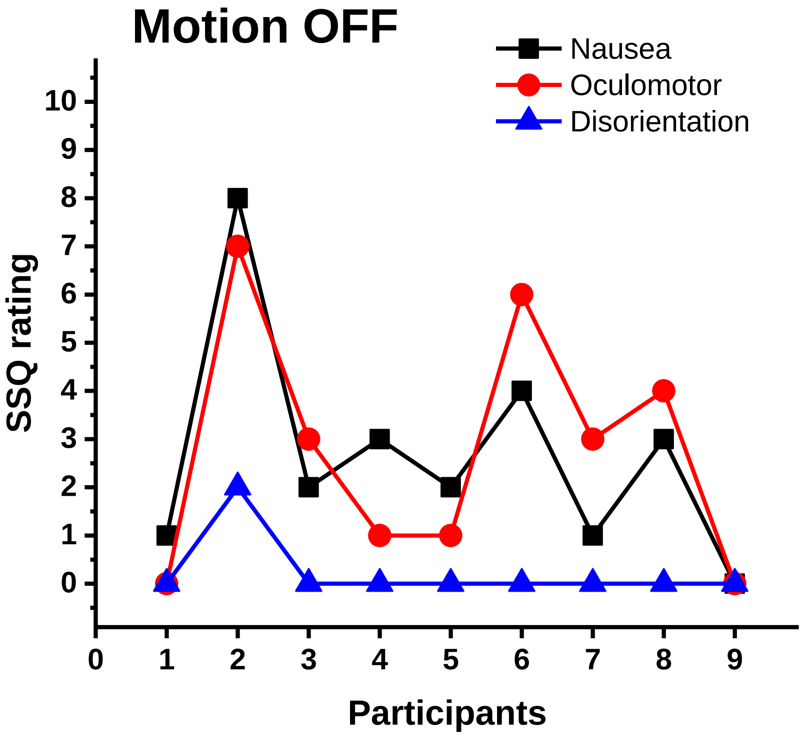
6.3.2 Subjective data

6.3.2.1 The effects of motion on simulation sickness

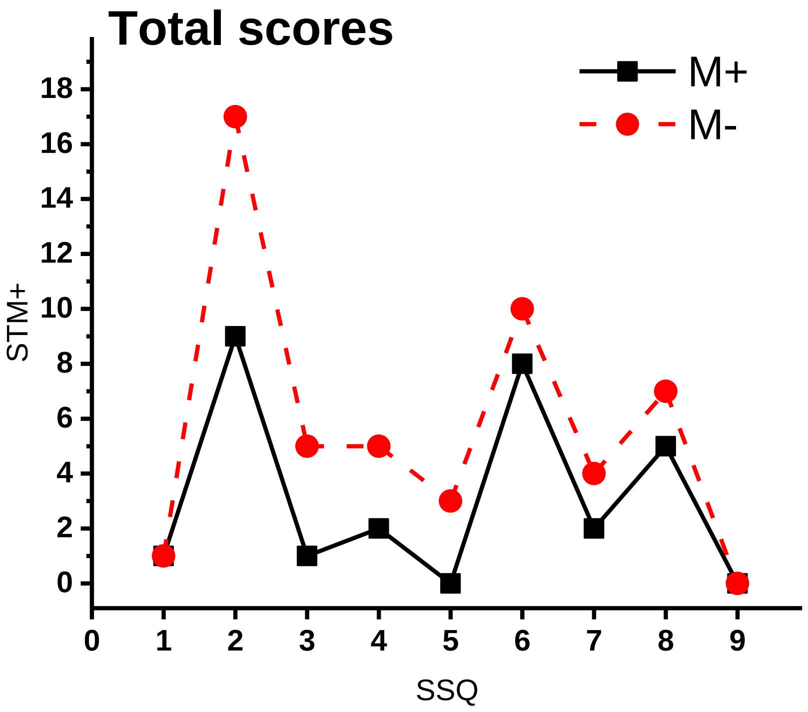
The levels of simulation sickness were recorded after each experimental block, where motion cues were either present or absent. Figure 41 shows the overall sickness rating scores on participant's level with and without motion cuing. The graphs suggest that the levels of simulator sickness increased when motion cues were not provided.



A



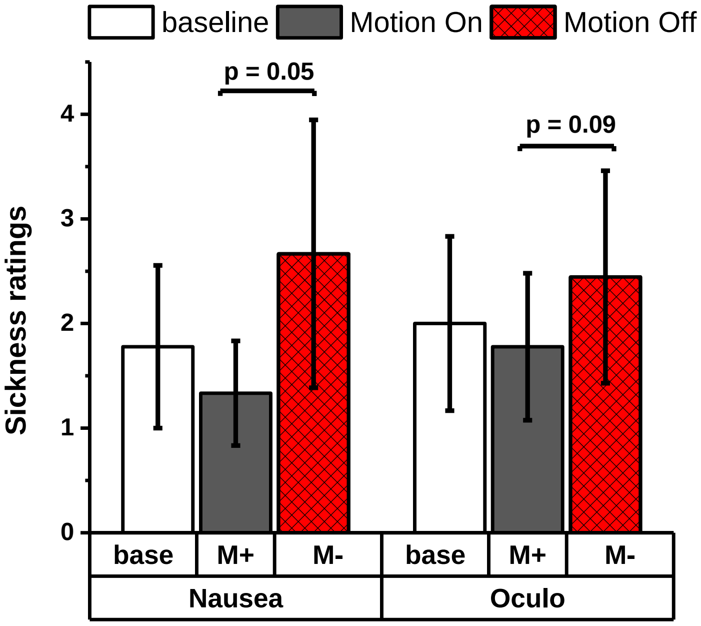
B



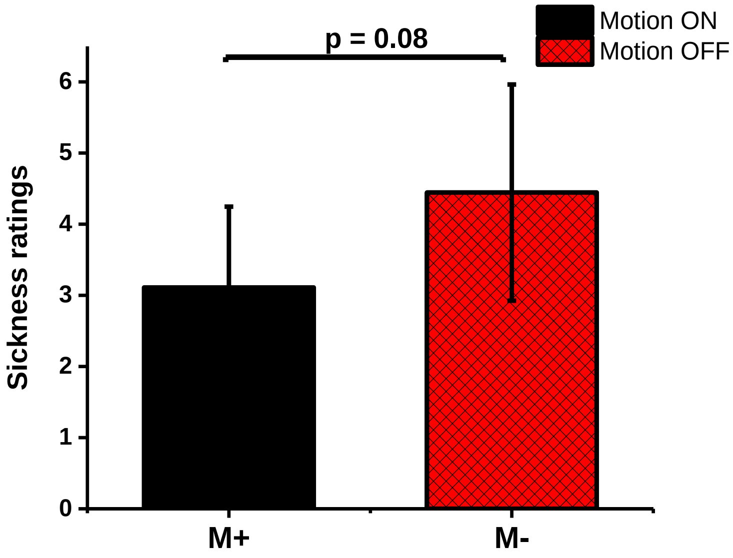
C

Figure 41. Mean sickness ratings on SSQ subscales across participants with motion cuing (A) and without motion cuing (B). The total severity scores graph (C) shows that participants experienced higher levels of simulation sickness in the condition where no motion feedback was provided.

All subscales as well as the total severity scores were included in the statistical analysis. A non-parametric Freidman ANOVA was used to analyse whether the SSQ scores were affected by the motion feedback with three dependent variables: nausea, oculomotor, disorientation and total SSQ scores. The analysis revealed a significant difference in sickness scores across all subscale and in the total severity scores (x^2 (8) = 18.437, p= 0.018, W = 0.25). As it was predicted that the absence of motion feedback will produce increased levels of simulation sickness, one tailed planned comparisons with Wilcoxon sign rank tests were performed on the SSQ subscales and on total SSQ scores. The mean sickness scores were always higher in all subscales of SSQ and on the total scores when motion cuing was not provided. Significant differences were recorded on Nausea subscale (Z=-1.633, p = 0.05, η = 0.38) and there was a trend towards significance on the Oculomotor subscale (Z=-1.342, p = 0.09, n = 0.31) and on the total sickness scores (Z = -1.298, p = 0.08, η = 0.30) (Figure 42). No difference was recorded in Disorientation subscale (p = 0.341).



A



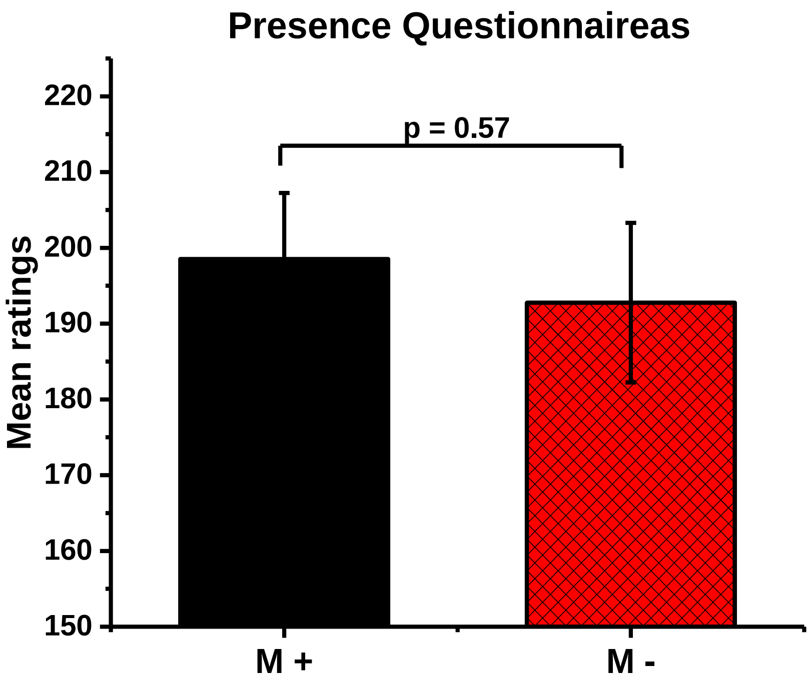
B

Figure 42. Overall cores on the Nausea and Oculomotor subscales with and without motion feedback (A). Total severity scores as reported by SSQ with and without motion feedback (B).

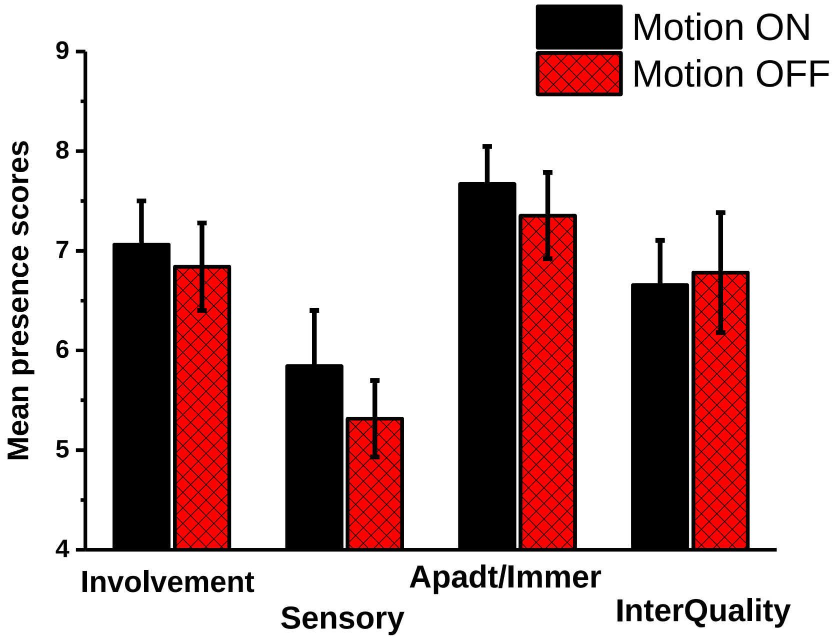
6.3.2.2 Presence Questionnaire data

To investigate the effects of motion cuing on the perceived sense of presence, the Immersive Tendencies Questionnaire (ITQ) and the Presence Questionnaire (PQ) designed by Witmer and Singer (1998) were used; ITQ was filled in before the task and PQs were filled in after each experimental block (Motion On and Motion Off). The levels of presence were recorded after the participants completed all three levels of perceived task demand. The reliability of the questionnaire was assessed by calculating the coefficient Cronbach alpha. The standardised alpha of the PQ questionnaire was 0.0912, which suggests a good reliability as the coefficient is greater than the suggested value of 0.70 (Nam et al., 2008).

The overall presence ratings ranged from 136 to 229. Participants reported a higher sense of presence in the motion condition ($M = 202.4$, $SD = 29.5$) than in the no motion condition ($M = 192.7$, $SD = 31.5$), however the significant differences between the presence ratings were not observed ($p = 0.57$). No significant correlations were found between ITQ and PQ ($p = 0.267$). All of the questionnaire data were grouped into a 4-factor model suggested by Witmer et al. (2005) to investigate the reliability of each subscale and the overall effects of motion feedback on presence ratings. The descriptive statistics and reliability analysis for each PQ subscale can be seen in the Table 10. Repeated measures ANOVA revealed a significant main effect of PQ subscale ($F(3, 24) = 11.546$, $p < 0.001$, $\eta^2 = 0.59$) but no significant effect of motion ($F(1,8) = 2.05$, $p = 0.19$, $\eta^2 = 0.21$) and no significant interaction ($F(3,24) = 0.572$, $p = 0.639$, $\eta^2 = 0.06$). The results show that motion feedback influenced the levels of presence in a positive manner, however no significant differences between two motion cuing conditions was found (Figure 43).



A



B

Figure 43. Mean overall scores for experimental block with and without the motion (A); Mean presence scores in each subscale of PQ (B). Error bars represent standard error. M+ = Motion On; M- = Motion Off.

6.3.2.3 NASA TXL workload ratings

NASA TXL scale measures workload on five dimension: mental demand refers to the easiness or simplicity of the task; physical demand refers to the amount of physical activity such as pushing, pulling, turning, controlling; temporal demand refers to the time pressure experienced during the task, such as slow or fast pace of the task; effort refers to how hard participant had to work to accomplish the task; frustration refers to the levels of insecurity, irritation or annoyance experienced during the task; performance refers to user's satisfaction with their performance (Hart & Staveland, 1988). Participants were instructed to place a mark on the line from 0 (low) to 10 (high) to indicate the levels of workload in each dimension. Higher ratings indicate an increased task demand, except for performance where higher scores indicate an increased satisfaction; therefore the data from the performance category were converted. The workload ratings for each pole width separation when motion cues were present and absent can be seen on Figure 44.

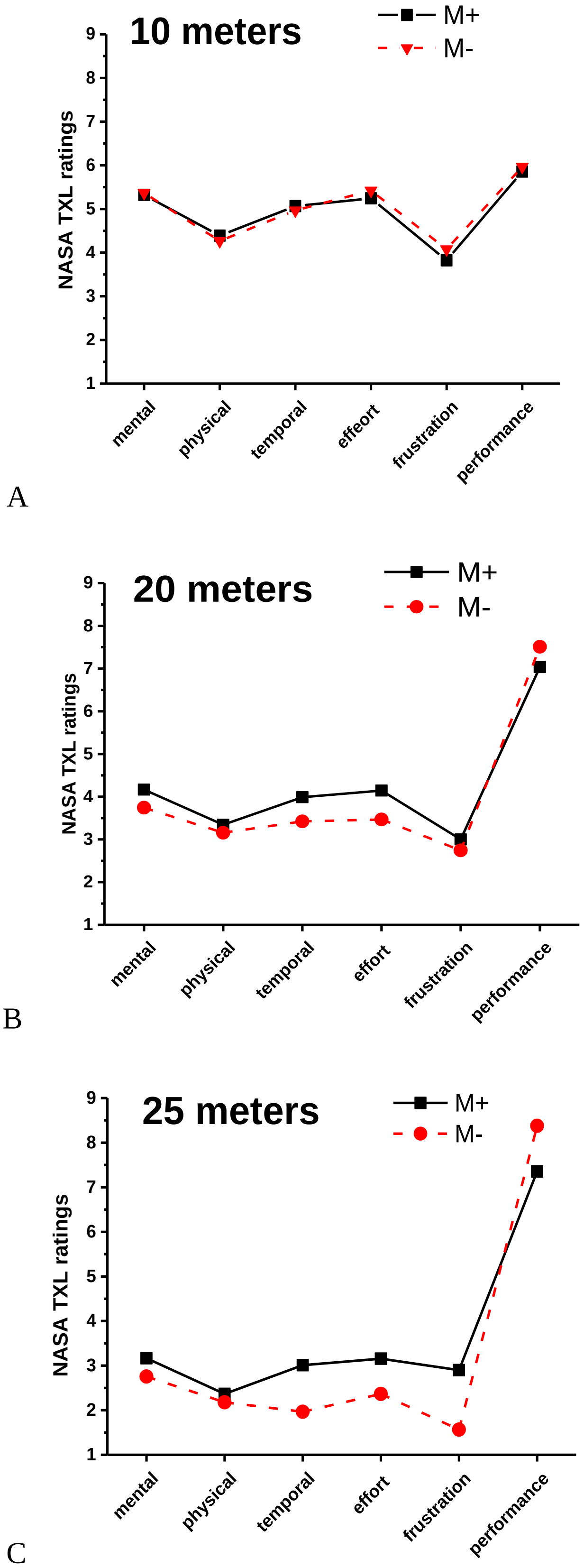


Figure 44. NASA TXL workload rating scale. Above graphs indicate ratings in each dimension for every task condition, 10m (A), 20m (B), 25m (C); with motion feedback (solid line) and without motion feedback (dash line).

Repeated measures ANOVA was performed using a 2 (motion on and off) x 3 (pole width separations) experimental design to investigate the effects of motion cuing and task demand on workload scores. The analysis of the overall mean workload scores showed that motion cuing produced the same levels of workload as no motion cuing (F(1.8) = 0.933, p = 0.36, η^2 = 0.104). The significant difference in workload scores was recorded between the different task demand conditions (F(1.14, 9.12) = 17.75, p = 0.002, η^2 = 0.689). There was also a trend towards a significant interaction between the motion and the task demand (F(1.74, 13.9) = 3.110, p = 0.08, η^2 = 0.28). Planned comparison with adjusted significance level ($p < 0.008$) revealed that in the motion condition the workload scores differed significantly between the 10m and 25m conditions (t(8) = 5.11, p = 0.001, d = 1.7) and the 10m and 20m conditions (t(8) = 4.52, p = 0.002, d = 1.5). In the no motion condition significant differences in the workload scores were recorded between the 10m and 25m conditions (t(8) = 4.06, p = 0.004, d = 1.64) and between 20m and 25m conditions (t(8) = 3.44, p = 0.008, d = 0.88) (Figure 45).

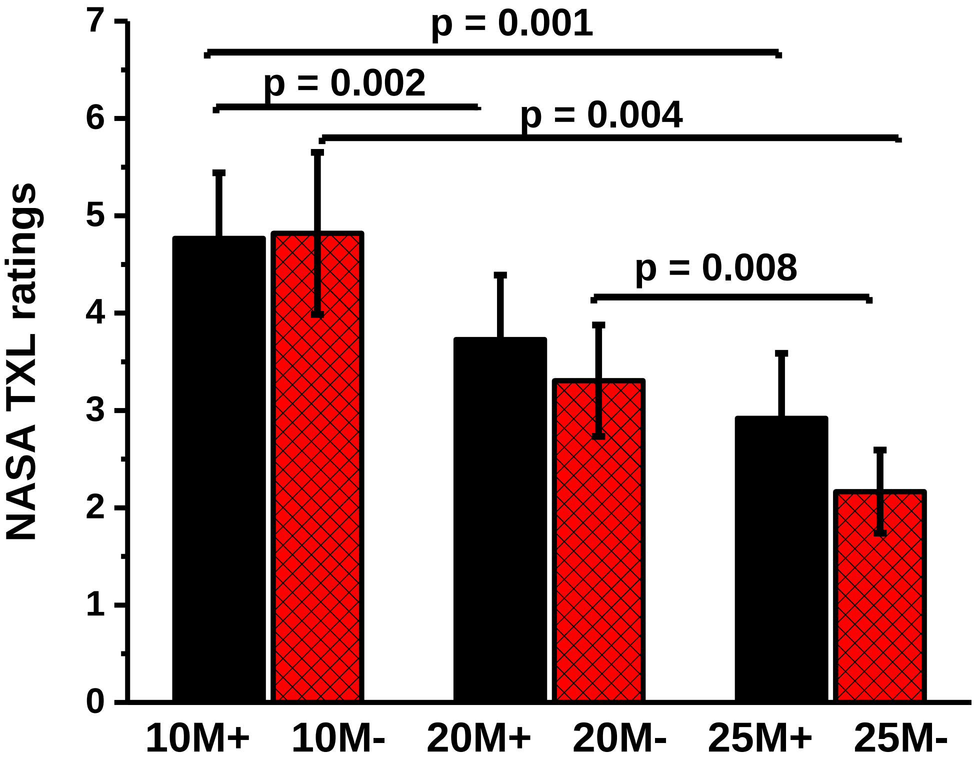


Figure 44. Mean and standard errors of NASA TXL workload ratings in three task demand conditions with (M+) and without (M-) the motion cuing.

A series of a 2 x 3 repeated measures analyses of variance was also performed on each dimension of the NASA TXL workload scale to investigate which dimension was most affected by the increased task demand (Bell and Grant, 2011). The analysis of mental demand found a significant effect of task demand (F(2,16) = 18.183, p = 0.001, η^2 = 0.694) but no significant effect of motion (F(1/8) = 0.423, p = 0.53, η^2 = 0.05), or interaction (F(2,16) = 0.490, p = 0.621, η^2 = 0.058). Planned comparisons, with adjusted significance level revealed that in the motion condition the workload scores differed significantly between the 10m and 20m conditions (t(8) = 4.32, p = 0.002, d = 1.49) and the 10m and 25m conditions (t(8) = 4.58, p = 0.002, d = 1.58). In the no motion condition significant differences in the workload scores were recorded between the 10m and 25m conditions (t(8) = 3.68 p = 0.006, d = 1.34).

The analysis of physical demand found a significant effect of task demand (F(2,16) = 9.656, p = 0.009, $\eta^2 = 0.547$), but no significant effect of motion (F(1/8) = 0.254, p = 0.628, $\eta^2 = 0.031$) or interaction (F(2,16) = 0.017, p = 0.983, $\eta^2 = 0.002$). Planned comparisons with adjusted significance level revealed that in the motion condition workload scores differed significantly between the 10m and 25m conditions (t(8) = 3.94, p = 0.004, d = 1.32). In the no motion condition significant differences were recorded between the 20m and 25m conditions (t(8) = 4.41, p = 0.003, d = 1.39).

The analysis of temporal demand found a significant effect of task demand (F(2,16) = 16.397, p = 0.002, $\eta^2 = 0.672$), but no significant effect of motion ((F(1/8) = 1.431, p = 0.266, $\eta^2 = 0.152$) or interaction (F(2,16) = 1.25, p = 0.287, $\eta^2 = 0.144$). Planned comparisons with adjusted significance level revealed that in the motion condition workload scores differed significantly between the 10m and 20m conditions (t(8) = 3.85, p = 0.005, d = 1.33) and the 10m and 25m conditions (t(8) = 5.28, p = 0.001, d = 1.81). In the no motion condition significant differences were recorded between 20m and 25m conditions (t(8) = 3.65, p = 0.006, d = 1.34).

The effort component of workload showed a significant effect of task demand (F(2,16) = 12.787, p = 0.005, $\eta^2 = 0.615$), but no significant effect of motion ((F(1/8) = 0.794, p = 0.399, $\eta^2 = 0.09$) or interaction (F(2,16) = 1.49, p = 0.255, $\eta^2 = 0.157$). Planned comparisons with adjusted significance level revealed that in the workload scores differed significantly between the 10m and 25m conditions (t(8) = 3.55, p = 0.007, d = 1.45) in the no motion condition only.

The frustration component of workload showed a significant effect of task demand (F(2,16) = 12.467, p = 0.002, $\eta^2 = 0.609$), but no significant effect of motion ((F(1/8) = 0.673, p = 0.436, $\eta^2 = 0.078$) and there was a trend towards a significant interaction (F(2,16) = 3.268, p = 0.067, $\eta^2 = 0.29$). Planned comparisons with adjusted significance level revealed that the workload scores differed significantly between the 10m and 25m conditions (t(8) = 4.52, p = 0.002, d = 1.81) in the no motion condition only.

6.3.2.4 Objective and subjective measures

Correlation analyses were performed on objective measures of performance and subjective measures of presence, sickness and workload. During the motion cuing the overall accuracy in performance was significantly affected by the levels of sickness (r = 0.715, p = 0.015) and by the levels of workload (r = 0.732, p = 0.024). The positive correlation implies that as the levels of discomfort and sickness increased the deviations from the desired trajectory also increased (Figure 46, A). There was a trend towards a significant effect of sickness levels on presence ratings (r = -0.551, p = 0.062). Here, a negative correlation implies that as the levels of simulation sickness decreased, the reported levels of immersion and presence increased. Higher levels of presence moderately influenced the amount of deviances from trajectory pattern, however the association was found to be not significant (r = -0.36, p = 0.32). When motion feedback was not provided, no significant correlations between the variables were found (Figure 46).

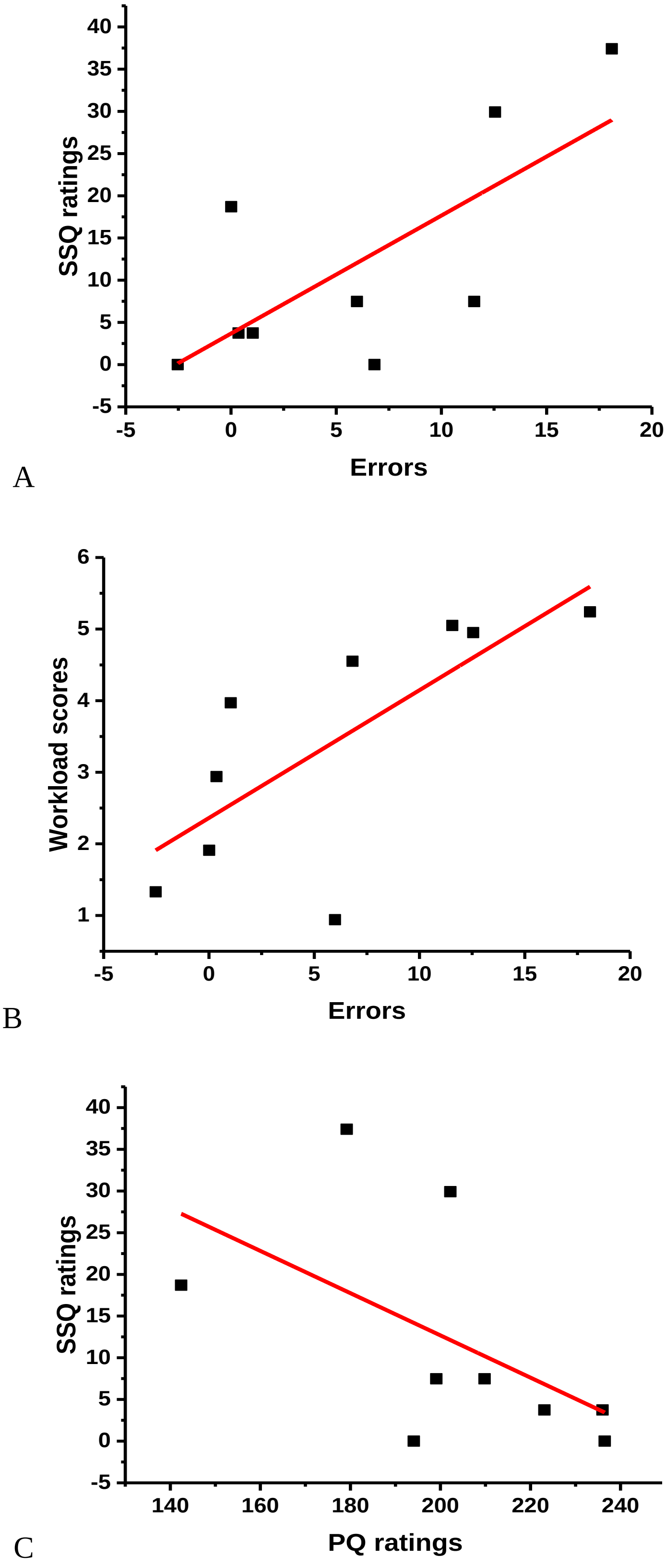


Figure 46. Correlational analyses during motion cuing. A significant positive correlation implies that the trajectory deviances were negatively affected by the sickness ratings (A) and by the workload scores (B). There was a trend towards significant negative correlation between SSQ and Presence ratings (C).

6.4 Discussion

This study investigated the effects of motion cuing and perceived task demand on performance and psychological state of the users by assessing the levels of workload, sickness and presence. Pilots were required to perform a slalom manoeuvre in the 6 DOF helicopter simulator with and without the motion feedback, where the task demand was varied in three conditions. The increase and decrease in task demand was achieved by the variation in the pole width separations that pilots have to fly through to successfully complete a slalom manoeuvre. The deviances from the desired trajectory pattern were recorded as the magnitude of errors. Overall, the presence of the motion feedback facilitated the accuracy of performance, however this was influenced by the perceived task demand condition. The motion feedback also modulated pilot's subjective levels of workload and discomfort but not the sense of presence. The following sections will provide a more detailed discussion of the found main effects.

6.4.1 Motion cuing and the accuracy of performance

The results revealed that the presence of motion feedback influenced the performance accuracy, i.e. the ability to follow predesigned trajectory path. These findings are in line with previous research that suggested the beneficial effect of motion cuing on performance (Feenstra et al., 2011, Hogema et al, 2012; Siegler et al, 2001). Even though no significant main effect was recorded, a significant interaction between the motion feedback and the task demand conditions was found, which suggests that both of them influenced the accuracy of task performance. The results showed that the magnitudes in errors recorded during the task were higher when motion feedback was not provided, however they were moderated by the perceived task demand. In particular, motion feedback facilitated the magnitudes of errors in the 10m and 20m conditions, but not in the 25m condition. The performance was significantly reduced during low task demand condition when motion feedback was present. One possible explanation for this finding can be in the level of control that is clarified in the pilot's control model proposed by Gray (2005).

According to Gray's (2005) boundary avoidance model for tracking tasks, the pilot's level of control increases linearly as the boundary is approached. This practice is influenced by other factors, such as flight control system, aircraft dynamics and other sources of sensory information (Gray, 2005; Lu & Jump, 2013). Lu and Jump (2013) investigated whether pilots behaviour can be affected by the perceived boundaries and concluded that boundary avoidance might result in increased tracking error in the task, however they also suggested that this error could be modulated by the inclusion of vestibular and proprioceptive cues i.e. the motion feedback. Similarly, Hess and Marchesi (2009) proposed a pilot model in which the control behaviour of pilots is predicted by taking into account the role of visual, vestibular and proprioceptive cues. All of these studies suggest that sensory cues could be stimulated by the activation of motion platform. In this study, the presence of motion feedback did modulate the accuracy in the task performance, thus confirming previous research (Lu & Jump, 2013).

However, the results further show that the modulation in the perceived task demand can influence pilot's flying strategies and changed their behaviour. The beneficial effects of motion platform can be seen only during the high and medium task demand conditions, but not in low demand condition. In this task, the poles on the slalom track represented the boundary that needs to be avoided in order to complete the manoeuvre successfully. When the poles were close together, pilot's levels of control and concentration seems to increase when they approached the boundaries (poles), as predicted by Gray’s boundary avoidance model. This kept their levels of arousal high and enabled them to perform the task to the best of their abilities. However, when the poles were far apart, the danger of hitting the poles was decreased; and so were their levels of control which resulted in the decreased levels in the accuracy of performance. The results showed that participants performed significantly more errors in the low demand condition with motion feedback as oppose to the condition without motion feedback.

The levels of subjective workload could be proposed as one explanation for these findings. Subjective measures showed that pilots experienced lower levels of workload, which decreased their levels of arousal and consequently led to the increase in the magnitude of errors. Jackson et al., (2014) investigated the effect of low cognitive workload on a number of performance tasks and suggested that low cognitive load and low arousal could be considered as factors in performance decline, including decision-making. This study confirms previous findings as during the condition where low cognitive load was reported the accuracy of the performance decreased. Therefore, it is suggested that further investigations of the effects of workload levels on performance accuracy are necessary, thus supporting the notion of using objective and subjective measures of performance.

6.4.2 Motion cuing and workload

When examining the self-reported workload scores collected from participants across all task demand conditions, the workload levels differed between motion conditions. As predicted, during the low task demand conditions participants reported lower workload and in the high task demand conditions participants reported higher workload. Overall, the cognitive workload was unaffected by the motion feedback in all task demand conditions. This result is in contrast to previous studies such as Bell and Grant (2011) who compared workload levels across three different technologies and found that a 6 DOF motion platform produced the highest workload scores as compared to motion seat and no motion condition. The results suggest that the availability of motion feedback did not affected the levels of workload and thus provide support for previous research that suggested the possible advances of motion feedback in VR (Hodge et al., 2015, Wang et al., 2013). When looking at the perceived task demand conditions, participants systematically reported lower levels of workload in the decreased task demand condition. Significant differences were observed between 10m, 20m and 25m pole width separations and as such the conclusion that subjective workload levels are moderated by the perceived task demand can be made.

6.4.3 Motion cuing and the levels of presence and discomfort

The findings in this study revealed that the absence of motion cuing modulated the levels of simulation sickness. Higher levels of sickness were recorded on Nausea and Oculomotor subscales of SSQ when motion cuing was not provided. These findings support the Sensory Conflict theory that predicts that inconsistencies between sensory modalities create conflict that contributes to sickness symptoms (Reason & Brand, 1975).

The overall scores on PQ showed that participants reported an increased sense of presence during motion feedback as oppose to no motion feedback, however the PQ scores between the two motion conditions failed to reach significance. These results are only partly in line with previous research that suggested that simulated motion could result in positive effects on presence and immersion (Mulder et al, 2012).

When investigating relationships between the variables it was found that the relationship exists between objective and subjective measures. Steven and Kincaid (2015) investigated possible relationship between the levels of presence and performance and they found a moderate relationship between these two variables. Correlational analyses performed in this study showed that the levels of presence only moderately influenced the task performance as no significant difference between motion conditions was found. It seems that in this case the motion feedback added little to the overall sense of presence perceived by the pilots. This finding may be due to the high fidelity visual information that was able to produce sufficient levels of involvement and immersion and as a result participants felt an increased sense of presence even when no motion feedback was provided.

A significant positive relationship was found between the sickness ratings and the performance accuracy. The positive relationship implies that when pilots reported increased levels of sickness their magnitude of errors also increased. Similarly, the levels of cognitive workload significantly affected the amount of errors recorded during motion cuing as the increase in workload scores was related to the increase in error rates. These findings are in line with previous research that reported that the presence of motion feedback modulates objective task performance as well as the subjective measures of increased sickness and workload ratings (Milleville-Pennel & Charon, 2015; Stevens & Kincaid, 2015). The absence of significant relationships between variables during the no motion condition suggests that motion feedback may be an important factor in the overall effectiveness of motion simulations. These findings further suggest that the use of objective and subjective measures is necessary for effective evaluations of virtual environments.

6.4.4 Limitation of the study

One of the main limitations of this study was the technological set up of the task. This study found no significant differences between the motion cuing conditions. This effect can be explained by the technological set up of the study. In order to enable a recruitment of the less experienced participants the flight altitude and the flight speed were kept constant. This might have contributed to no observable effect of motion between the two conditions. Another limitation of this study was the sample size. Even though some participants had previous flight experience, most of them had not flown a helicopter simulator before which could have influenced their subjective responses to the contribution of motion cuing in fidelity assessments. Individual differences within the participants also needed to be taken into account, as there was no common standard introduced for the evaluation of sense of presence and immersion. This meant that each participant rated their experience on an internal scale for evaluation of their subjective experience in terms of presence and immersion. Another limitation of this study is the lack of the physiological measures, which could of have provided more robust results when investigating the relationship between objective and subjective measures.

Chapter 7

Overall Discussion and future recommendations

7.1 Summary

The main aim of this thesis was to explore the role of sensory feedback in virtual reality environments and provide a conceptual framework for effective fidelity evaluation. In order to achieve this aim I designed and conducted a series of experimental studies where I have investigated the effects of visual, audio, haptic and motion sensory cues on the task performance and psychological state of the users. This was accomplished by collecting the objective measures of performance, such as postural responses, task completion time, error rate and accuracy of performance as well as the subjective measures of user perception, which included the perceived sense of presence and immersion, simulation sickness and cognitive workload. The overall contribution and the main conclusions of this thesis are summarised below.

7.2 Overall contribution of this thesis

This thesis main contribution lies in the investigation of the role of multisensory cues in fidelity assessments in virtual reality environments. Through the series of studies I have shown that multisensory cues, in particular visual, audio, tactile and haptic sensory cues all play an important part when investigating the effectiveness of VR system. By collecting objective measures of performance and subjective measures of user satisfaction I have shown that user experience needs to be included when investigating the usability of the feedback signals provided by VR technologies. This thesis has expanded previous research on fidelity in three points:

- By investigating how visual disparity cues and physical properties of the reference point within the environment affect postural responses to visual motion stimulus i.e. to see whether postural stability can be used as an objective measures of presence and fidelity.
- By using augmented multisensory cues during manual assembly task in high fidelity VR environment and additionally exploring whether the benefits of the virtual training with these cues transfer into performance improvements on the real task i.e. to investigate the contribution of physical versus information fidelity in virtual and real environments.
- By investigating benefits of motion cuing during simulated flight task during the manipulation of the perceived task demand i.e. to investigate the relationship between motion cuing and cognitive workload during fidelity assessments.

7.3 The thesis main conclusions

7.3.1 Environmental and perceptual factors in fidelity assessments

Visual and postural signals that are part of the integrated system to maintain balance also form the basis of perceived vection and presence in virtual reality. A previous study has shown that postural responses to visual motion differ in the presence of virtual and real environmental anchors (Meyer et al., 2013). I have extended previous research and shown that for the current VR system, with improved graphical image rendering, tracking accuracy and reduced tracking and visualisation lag, postural responses for real and virtual foreground environmental objects are modulated in a different ways. In particular, the fidelity of the virtual environment was affected by the position of the convergent plane and influenced by the vergence-accommodation conflict that is often the cause of visual fatigue and discomfort. Similarly, the concurrent evaluation of the positional stability of environmental anchors presented in the scene also contributed to postural responses.

The unpleasant aspects of illusionary self-motion, such as visual discomfort that can cause headache and nausea should also be addressed in an effective manner. The results from my first study confirmed previous findings that the presentation of real static reference points in the scene and reduction of non-informative signals, such as background motion, could reduce postural responses. I also found that even virtual static reference points can reduce the amount of postural sway assuming a high quality motion tracking system with high quality image rendering is used.

The improved fidelity of VR can also influence users' subjective feelings of presence and immersion, which has previously been found to increase interaction and performance in virtual reality environments. However, consistent subjective ratings of fidelity in VR environments are difficult to obtain, in particular when using a small set of participants who might be biased due to previous experiences (Meyer et al., 2012) or influenced by demand characteristics (Young et al., 2007). As the VEPR component of postural sway is performed in an unconscious manner, I propose that postural behavioural responses to multisensory stimulation can provide a viable option for the objective assessment of fidelity in VR that can later be compared to subjective ratings. As the assessment and configuration of the environment (virtual or real) is done on high levels of the explicit judgment of surroundings, I suggest that in order to avoid breaking the levels of immersion and presence when they are required, a careful configuration of the VR set up and laboratory space is necessary. When the levels of immersion and presence, as well as illusionary self-motion are not required, I propose the addition of real objects with static physical properties within the environment to prevent undesirable effects of motion sickness. Future research should concentrate on the investigations of other factors that could influence postural responses; for example, the effects of different foreground objects on postural stability, the introduction of system latency in the simulation and the addition of more interactive tasks during which participants' cognitive load, presence and discomfort could be measured and compared.

7.3.2 The effects of augmented sensory cues on fidelity assessments

The beneficial effects of augmented multisensory cues have been investigated in two studies in this thesis. It was found that augmented visual, audio and haptic cues presented as unimodal, bimodal and multimodal feedback can improve task performance and users' subjective perception of the virtual environment. The results from the studies presented in this thesis indicate that when additional information is presented in VR simulation, the information that is task-relevant is valued more than the decreased fidelity of the VR environment. Previous research has suggested that higher fidelity results in higher performance, however some research studies have suggested that instead of upgrading the whole VR system, increasing the fidelity of individual components of virtual simulation might be sufficient for overall task improvements (McMahan et al, 2006, Dahlstrom et al., 2012). In this thesis I have shown that the presented sensory feedback does not need to have high levels of fidelity as long as the information presented is useful and relevant to the task in hand; as a result these sensory cues can not only enhance overall task performance but also facilitate the user's acceptability of the VR system.

The necessity to provide appropriate, task-relevant sensory cues is also important in order to support training in the VR environment. I have shown that virtual training with augmented sensory cues does not hinder user perception and acceptability of the virtual environment; it can, in fact transfer into the performance improvements in real life tasks. Additional cues provided in this study (vibration) were chosen to represent cues that are sometimes difficult to simulate in VR environments (torque and weight). I propose that the additional, task relevant cues can be effective and efficient substitutes for the cues that are not so easily achieved in VR environments and I strongly encourage their use during virtual training in future research.

7.3.2.1 Motion tracking and augmented sensory cues

Accurate motion tracking in VR environments is necessary to allow for natural interaction between the user and the environment. The latency, lag and inaccuracy of the tracking system used in VR are all important parameters for the understanding of the effectiveness of VR systems in terms of sickness symptoms, performance and presence. In this study I have shown that positional inaccuracy that is often present in low cost motion tracking system, and that was implemented as the visual sway of the virtual scene, can affect performance on a manual assembly task in VR; yet this effect can be mediated by the presentation of augmented sensory cues that carry information relevant to the task in hand. These findings further confirm the advantages of augmented multisensory presentation. The main implication arising from this research suggests that even though the overall fidelity of the environment can be decreased to aid performance, the overall quality of the motion tracking system, in terms of accuracy used for human motion tracking should not be negotiable. This is especially important in order to avoid negative effects of simulation sickness on user experience. The reduction of the negative effects of tracking inaccuracy is the main goal of many VR applications and thus a precise and accurate motion tracking system is required to effectively track human motion, improve presence and eliminate sickness symptoms. As simulation sickness is still the biggest issue of VR systems, there is a need for further investigation, evaluation and the development of effective measures that can capture simulation sickness symptoms in order to minimise its effects and develop an efficient virtual reality set up to encourage the use of VR systems in future applications. In this study I have added extra knowledge in terms of the advances of augmented sensory cues and the necessity to implement a precise and accurate motion tracking system for the effective evaluation of fidelity in VR environments.

7.3.3 Motion cuing and fidelity assessments

The results of the study that investigated the effects of motion cuing on performance and cognitive workload suggest that the presence of motion feedback can facilitate task performance when motion is expected and as such can be used as a suitable training tool. The presence of motion feedback facilitated subjective levels of perceived sense of presence and decreased the overall levels of discomfort. This finding is of great importance as motion sickness, as mentioned before, is still a major drawback of many VR applications. The high levels of cognitive demand may lead to an increased amount of errors that can be detrimental for training and for real operations. In this study I have shown that motion cuing can have differential effects on different levels of cognitive load; motion cuing aided performance during the medium and high demand conditions, however it had the opposite effect during the low demand condition. This suggests that there is a link between objective measures of performance and subjective measures of users perception. Therefore, future research should continue to use objective and subjective measures of performance simultaneously and further investigate the conditions under which motion feedback is perceived as useful and beneficial not only to the overall performance but also to the psychological state of the users.

7.4 Objective and subjective measures in fidelity assessments

In all of the studies reported in this thesis the objective measures of performance have been collected whilst at the same time users' subjective perception of the environment were also obtained. The correlational analyses were conducted in every study reported in this thesis in order to investigate the relationship between the objective and subjective measures in the task where multisensory information was presented during the VR interaction. Sensory information provided by multimodal interfaces had beneficial effects on performance and users' acceptability of the virtual environment and as such it can be concluded that multisensory presentation is beneficial and can influence the fidelity assessments of VR environments. The addition of augmented sensory cues into the virtual environment was acceptable even when the presented cues decrease the overall fidelity of the VR environment. I have also shown that the presentation of augmented cues can facilitate the negative effects of positional inaccuracy that can be present when using a low cost motion tracking system during the VR exposure. Subjective measures recorded during the task provided highly reliable results as they show consistent and significant differences that are mirrored in our objective measures.

Additionally, the presence of augmented cues also enhanced users' subjective perceptions of presence. As the technological part of the VR equipment, also referred to as the immersion of VR, was slightly compromised by the presentation of unrealistic sensory cues, but the perceptual state of the users was not, it can be concluded that in these studies the sense of presence was more important to the users than the overall immersion of the virtual environment. The results of the motion tracking study showed that the overall fidelity of the environment can be degraded to aid performance on a manual assembly task, however the importance of an accurate motion tracking system should not be compromised in order to avoid negative effects of simulation sickness.

In the same way, motion cuing that provided by a high fidelity helicopter flight simulator, also influenced task performance and the subjective experience of the users, however the levels of perceived task demand modulated this effect. Positive effects of motion cuing were evident during medium and high task demand conditions, however motion cuing with low demand caused a decline in task performance in terms of accuracy. This suggests that the levels of cognitive workload are dependent on the available sensory cues presented during the task. This suggests that the implementation and investigation of objective and subjective measures of performance are necessary to gain a deeper understanding of the environmental and cognitive factors that can influence user experience in VR in terms of fidelity assessments.

7.5 Implications

One of the main implications of this thesis is to offer new insight and inspire future research in multisensory perception, sensory cuing and sensory feedback that are provided during VR interactions. The research findings in chapter 3 suggest that the presentation of visual information in VR is an important parameter. By making careful and informed decisions about the configuration of the VR environment, the findings show how the undesirable effect of illusionary self-motion that can give rise to motion sickness, can be lessened or even eliminated.

The research findings from chapter 4 and 5 showed potential advantages of multisensory cuing in terms of improved performance. In particular, audio and tactile feedback presented unimodally have shown great advantages and thus future research could explore these modalities in more detail. In addition the research findings also showed that informational content of the sensory cues or informational fidelity, is valued more by the users that the overall physical fidelity of the cues. Findings from Chapter 5 further confirmed that virtual reality can be used as an efficient training platform where the performance can be improved and transferred into real life scenarios. The understanding of the factors and conditions of multisensory cuing under which VR users experience an enhanced sense of presence and immersion can help designers to allocate computational resources proportionally when building future designs of virtual systems with multimodal feedback. Therefore, more detailed research into the use of realistic as well as additional augmented multisensory cues and its effects on performance and user subjective performance in virtual environments is recommended.

The main implication stemming form the research findings described in chapter 6 suggest that acceleration cues are beneficial for performance when using motion platform in a simulated flight task. Additionally it was show that when considering the cognitive workload of pilots, a certain amount of workload is desired as it keeps adequate levels of arousal and concentration on the task in hand.

The main argument stemming from this thesis suggests that user experience needs to be included when investigating the usability of feedback signals in the virtual reality environment. This thesis has also provided evidence for further exploration of the relationship between objective and subjective measures. Throughout this thesis I have consistently shown that subjective measures of user perception in VR provide consistent results that are directly comparable with objective measures of performance. One of the implications of this thesis emphasises how to deal with the high inter-variability between participants when investigating the relationship between objective and subjective measures. Normalisation of the data was used as one of the ways to overcome the inter-subjective variability. This method could be used in future studies that wish to address this problem in more detail.

The main implication arising from this research is the importance of combining the objective measures of performance and subjective measures of a user perceptual experience in order to achieve an effective and efficient evaluation of fidelity in virtual reality environments. As such, a continuous use of metrics that can recorded data from both subjective and objective measures is recommended as it can provide a holistic view of user experience i.e. getting a deeper understanding of environmental and cognitive factors that can influence user experience in VR. Additionally, the findings from the research studies described in this thesis could provide future designers with a new way to approach VR design from a more informed perspective especially in the context of sensory cuing and interactivity.

7.6 Directions for future research

The research work presented in this thesis contributed to the existing knowledge in a number of ways, as described in Chapter 1. As with any other experimental research, limitations of the research studies described in this thesis were identified and have been discussed in more detail in the discussion section of each chapter. This section will offer suggestions for future research work. Firstly, the overall improvement for future studies will be provided, following with the future research recommendations for each study described in this thesis.

7.6.1 Overall future recommendations

7.6.1.1 Methodology

Quantitative and qualitative methodologies were adopted in all of the studies presented in this thesis. Even though measures to assess user subjective experience were used, future research would benefit by including more qualitative methodology across the studies. These could include focus groups after the study to allow participants to express their feelings in a less formal environment and enable further discussion about the study, the aftereffects and the task itself. Semi-structured interviews could also be performed, especially with professional test pilots who are trained to provide a detailed description of their experience when using VR systems. This additional data would then provide a basis for a more informative approach to the assessment of fidelity in VR environments.

7.6.1.2 Physiological measures

The inclusion of physiological measures in all of the studies presented in this thesis would provide another facet that could support the reliability of the findings. For example, previous studies such as that of Bertin et al. (2005) investigated physiological reaction during induced simulation sickness and found that heart rate and skin temperature correlated with simulation sickness. Additionally the use of distractors, such as multisensory reference points within the environment could be included in the future studies to investigate whether the levels of overall discomfort could be decreased. Finding a link between physiological measures, subjective ratings and overall performance could provide a clearer picture of how to prevent or eliminate simulation sickness symptoms. The collected data and its subsequent analysis could yield interesting findings that would help us to understand in more detail what sensory and physical changes occur and how these might relate to behavioural changes during the task. Additionally, finding a link between physiological measures, objective measures of task performance and subjectively reported measures of satisfaction would enable a deeper understanding of the environmental factors and overall user experience in VR environments.

7.6.1.3. Age and Gender

Another improvement of the studies reported in this thesis would be to consider performing the task with participants of different ages. None of the studies reported in this thesis investigated the effects of age or gender, due to limited sample sizes and time constrains. Empirical research, such as that of Barr and Giambra (1990) and more recently Novak et al. (2016), reported age related deterioration in the sensitivity of auditory stimuli, which can be affected by procedural factors. These findings therefore suggest that older participants might have preferences for other sensory cues, such as visual or haptic. Future research could therefore investigate whether the robustness of the findings reported in this thesis could be influenced by the age of the participants. Additionally, the effects of gender and personality variables could be investigated in more details and possibly provide some interesting findings.

7.6.2 Future recommendations based on each chapter

7.6.2.1 Postural response to visual motion

One of the main limitations in the studies described in Chapter 3 was the lack of subjective evaluations. Even though the main task consisted of focusing one's attention on the fixation target, subjective evaluation of cognitive and physical workload would provide a deeper understanding of user experience rather than just recording behavioural postural responses. Future studies could also include simple cognition task, such as performing simple maths sums or answering general knowledge questions. This would allow future studies to investigate whether higher levels of cognitive processing can influence the magnitude of postural responses to visual motion stimulation.

Another limitation of this study was the set up of the virtual environment. Due to the VR set up, the order of the conditions during the task was always the same, and as such the result could be influenced by the order effect. Also, during the studies, only simple stimulus was presented, vertical black and white stripes. Further research could explore the effects of other, noisier background and different fixation points on postural adjustments. In the same way, other parameters could be considered, such as other foreground objects, precise time lag in motion tracking, foreground stimulus characteristics, and behaviour and stability of object in the foreground, which would provide a further understanding of the environmental effects on postural responses to visual motion stimulus.

7.6.2.2 Additional augmented multisensory cues

Although a high quality virtual reality set up was use that allowed a high level of interactivity, some interferences during task were observed, which was one of the limitations of the study described in chapter 4. Future research could therefore explore whether the beneficial effects of additional augmented sensory cues will be evident when the task is performed in VR environments with very low or very high fidelity levels. For example, different VR technologies could be used to perform the task, such as HMD devices, CAVE system or simpler devices such as a PC with a video game set up. Additionally, a more complex task could be utilised in which participants would be required to give more attention to subsequent stages of the task, rather than perform a simple task as fast as possible. The comparison of the findings between VR environments that have different levels of fidelity as well as different levels of task complexity would provide further understanding of how additional sensory cues could influence task performance and user experience.

The investigation of the positional inaccuracy during the task showed that participants felt increased levels of discomfort, nausea and dizziness; however these symptoms were quite mild; no participant experienced extreme symptoms of simulation sickness where testing had to be terminated. This suggests that projection based VR technologies are suitable platforms for training, however future research could explore scenarios with enhanced manipulation of motion tracking latency to further understand the limits of the negative effects of VR exposure.

As the task in the current study was a manual assembly task, the level of cognitive demand may have been increased. This study collected questionnaire-based responses mainly concerning the experienced sense of presence and simulation sickness but no workload rating scale. Future studies could therefore include this workload rating scale which would enable participants to rate their levels of mental and temporal demand as well as their levels of frustration or satisfaction with the VR system. This would provide further facet in the overall assessment process of the effectiveness of the VR system.

7.6.2.3 Transfer of training

The workload score in this study was relatively low, however the VR training group with multisensory cuing reported higher levels of workload during the VR training. Therefore future research could investigate how different display types (HMD or CAVE) would affect the transfer of skills from virtual to real environments. Additionally these VR technologies could be used to investigate their effectiveness in terms of reported workload levels. Paring the workload levels with different sensory cues using various VR devices could yield some interesting findings.

This study showed transferable performance improvements from virtual to real task, however the performance was assessed through time completion and error penalties. The error penalties were given by the experimenter; however future studies could provide more robust measurements of bolt tightness, similar to those proposed in Lin et al. (2014). Virtual training in this study resulted in a similar performance on the real task, however this finding may not apply to other training tasks. Future studies could further explore other task scenarios and additional methods for assessing differences in behaviour between real and virtual environments to provide comprehensive knowledge about the transfer of skills. Additionally, future research should be directed towards establish a comprehensive knowledge about what is being transferred (different type of tasks), in what way and under what circumstances (high and low workload conditions) from virtual to real world environments.

7.6.2.4 Motion cuing

This study found no significant effects of motion cuing on overall performance, which could be explained by one of the limitations of this study. This limitation was in the technical set up of the task. Although a high fidelity 6-DOF helicopter simulator was used, certain parameters within the simulation were fixed. These were flight altitude and flight speed. The main reason for keeping these parameters constant was to allow recruitment of participants who were less experienced with full motion helicopter simulators. In this task all of the participants had previous simulation flying experience however only a few of them had previously flown on fixed wing aircrafts simulators. Therefore, future studies, where these parameters will be available, should be conducted; which will allow for further exploring of the effects of motion cuing during flight task. The recruitment of professional test pilots who could provide more informed and accurate answers and comments about the advantages and disadvantages of motion cuing in low and high workload situations would be another advancement of future studies.

Pilots recruited in this study reported higher levels of presence and immersion when motion cuing was provided. However as motion cuing conditions failed to reach significance, the notion that the availability of motion cuing generates better task performance can be only partially supported. As mentioned by McCauley (2006), during the simulated flight, pilots dislike no motion; however the addition of simple vibration cues can be as beneficial to performance as full motion platform. Future studies could therefore investigate two different conditions: one with full motion cuing and one with higher intensity vibrations. The manipulation of additional simulator variables such as enhanced/decreased visual cues, additional test points, auditory motion cues and adding the manipulation of washout algorithms could also be included in future studies, as it would enable pilots to observe the surroundings which could benefit to their situational awareness, sense of presence and immersion. In this way the findings from future studies could provide a further understanding and support for the use of full motion platforms for aviation training and research.

References

Adams, R.J., Klowden, D. & Hannaford, B. (2001). Virtual training for a manual assembly task. Haptics-e. 2(2) (http://www.haptics-e.org).

Ahlberg, G., Enochsson, L., Gallagher, A.G., Hedman, L., Hogman, C., McClusky, D.A., Ramel, S., Smith, C.D. & Arvidsson, D. (2007). Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies. *The American journal of surgery*, *193*(6), 797-804.

Albert, W., & Tullis, T. (2013). *Measuring the user experience: collecting, analyzing, and presenting usability metrics*. Newnes.

Akamatsu, M., MacKenzie, I.S. & Hasbroucq, T. (1995). A comparison of tactile, auditory and visual feedback in a pointing task using a mouse-typedevice. *Ergonomics*, *38*(4), 816-827.

Allen, B., Hanley, T., Rokers, B. & Green, C.S. (2016). Visual 3D motion acuity predicts discomfort in 3D stereoscopic environments. *Entertainment Computing*, *13*, 1-9.

Anzanello, M.J. & F. S. Fogliatto. (2011). Learning curve models and applications: Literature review and research direction. *International Journal of Industrial Ergonomics*, 41, 573-583.

Atkins, R. (2008). Enhance virtual reality - Engineering Challenges. Retrieved on 8th of September 2016 from http://www.engineeringchallenges.org/cms/8996/9140.aspx

Aykent, B., Paillot, D., Merienne, F., Fang, Z. & Kemeny, A. (2011). Study of the influence of different washout algorithms on simulator sickness for a driving simulation task. In *ASME 2011 World Conference on Innovative Virtual Reality* (331-341). American Society of Mechanical Engineers.

Balaji, S., Singh, P., Sodergren, M.H., Corker, H.P., Kwasnicki, R.M., Darzi, A. & Paraskeva, P. (2015). A Randomized Controlled Study to Evaluate the Impact of Instrument and Laparoscope Length on Performance and Learning Curve in Single-Incision Laparoscopic Surgery. *Surgical innovation*, *22*(6), 621-628.

Barfield, W. & Hendrix, C. (1995). The effect of update rate on the sense of presence within virtual environments. *Virtual Reality*, *1*(1), 3-15.

Barr, R. A., & Giambra, L. M. (1990). Age-related decrement in auditory selective attention. *Psychology and Aging*, *5*(4), 597.

Baskett, B.J. (2000). *Aeronautical Design Standard performance specification Handling Qualities requirements for military rotorcraft* (No. ADS-33E-PRF). Army Aviation And Missile Command Redstone Arsenal AL.

Bell, J. & Grant, S.C. (2011). Effects of Motion Cueing on Components of Helicopter Pilot Workload. In *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference. Retrieved February, 16*.

Berger, D.R., Terzibas, C., Beykirch, K. & Bülthoff, H.H. (2007). The role of visual cues and whole-body rotations in helicopter hovering control. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit (AIAA 2007), Reston, VA, USA. American Institute of Aeronautics and Astronautics*.

Berthoz, A., Lacour, M., Soechting, J.F. & Vidal, P.P. (1979). The role of vision in the control of posture during linear motion. *Progress in brain research*, *50,* 197-209.

Bertin, R. J. V., Collet, C., Espié, S., & Graf, W. (2005). Objective measurement of simulator sickness and the role of visual-vestibular conflict situations. *Proceedings of the Driving Simulation Conference*, Orlando, FL, 280-293.

Billinghurst, M. (2002). Augmented reality in education. *New horizons for learning*, *12*.

Billinghurst, M., Clark, A. & Lee, G. (2015). A survey of augmented reality. *Foundations and Trends® Human–Computer Interaction*, *8*(2-3), 73-272.

Bland, J.M. & Altman, D.G., (1995). Statistics notes: Calculating correlation coefficients with repeated observations: Part 1—correlation within subjects. *Bmj*, *310*(6977), 446.

Bowman, D.A. & McMahan, R.P. (2007). Virtual reality: how much immersion is enough? *Computer*, *40*(7), 36-43.

Bronstein, A.M. & Buckwell, D. (1997). Automatic control of postural sway by visual motion parallax. *Experimental Brain Research*, *113*(2), 243-248.

Brooks, J.O., Goodenough, R.R., Crisler, M.C., Klein, N.D., Alley, R.L., Koon, B.L., Logan, W.C., Ogle, J.H., Tyrrell, R.A. & Wills, R.F. (2010). Simulator sickness during driving simulation studies. *Accident Analysis and Prevention*, *42*(3), 788-796.

Bruck, S. & Watters, P.A. (2009). Estimating cybersickness of simulated motion using the simulator sickness questionnaire (SSQ): A controlled study. In *Computer Graphics, Imaging and Visualization, 2009. CGIV'09. Sixth International Conference on* (486-488). IEEE.

Burke, J.L., Prewett, M.S., Gray, A.A., Yang, L., Stilson, F.R., Coovert, M.D., Elliot, L.R. & Redden, E. (2006). Comparing the effects of visual-auditory and visual-tactile feedback on user performance: a meta-analysis. In *Proceedings of the 8th international conference on Multimodal interfaces*, 108-117, ACM.

Cao, C.G., Zhou, M., Jones, D.B. & Schwaitzberg, S.D. (2007). Can surgeons think and operate with haptics at the same time? *Journal of Gastrointestinal Surgery*, *11*(11), 1564-1569.

Carlson, P., Peters, A., Gilbert, S.B., Vance, J.M. & Luse, A. (2015). Virtual Training: Learning Transfer of Assembly Tasks. *IEEE transactions on visualization and computer graphics*, *21*(6), 770-782.

Casas, S., Coma, I., Riera, J. V., & Fernández, M. (2015). Motion-cuing algorithms: Characterization of users' perception. *Human factors, 57*(1), 144-162.

Cohen, J. (1992). A power primer. *Psychological bulletin*, *112*(1), 155.

Cooper, N., Milella, F., Cant, I., Pinto, C., White, M. & Meyer, G. The Effects of Multisensory Feedback on Task Performance and the Sense of Presence in Virtual Reality. PloS One, In press.

Cierniak, G., Scheiter, K., & Gerjets, P. (2009). Explaining the split-attention effect: Is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load?. *Computers in Human Behavior*, *25*(2), 315-324.

Curtis, M.K., Dawson, K., Jackson, K., Litwin, L., Meusel, C., Dorneich, M.C., Gilbert, S.B., Kelly, J., Stone, R. & Winer, E. (2015). Mitigating Visually Induced Motion Sickness A virtual hand-eye coordination task. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *59*(1), 1839-1843. SAGE Publications.

Dahlman, J., Falkmer, T. & Forsman, F. (2012). Perceived motion sickness and effects on performance following naval transportation. *Journal of human performance in extreme environments*, *10*(1), 3.

Dahlstrom, N., Dekker, S., Van Winsen, R. & Nyce, J. (2009). Fidelity andvalidity of simulator training. *Theoretical Issues in Ergonomics Science*, *10*(4), 305-314.

Dakin, C.J., Son, G.M.L., Inglis, J.T. & Blouin, J.S. (2007). Frequency response of human vestibular reflexes characterized by stochastic stimuli. *The Journal of physiology*, *583*(3), 1117-1127.

Day, B.L. & Guerraz, M. (2007). Feed-forward versus feedback modulation of human vestibular-evoked balance responses by visual self-motion information. *The Journal of physiology*, *582*(1), 153-161.

De Korte, E.M., Huysmans, M.A., De Jong, A.M., Van de Ven, J.G. & Ruijsendaal, M. (2012). Effects of four types of non-obtrusive feedback on computer behaviour, task performance and comfort. *Applied ergonomics*, *43*(2), 344-353.

De Winter, J.C., Dodou, D. & Mulder, M. (2012). Training effectiveness of whole body flight simulator motion: A comprehensive meta-analysis. *The International Journal of Aviation Psychology*, *22*(2), 164-183.

Deniaud, C., Mestre, D., Honnet, V. & Jeanne, B. (2014). The concept of presence used as a measure for ecological validity in driving simulators. In *Proceedings of the 2014 European Conference on Cognitive Ergonomics*, ACM.

Dinh, H.Q., Walker, N., Hodges, L.F., Song, C. & Kobayashi, A. (1999). Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. In *Virtual Reality, 1999. Proceedings., IEEE* (222-228).

Dobrian, C., & Bevilacqua, F. (2003). Gestural control of music: using the vicon 8 motion capture system. In *Proceedings of the 2003 conference on New interfaces for musical expression*, 161-163, National University of Singapore.

Dokka, K., Kenyon, R.V., Keshner, E.A. & Kording, K.P. (2010). Self versus environment motion in postural control. *PLoS Comput Biol*, *6*(2), p.e1000680.

Dunn, N. & Williamson, A. (2011). Monotony in the rail industry: The role of task demand in mitigating monotony-related effects on performance. In *Ergonomics Australia–HFESA 2011 Conference Edition*, 11.

Efron, B. & Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical science,* 54-75.

Ellis, S.R., Wolfram, A. & Adelstein, B.D. (2002). Large amplitude three-dimensional tracking in augmented environments: a human performance trade-off between system latency and update rate. In *Proc. of the 46th Annual Human Factors and Ergonomics Society meeting*.

Faas, D., Bao, Q., Frey, D.D. & Yang, M.C. (2014). The influence of immersion and presence in early stage engineering designing and building. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, *28*(02), 139-151.

Farrell, M.J., Arnold, P., Pettifer, S., Adams, J., Graham, T. & MacManamon, M. (2003). Transfer of route learning from virtual to real environments. *Journal of Experimental Psychology: Applied*, *9*(4), 219.

Faul, F., Erdfelder, E., Buchner, A. & Lang, A.G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behaviour research methods*, *41*(4), 1149-1160.

Feenstra, P.J., Bos, J.E. & van Gent, R.N. (2011). A visual display enhancing comfort by counteracting airsickness. *Displays*, *32*(4), 194-200.

Feintuch, U., Raz, L., Hwang, J., Josman, N., Katz, N., Kizony, R., Rand, D., Rizzo, A.S., Shahar, M., Yongseok, J. & Weiss, P.L. (2006). Integrating haptic-tactile feedback into a video-capture-based virtual environment for rehabilitation. *CyberPsychology andBehavior*, *9*(2), 129-132.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Sage

Fitts, P.M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology*, *47*(6), 381.

Fjeld, M. & Voegtli, B.M. (2002). Augmented chemistry: An interactive educational workbench. In *Mixed and Augmented Reality, 2002. ISMAR 2002. Proceedings. International Symposium on* (259-321). IEEE.

Friston, S., Karlström, P. & Steed, A. (2016). The Effects of Low Latency on Pointing and Steering Tasks. *IEEE transactions on visualization and computer graphics*, *22*(5), 1605-1615.

Gallagher, A.G., Seymour, N.E., Jordan-Black, J.A., Bunting, B.P., McGlade, K. & Satava, R.M. (2013). Prospective, randomized assessment of transfer of training (ToT) and transfer effectiveness ratio (TER) of virtual reality simulation training for laparoscopic skill acquisition. *Annals of surgery*, *257*(6), 1025-1031.

Gavish, N., Gutiérrez, T., Webel, S., Rodríguez, J., Peveri, M., Bockholt, U. & Tecchia, F. (2015). Evaluating virtual reality and augmented reality training for industrial maintenance and assembly tasks. *Interactive Learning Environments*, *23*(6), 778-798.

Go, T.H., Bürki-Cohen, J., Chung, W.W., Schroeder, J., Saillant, G., Jacobs, S. & Longridge, T. (2003). The effects of enhanced hexapod motion on airline pilot recurrent training and evaluation. In *AIAA Modelling and Simulation Technologies Conference*.

Gray, W.R. (2005). Boundary-avoidance tracking: a new pilot tracking model. In *AIAA Atmospheric Flight Mechanics Conference and Exhibit*, 86-97.

Greuter, S. & Roberts, D.J. (2014). Spacewalk: Movement and interaction in virtual space with commodity hardware. *In Proceedings of the 2014 Conference on Interactive Entertainment*, 1-7. ACM.

Groen, E. L., V. Valenti Clari, M. S., & AW Hosman, R. J. (2001). Evaluation of perceived motion during a simulated take off run. *Journal of Aircraft*, *38*(4), 600-606.

Guimarães, L. D. M., Anzanello, M. J., & Renner, J. S. (2012). A learning curve-based method to implement multifunctional work teams in the Brazilian footwear sector. *Applied ergonomics*, *43*(3), 541-547.

Hamstra, S.J., Brydges, R., Hatala, R., Zendejas, B. & Cook, D.A. (2014). Reconsidering fidelity in simulation-based training. *Academic Medicine*, *89*(3), 387-392.

Hancock, P.A., Williams, G. & Manning, C.M. (1995). Influence of task demand characteristics on workload and performance. *The International Journal of Aviation Psychology*, *5*(1), 63-86.

Hart, S.G. (2006). NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, *50*(9), 904-908. Sage Publications.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology*, *52*, 139-183.

Hays, R.T., Jacobs, J.W., Prince, C. & Salas, E. (1992). Flight simulator training effectiveness: A meta-analysis. *Military psychology*, *4*(2), 63.

Henrysson, A., Billinghurst, M. & Ollila, M. (2005). Face to face collaborative AR on mobile phones. In *Mixed and Augmented Reality, 2005. Proceedings. Fourth IEEE and ACM International Symposium,* 80-89.

Hess, R.A. & Marchesi, F. (2009). Analytical assessment of flight simulator fidelity using pilot models. *Journal of Guidance, Control, and Dynamics, 32*(3), 760-770.

Ho, C., Tan, H.Z. & Spence, C. (2005). Using spatial vibrotactile cues to direct visual attention in driving scenes. *Transportation Research Part F: Traffic Psychology and Behaviour, 8*(6), 397-412.

Hodge, S.J., Perfect, P., Padfield, G.D. & White, M.D. (2015). Optimising the roll-sway motion cues available from a short stroke hexapod motion platform. *The Aeronautical Journal*, *119*(1211), 23-44.

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*(3), 33-33.

Hogema, J., Wentink, M. & Bertollini, G. (2012), January. Effects of Yaw Motion on Driving Behaviour, Comfort and Realism. In *Proceeding of the Driving Simulation Conference,* 149-158, Paris, France.

Hopp, P.J., Smith, C.A., Clegg, B.A. & Heggestad, E.D. (2005). Interruption management: The use of attention-directing tactile cues. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 47*(1), 1-11.

Horey, J.D. (1992). Estimating the impact of restricting simulated motion on transfer of training in rotary wing aircraft. In *Proceedings of the 14th Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)*. San Antonio: TX.

Howarth, P.A. (2011). Potential hazards of viewing 3-D stereoscopic television, cinema and computer games: a review. *Ophthalmic and Physiological Optics*, *31*(2), 111-122.

Jacko, J., Emery, V.K., Edwards, P.J., Ashok, M., Barnard, L., Kongnakorn, T., Moloney, K.P. & Sainfort, F. (2004). The effects of multimodal feedback on older adults' task performance given varying levels of computer experience. *Behaviour and Information Technology*, *23*(4), 247-264.

Jackson, S.A., Kleitman, S. & Aidman, E. (2014). Low cognitive load and reduced arousal impede practice effects on executive functioning, metacognitive confidence and decision making. *PloS one*, *9*(12), p.e115689.

Jeon, M., Davison, B.K., Nees, M.A., Wilson, J. & Walker, B.N. (2009). Enhanced auditory menu cues improve dual task performance and are preferred with in-vehicle technologies. In *Proceedings of the 1st international conference on automotive user interfaces and interactive vehicular applications*, 91-98, ACM.

Jerome, C. J. (2006). Orienting of visual-spatial attention with augmented reality: Effects of spatial and non-spatial multi-modal cues. *Unpublished Doctoral dissertation*, University of Central Florida Orlando, Florida.

Jia, D., Bhatti, A. & Nahavandi, S. (2011). User-centered design and evaluation of an interactive visual-haptic-auditory interface: a user study on assembly. In *ASME 2011 World Conference on Innovative Virtual Reality*, 263-272, American Society of Mechanical Engineers.

Kapur, A., Tzanetakis, G., Virji-Babul, N., Wang, G., & Cook, P. R. (2005). A framework for sonification of vicon motion capture data. In *Conference on Digital Audio Effects*, 47-52.

Kemeny, A. & Panerai, F. (2003). Evaluating perception in driving simulation experiments. *Trends in cognitive sciences*, *7*(1), 31-37.

Kennedy, R.S., Drexler, J. & Kennedy, R.C. (2010). Research in visually induced motion sickness. *Applied ergonomics*, *41*(4), 494-503.

Kennedy, R.S., Lane, N.E., Berbaum, K.S. & Lilienthal, M.G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, *3*(3), 203-220.

Kennedy, R.S., Stanney, K.M. & Dunlap, W.P. (2000). Duration and exposure to virtual environments: sickness curves during and across sessions. *Presence 9*(5), 463–472.

Kern, C. & Refflinghaus, R. (2013). Cross-disciplinary method for predicting and reducing human error probabilities in manual assembly operations. *Total Quality Management and Business Excellence*, *24*(7-8), 847-858.

Khoshelham, K. & Elberink, S.O. (2012). Accuracy and resolution of Kinect depth data for indoor mapping applications. *Sensors*, *12*(2), 1437-1454.

Klein, M.I., Lio, C.H., Grant, R., Carswell, C.M. & Strup, S. (2009). A mental workload study on the 2d and 3d viewing conditions of the da Vinci surgical robot. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 53(18), 1186-1190, SAGE Publications.

Klüver, M., Herrigel, C., Heinrich, C., Schöner, H.P. & Hecht, H. (2016). The behavioral validity of dual-task driving performance in fixed and moving base driving simulators. *Transportation research part F: traffic psychology and behaviour, 37*, 78-96

Kolasinski, E.M. (1995). *Simulator Sickness in Virtual Environments* (No. ARI-TR-1027). Army Research Institute for the Behavioural and Social Sciences ALEXANDRIA VA.

Konrad, R. (2015). What is the vergence-accommodation conflict and how do we fix it? *XRDS: Crossroads, The ACM Magazine for Students*, *22*(1), 52-55.

Kramida, G. (2016). Resolving the Vergence-Accommodation Conflict in Head-Mounted Displays. *IEEE transactions on visualization and computer graphics*, *22*(7), 1912-1931.

Krausert, A. (2015). The timing of training effects: A learning curve perspective. In John Humphreys (Ed.), *Proceedings of the Seventy-fifth Annual Meeting of the Academy of Management.* Online ISSN: 2151-6561.

Laha, B., Sensharma, K., Schiffbauer, J.D. & Bowman, D.A. (2012). Effects of immersion on visual analysis of volume data. *IEEE Transactions on Visualization and Computer Graphics*, *18*(4), 597-606.

Lambooij, M.T., IJsselsteijn, W.A. & Heynderickx, I. (2007). Visual discomfort in stereoscopic displays: a review. In *Electronic Imaging 2007* (64900I-64900I). International Society for Optics and Photonics.

Lapré, M. A., & Nembhard, I. M. (2010). Inside the organizational learning curve: Understanding the organizational learning process. Foundations and Trends in Technology, *Information and Operations Management, 4,* 1-103.

Lee, C., Bonebrake, S., Bowman, D.A. & Höllerer, T. (2010). The role of latency in the validity of AR simulation. In *2010 IEEE Virtual Reality Conference (VR)*, 11-18.

Lee, M., Billinghurst, M., Baek, W., Green, R. & Woo, W. (2013). A usability study of multimodal input in an augmented reality environment. *Virtual Reality*, *17*(4), 293-305

Leung, G.T., Yucel, G. & Duffy, V.G. (2010). The effects of virtual industrial training on mental workload during task performance. *Human Factors and Ergonomics in Manufacturing and Service Industries*, *20*(6), 567-578.

Lin, G., Pang, H., Zhang, W., Wang, D., & Feng, L. (2014). A self-decoupled three-axis force sensor for measuring the wheel force. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, *228*(3), 319-334.

Lin, 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 *Virtual Reality, 2002. Proceedings. IEEE,* 164-171.

Loftin, R.B., Scerbo, M.W., McKenzie, R., Catanzaro, J.M., Bailey, N.R., Phillips, M.A. & Perry, G. (2004). *Training in peacekeeping operations using virtual environments*. Conference paper, VA Virginia Modelling Analysis and Simulation Center. University Suffolk

Lu, L. & Jump, M. (2014). Multiloop Pilot Model for Boundary-Triggered Pilot-Induced Oscillation Investigations. *Journal of Guidance, Control, and Dynamics*, *37*(6), 1863-1879.

Lubeck, A.J., Bos, J.E. & Stins, J.F. (2015). Motion in images is essential to cause motion sickness symptoms, but not to increase postural sway. *Displays*, *38*, 55-61.

Lugrin, J.L., Wiebusch, D., Latoschik, M.E. & Strehler, A. (2013). Usability benchmarks for motion tracking systems. In *Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology*, 49-58, ACM.

Ma, R. & Kaber, D.B. (2006). Presence, workload and performance effects of synthetic environment design factors. *International Journal of Human-Computer Studies, 64*(6), 541-552.

Martínez, J., Martínez, D., Molina, J.P. & Garcia, A. (2011). Comparison of force and vibrotactile feedback with direct stimulation for texture recognition. In *Cyberworlds (CW), 2011 International Conference,* 62-68, IEEE.

Mayol- Cuevas, W. W., Davison, A. J., Tordoff, B. J., & Murray, D. W. (2005). Applying active vision and slam to wearables. *Robotics Research*, 325-334.

McCauley, M. E. (2006). *Do Army helicopter training simulators need motion bases?* Monterey, California. Naval Postgraduate School, Monterey CA Department of Operations Research.

McMahan, R. P., Gorton, D., Gresock, J., McConnell, W., & Bowman, D. A. (2006). Separating the effects of level of immersion and 3D interaction techniques. In *Proceedings of the ACM symposium on Virtual reality software and technology*, 108-111, ACM.

Meehan, M., Razzaque, S., Whitton, M.C. & Brooks, F.P. (2003). Effect of latency on presence in stressful virtual environments. In *virtual reality, 2003. Proceedings. IEEE* ,141-148.

Menzies, R.J., Rogers, S.J., Phillips, A.M., Chiarovano, E., de Waele, C., Verstraten, F.A.J. & MacDougall, H. (2016). An objective measure for the visual fidelity of virtual reality and the risks of falls in a virtual environment. *Virtual Reality*, 1-9.

Mergner, T., Schweigart, G., Maurer, C. & Blümle, A. (2005). Human postural responses to motion of real and virtual visual environments under different support base conditions. *Experimental Brain Research, 167*(4), 535-556.

Mestre, D., Fuchs, P., Berthoz, A. & Vercher, J.L. (2006). Immersion et présence. *Le traité de la réalité virtuelle. Paris: Ecole des Mines de Paris*, 309-38.

Meyer, G., Clarke, E. & Robotham, T. (2012). Multisensory interactions in the automatic control of postural sway. *Seeing and Perceiving*, *25*, 77-77.

Meyer, G.F., Shao, F., White, M.D., Hopkins, C. & Robotham, A.J. (2013). Modulation of visually evoked postural responses by contextual visual, haptic and auditory information: a 'virtual reality check'. *PloS one, 8*(6), doi:10.1371/journal.pone.0067651

Miles, H. C., Pop, S. R., Watt, S. J., Lawrence, G. P., & John, N. W. (2012). A review of virtual environments for training in ball sports. *Computers & Graphics, 36*(6), 714-726.

Milleville-Pennel, I. & Charron, C. (2015). Do mental workload and presence experienced when driving a real car predispose drivers to simulator sickness? An exploratory study. *Accident Analysis and Prevention*, *74*, 192-202.

Moorthy, K., Munz, Y., Sarker, S. K., & Darzi, A. (2003). Objective assessment of technical skills in surgery. *BMJ: British Medical Journal, 327*(7422), 1032.

Mowbray, C.T., Holter, M.C., Teague, G.B. & Bybee, D. (2003). Fidelity criteria: Development, measurement, and validation. *American journal of evaluation*, *24*(3), 315-340.

Mulder, F.A., Verlinden, J.C. & Dukalski, R.R. (2012). The effect of motion on presence during virtual sailing for advanced training. In *Presence 2012 ISPR 2012, Philadelphia (USA), 24-26 Oct., 2012*. International Society for Presence Research.

Murata, A. (2004). Effects of duration of immersion in a virtual reality environment on postural stability. *International Journal of Human-Computer Interaction*, *17*(4), 463-477.

Murray, N., Goulermas, J.Y. & Fernando, T. (2003). Visual tracking for a virtual environment. In *Proceedings of HCI International, 1*, 1198-1202.

Nam, C.S., Shu, J. & Chung, D. (2008). The roles of sensory modalities in collaborative virtual environments (CVEs). *Computers in Human Behaviour*, *24*(4), 1404-1417.

Navab, N., Feuerstein, M. & Bichlmeier, C. (2007). Laparoscopic virtual mirror new interaction paradigm for monitor based augmented reality. In *Virtual Reality Conference, 2007. VR'07. IEEE*, 43-50.

Nichols, S., Haldane, C. & Wilson, J.R. (2000). Measurement of presence and its consequences in virtual environments. *International Journal of Human-Computer Studies*, *52*(3), 471-491.

Novak, D., Mihelj, M. & Munih, M. (2012). A survey of methods for data fusion and system adaptation using autonomic nervous system responses in physiological computing. *Interacting with computers*, *24*(3), 154-172.

Nowak, K., Oron, A., Szymaszek, A., Leminen, M., Näätänen, R., & Szelag, E. (2016). Electrophysiological Indicators of the Age-Related Deterioration in the Sensitivity to Auditory Duration Deviance. *Frontiers in aging neuroscience, 8*.

Oakley, B. (2009). The effects of multimodal feedback and age on a mouse pointing task, *Unpublished Doctoral dissertation,* University of Central Florida Orlando, Florida).

Oikonomidis, I., Kyriazis, N. & Argyros, A.A. (2011). Efficient model-based 3D tracking of hand articulations using Kinect. In *BmVC*, *1*(2), 3.

Oskarsson, P.A., Eriksson, L. & Carlander, O. (2012). Enhanced perception and performance by multimodal threat cueing in simulated combat vehicle. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *54*(1), 122-137.

Oviatt, S., Coulston, R. & Lunsford, R. (2004). When do we interact multimodally?- Cognitive load and multimodal communication patterns. In *Proceedings of the 6th international conference on Multimodal interfaces*, 129-136, ACM.

Pallant, J. (2013). *SPSS survival manual*. McGraw-Hill Education (UK).

Palmisano, S. (2002). Consistent stereoscopic information increases the perceived speed of vection in depth. *Perception*, *31*(4), 463–480.

Papadakis, G., Mania, K. & Koutroulis, E. (2011). A system to measure, control and minimize end-to-end head tracking latency in immersive simulations. In *Proceedings of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry*, 581-584, ACM.

Park, J., MacRae, H., Musselman, L.J., Rossos, P., Hamstra, S.J., Wolman, S. & Reznick, R.K. (2007). Randomized controlled trial of virtual reality simulator training: transfer to live patients. *The American journal of surgery*, *194*(2), 205-211.

Pasma, D.L., Grant, S.C., Gamble, M., Kruk, R.V. & Herdman, C.M. (2011). Utility of motion and motion-cueing to support simulated in-flight rotary-wing emergency training. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *55*(1), 133-137, SAGE Publications.

Perfect, P., White, M., Padfield, G.D., Gubbels, A. & Berryman, A. (2010). Integrating predicted and perceived fidelity for flight simulators. *In 36th European Rotorcraft Forum*, 1-20, Paris: CEAS.

Peterka, R.J. (2002). Sensorimotor integration in human postural control. *Journal of neurophysiology*, *88*(3), 1097-1118.

Piekarski, W. & Thomas, B. (2002). ARQuake: the outdoor augmented reality gaming system. *Communications of the ACM*, *45*(1), 36-38.

Pietschmann, D. & Rusdorf, S. (2014). Matching levels of task difficulty for different modes of presentation in a VR table tennis simulation by using assistance functions and regression analysis. In *International Conference on Virtual, Augmented and Mixed Reality*, 406-417, Springer International Publishing.

Ramsamy, P., Haffegee, A., Jamieson, R. & Alexandrov, V. (2006). Using haptics to improve immersion in virtual environments. In *International Conference on Computational Science*, 603-609, Springer Berlin Heidelberg.

Ramsay, C.R., Grant, A.M., Wallace, S.A., Garthwaite, P.H., Monk, A.F. & Russell, I.T. (2001). *Statistical assessment of the learning curves of health technologies*. Health Technology Assessment Programme.

Reason, J.T. & Brand, J.J. (1975). *Motion sickness*. Academic press.

Richard, P., Chamaret, D., Inglese, F. X., Lucidarme, P., & Ferrier, J. L. (2006). Human-scale virtual environment for product design: Effect of sensory substitution. *The International Journal of Virtual Reality*, *2*(5), 37-44.

Riccio, G. E., & Stoffregen, T. A. (1991). An ecological theory of motion sickness and postural instability. *Ecological psychology*, *3*(3), 195-240.

Riecke, B. E., Schulte-Pelkum, J., Avraamides, M. N., Heyde, M. V. D., & Bülthoff, H. H. (2006). Cognitive factors can influence self-motion perception (vection) in virtual reality. *ACM Transactions on Applied Perception (TAP), 3*(3), 194-216.

Rolland, J.P., Davis, L. & Baillot, Y. (2001). A survey of tracking technology for virtual environments. *Fundamentals of wearable computers and augmented reality, 1*(1), 67-112.

Roscoe, A.H. (1992). Assessing pilot workload: Why measure heart rate, HRV and respiration? *Biological psychology*, *34*(2), 259-287.

Roscoe, A. H., & Ellis, G. A. (1990). *A subjective rating scale for assessing pilot workload in flight: A decade of practical use* (No. RAE-TR-90019). Royal Aerospace Establishment Farnborough (United Kingdom).

Rubio, S., Díaz, E., Martín, J. & Puente, J.M. (2004). Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and workload profile methods. *Applied Psychology*, *53*(1), 61-86.

Salas, E., Wildman, J.L. & Piccolo, R.F. (2009). Using simulation-based training to enhance management education. *Academy of Management Learning and Education*, *8*(4), 559-573.

Sanchez, A., Mat, S., Goh, Y., & Case, K. (2013). Human variability, task complexity and motivation contribution in manufacturing. In *Proceedings of the 11th International Conference on Manufacturing Research (ICMR2013),* Cranfield University, UK, 325-330.

Santangelo, V. & Spence, C. (2007). Multisensory cues capture spatial attention regardless of perceptual load. *Journal of Experimental Psychology: Human Perception and Performance, 33*(6), 1311.

Scerbo, M.W. & Dawson, S. (2007). High fidelity, high performance? *Simulation in Healthcare*, *2*(4), 224-230.

Schäfer, P., Koller, M., Diemer, J. & Meixner, G. (2015). Development and Evaluation of a Virtual Reality-System with integrated Tracking of Extremities under the Aspect of Acrophobia. In *SAI Intelligent Systems Conference (IntelliSys),* 408-417, IEEE.

Schoonmaker, R.E. & Cao, C.G. (2006). Vibrotactile feedback enhances force perception in minimally invasive surgery. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *50*(10), 1029-1033, SAGE Publications.

Schroeder, J.A. (1999). Helicopter flight simulation motion platform requirements. NASA/TP-1999-208766.

Schuchardt, P. & Bowman, D.A. (2007). The benefits of immersion for spatial understanding of complex underground cave systems. In *Proceedings of the 2007 ACM symposium on Virtual reality software and technology*, 121-124, ACM.

Seaborn, K., Riecke, B.E. & Antle, A.N. (2010). Exploring the interplay of visual and haptic modalities in a pattern-matching task. In *Haptic Audio-Visual Environments and Games (HAVE), 2010 IEEE International Symposium on*, 1-6, IEEE.

Seay, A.F., Krum, D.M., Hodges, L. & Ribarsky, W. (2001). Simulator sickness and presence in a high FOV virtual environment. In*Virtual Reality, 2001. Proceedings. IEEE*, 299-300.

Seymour, N.E., Gallagher, A.G., Roman, S.A., O’Brien, M.K., Bansal, V.K., Andersen, D.K. & Satava, R.M. (2002). Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Annals of surgery*, *236*(4), 458-464.

Shi, Y., Ruiz, N., Taib, R., Choi, E. & Chen, F. (2007). Galvanic skin response (GSR) as an index of cognitive load. In *CHI'07 extended abstracts on Human factors in computing systems*, 2651-2656, ACM.

Shibata, T., Kim, J., Hoffman, D.M. & Banks, M.S. (2011). Visual discomfort with stereo displays: Effects of viewing distance and direction of vergence-accommodation conflict. In *ISANDT/SPIE Electronic Imaging*, 78630P-78630P. International Society for Optics and Photonics.

Shin, B., Venkatramani, R., Borker, P., Olch, A., Grimm, J. & Wong, K. (2013). Spatial accuracy of a low cost high resolution 3D surface imaging device for medical applications.

Siegler, I., Reymond, G., Kemeny, A. & Berthoz, A. (2001). Sensorimotor integration in a driving simulator: contributions of motion cueing in elementary driving tasks. In *Proceedings of Driving Simulation Conference*, 21-32.

Singh, I.L., Sharma, H.O. & Singh, A.L. (2005). Effect of training on workload in flight simulation task performance. *Journal of the Indian Academy of applied Psychology*, *31*(1-2), 81-90.

Slater, M. (2003). A note on presence terminology. *Presence connect*, *3*(3), 1-5.

Slater, M., Linakis, V., Usoh, M., Kooper, R. & Street, G. (1996). Immersion, presence, and performance in virtual environments: An experiment with tri-dimensional chess. In *ACM virtual reality software and technology (VRST)*, 163-172). New York, NY: ACM Press.

Sparto, P. J., Whitney, S. L., Hodges, L. F., Furman, J. M., & Redfern, M. S. (2004). Simulator sickness when performing gaze shifts within a wide field of view optic flow environment: preliminary evidence for using virtual reality in vestibular rehabilitation. *Journal of NeuroEngineering and Rehabilitation*, *1*(1), 14.

Spelmezan, D., Jacobs, M., Hilgers, A. & Borchers, J. (2009). Tactile motion instructions for physical activities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2243-2252, ACM.

Sprague, D.W., Po, B.A. & Booth, K.S. (2006). The importance of accurate VR head registration on skilled motor performance. In *Proceedings of Graphics Interface 2006*, 131-137, Canadian Information Processing Society.

Stanney, K.M., Kingdon, K.S., Graeber, D. & Kennedy, R.S. (2002). Human performance in immersive virtual environments: Effects of exposure duration, user control, and scene complexity. *Human Performance*, *15*(4), 339-366.

Steed, A. (2008). A simple method for estimating the latency of interactive, real-time graphics simulations. In *Proceedings of the 2008 ACM symposium on Virtual reality software and technology*, 123-129, ACM.

Stein, M. & Robinski, M. (2012). Simulator sickness in flight simulators of the German armed forces. *Aviation Psychology and Applied Human Factors*, *2*(1), 11-19.

Stevens, J.A. & Kincaid, J.P. (2015). The Relationship between Presence and Performance in Virtual Simulation Training. *Open Journal of Modelling and Simulation*, *3*(02), 41.

Stone, R.T., Watts, K.P., Zhong, P. & Wei, C.S. (2011). Physical and cognitive effects of virtual reality integrated training. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *53*(5), 558-572.

Stothard, P. (2008). Developing an Enhanced VR Simulation Capability for the Coal Mining Industry. *UNSW*, *Technological report*, School of Mining Engineering.

Streepey, J.W., Kenyon, R.V. & Keshner, E.A. (2007). Field of view and base of support width influence postural responses to visual stimuli during quiet stance. *Gait and posture*, *25*(1), 49-55.

Suma, E., Finkelstein, S., Reid, M., Babu, S., Ulinski, A. & Hodges, L.F. (2010). Evaluation of the cognitive effects of travel technique in complex real and virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, *16*(4), 690-702.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and instruction*, *4*(4), 295-312.

Swindells, C., Po, B.A., Hajshirmohammadi, I., Corrie, B., Dill, J.C., Fisher, B.D. & Booth, K.S. (2004). Comparing CAVE, wall, and desktop displays for navigation and way finding in complex 3D models. In *Computer Graphics International, 2004. Proceedings*, 420-427, IEEE.

Tadin, D. (2015). Suppressive mechanisms in visual motion processing: From perception to intelligence. *Vision research*, *115*, 58-70.

Torres-Gómez, A. & Mayol-Cuevas, W. (2014). Recognition and reconstruction of transparent objects for augmented reality. In *Mixed and Augmented Reality (ISMAR), 2014 IEEE International Symposium on*, 129-134, IEEE.

Treleaven, J., Battershill, J., Cole, D., Fadelli, C., Freestone, S., Lang, K. & Sarig-Bahat, H. (2015). Simulator sickness incidence and susceptibility during neck motion-controlled virtual reality tasks. *Virtual Reality*, *19*(3-4), 267-275.

Vaden, E.A. & Hall, S. (2005). The effect of simulator platform motion on pilot training transfer: A meta-analysis. *The International Journal of Aviation Psychology*, *15*(4), 375-393.

Van Der Linden, J., Johnson, R., Bird, J., Rogers, Y. & Schoonderwaldt, E. (2011). Buzzing to play: lessons learned from an in the wild study of real-time vibrotactile feedback. In *Proceedings of the SIGCHI Conference on Human factors in Computing Systems*, 533-542, ACM.

Van der Straaten, P. (2000). Interaction affecting the sense of presence in virtual reality. *Delft University of Technology, Faculty of Information Technology and System*, *67*.

Van Erp, J.B. & Van Veen, H.A. (2004). Vibrotactile in-vehicle navigation system. *Transportation Research Part F: Traffic Psychology and Behaviour*,*7*(4), 247-256.

Van Krevelen, D., & Poelman, R. (2007). Augmented Reality: Technologies, Applications, and Limitations. *The International Journal of Virtual Reality*, *9*(2), 1–20.

Vastenburg, M.H., Ross, P.R. & Keyson, D.V. (2007). A user experience-based approach to home atmosphere control. *Universal Access in the Information Society*, *6*(1), 1-13.

Viaud-Delmon, I., Warusfel, O., Seguelas, A., Rio, E. & Jouvent, R. (2006). High sensitivity to multisensory conflicts in agoraphobia exhibited by virtual reality. *European Psychiatry*, *21*(7), 501-508.

Vicon (2016) Vicon Motion Systems Ltd., *Vicon Motion Capture System*, http://www.vicon.com Accessed November 2016.

Villard, S.J., Flanagan, M.B., Albanese, G.M. & Stoffregen, T.A. (2008). Postural instability and motion sickness in a virtual moving room. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(2), 332-345.

Vitense, H.S., Jacko, J.A. & Emery, V.K. (2003). Multimodal feedback: an assessment of performance and mental workload. *Ergonomics, 46*(1-3), 68-87.

Wang, Y., White, M., Owen, I., Hodge, S. & Barakos, G. (2013). Effects of visual and motion cues in flight simulation of ship-borne helicopter operations. *CEAS Aeronautical Journal*, *4*(4), 385-396.

Webb, C.M., Bass, J.M., Johnson, D.M., Kelley, A.M., Martin, C.R. & Wildzunas, R.M. (2009). Simulator sickness in a helicopter flight training school. *Aviation, space, and environmental medicine*, *80*(6), 541-545.

Weber, B., Schätzle, S., Hulin, T., Preusche, C. & Deml, B. (2011). Evaluation of a vibrotactile feedback device for spatial guidance. In *World Haptics Conference (WHC), 2011 IEEE*, 349-354.

White, M.D., Perfect, P., Padfield, G.D., Gubbels, A.W. & Berryman, A.C. (2012). Acceptance testing and commissioning of a flight simulator for rotorcraft simulation fidelity research. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, p.0954410012439816.

Wickens, C.D. (2002). Multiple resources and performance prediction. *Theoretical issues in ergonomics science*, *3*(2), 159-177.

Wierinck, E., Puttemans, V., Swinnen, S. & Steenberghe, D. (2005). Effect of augmented visual feedback from a virtual reality simulation system on manual dexterity training. *European Journal of Dental Education*, *9*(1), 10-16.

Witmer, B.G. & Singer, M.J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and virtual environments*, *7*(3), 225-240.

Witmer, B.G., Jerome, C.J. & Singer, M.J. (2005). The factor structure of the presence questionnaire. *Presence*, *14*(3), 298-312.

Wong, L. T. (2014) Self-evaluation vs. objective performance measures: evaluation of fidelity, presence and training transfer in two helicopter simulator tasks. MPhil Dissertation paper, University of Liverpool.

Wright, T.P. (1936). Factors affecting the cost of airplanes. *Journal of the aeronautical sciences* 3, 122-128.

Yeh, Y.Y. & Wickens, C.D. (1988). Dissociation of performance and subjective measures of workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *30*(1), 111-120.

Young, S.D., Adelstein, B.D. & Ellis, S.R. (2007). Demand characteristics in assessing motion sickness in a virtual environment: or does taking a motion sickness questionnaire make you sick? *IEEE transactions on visualization and computer graphics*, *13*(3), 422-428

Youngblut, C. & Huie, O. (2003). The relationship between presence and performance in virtual environments: Results of a VERTS study. In*Virtual Reality, 2003. Proceedings. IEEE*, 277-278.

Yunus, F. A. N., Baser, J. A., Masran, S. H., Razali, N., & Rahim, B. (2011). Virtual reality simulator developed welding technology skills. *Journal of Modern Education Review*, *1*(1), 57-62.

Zielinski, D.J., Rao, H.M., Sommer, M.A. & Kopper, R. (2015). Exploring the effects of image persistence in low frame rate virtual environments. In *Virtual Reality (VR), 2015 IEEE*, 19-26.

Zeelenberg, R., & Pecher, D. (2015). A method for simultaneously counterbalancing condition order and assignment of stimulus materials to conditions. *Behavior research methods*, *47*(1), 127-133.

Zollhöfer, M., Martinek, M., Greiner, G., Stamminger, M. & Süßmuth, J. (2011). Automatic reconstruction of personalized avatars from 3D face scans. *Computer Animation and Virtual Worlds*, *22*(2-3), 195-202.

Appendices

Appendix A

Consent form

Title of Project: Investigation of sensory cues in virtual reality environments

Name of lead Researcher: Dr G Meyer

Email: Georg@liverpool.ac.uk
Telephone: 01517942579

Address: School of Psychology
Eleanor Rathbone Building,
Bedford Street South,
Liverpool,
L69 7ZA

Please tick initial box

1. I confirm that I have read and understood the information sheet for the above study. ☐

2. I confirm that I have normal or corrected to normal vision. ☐

3. I understand that my participation is voluntary and that I am free to withdraw at any time without giving a reason. ☐

4. I understand that none of my personal details will be recorded and that my responses are anonymous. ☐

5. I confirm that I have been given the opportunity to ask questions and have them answered. ☐

6. I understand that I can have access to the data, and ask for it to be destroyed if I so wish ☐

7. I agree to take part in the above study. ☐

____________________	____________	____________________
Name of Participant	Date	Signature
____________________	____________	____________________
Researcher	Date	Signature

Appendix B

Immersive Tendencies Questionnaire (ITQ)

(Witmer & Singer, Version 3.01, September 1996)* Revised by the UQO Cyberpsychology Lab (2004)
Name: ______________________ Age:____ Sex: male/female
Indicate your preferred answer by marking an "X" in the appropriate box of the seven point scale. Please consider the entire scale when making your responses, as the intermediate levels may apply. For example, if your response is once or twice, the second box from the left should be marked. If your response is many times but not extremely often, then the sixth (or second box from the right) should be marked.

1. Do you easily become deeply involved in movies or TV dramas?
|________|________|__________|_________|_________|_________|_________|
NEVER OCCASIONALLY OFTEN

2. Do you ever become so involved in a television program or book that people have problems getting your attention?
|________|________|__________|_________|_________|_________|_________|
NEVER OCCASIONALLY OFTEN

3. How mentally alert do you feel at the present time?
|________|________|__________|_________|_________|_________|_________|
NOT ALERT MODERATELY FULLYALERT

4. Do you ever become so involved in a movie that you are not aware of things happening around you?
|________|________|__________|_________|_________|_________|_________|
NEVER OCCASIONALLY OFTEN

5. How frequently do you find yourself closely identifying with the characters in a story line?
|________|________|__________|_________|_________|_________|_________|
NEVER OCCASIONALLY OFTEN

6. Do you ever become so involved in a video game that it is as if you are inside the game rather than moving a joystick and watching the screen?

|________|________|__________|_________|_________|_________|_________|
NEVER OCCASIONALLY OFTEN

7. How physically fit do you feel today?
|________|________|__________|_________|_________|_________|_________|
NOT FIT MODERATELY FIT EXTREMELY FIT

8. How good are you at blocking out external distractions when you are involved in something?
|________|________|__________|_________|_________|_________|_________|
NOT VERY SOMEWHAT GOOD VERY GOOD

9. When watching sports, do you ever become so involved in the game that you react as if you were one of the players?
|________|________|__________|_________|_________|_________|_________|
NEVER OCCASIONALLY OFTEN

10. Do you ever become so involved in a daydream that you are not aware of things happening around you?
|________|________|__________|_________|_________|_________|_________|
NEVER OCCASIONALLY OFTEN

11. Do you ever have dreams that are so real that you feel disoriented when you awake?
|________|________|__________|_________|_________|_________|_________|
NEVER OCCASIONALLY OFTEN

12. When playing sports, do you become so involved in the game that you lose track of time?
|________|________|__________|_________|_________|_________|_________|
NEVER OCCASIONALLY OFTEN

13. How well do you concentrate on enjoyable activities?
|________|________|__________|_________|_________|_________|_________|
NOT AT ALL MODERATELY WELL VERY WELL

14. How often do you play arcade or video games? (OFTEN should be taken to mean every day or every two days, on average.)
|________|________|__________|_________|_________|_________|_________|
NEVER OCCASIONALLY OFTEN

15. Have you ever gotten excited during a chase or fight scene on TV or in the movies?
|________|________|__________|_________|_________|_________|_________|
NEVER OCCASIONALLY OFTEN

16. Have you ever gotten scared by something happening on a TV show or in a movie?
|________|________|__________|_________|_________|_________|_________|
NEVER OCCASIONALLY OFTEN

17. Have you ever remained apprehensive or fearful long after watching a scary movie?
|________|________|__________|_________|_________|_________|_________|
NEVER OCCASIONALLY OFTEN

18. Do you ever become so involved in doing something that you lose all track of time?
|________|________|__________|_________|_________|_________|_________|
NEVER OCCASIONALLY OFTEN

Appendix C

Simulation Sickness Questionnaire

No______________ Date____________________

SIMULATOR SICKNESS QUESTIONNAIRE

Kennedy, Lane, Berbaum, & Lilienthal (1993)***

Instructions: Circle below how much each symptom below is affecting you <u>right now.</u>

1. General discomfort	<u>None</u>	<u>Slight</u>	<u>Moderate</u>	<u>Severe</u>
2. Fatigue	<u>None</u>	<u>Slight</u>	<u>Moderate</u>	<u>Severe</u>
3. Headache	<u>None</u>	<u>Slight</u>	<u>Moderate</u>	<u>Severe</u>
4. Eye strain	<u>None</u>	<u>Slight</u>	<u>Moderate</u>	<u>Severe</u>
5. Difficulty focusing	<u>None</u>	<u>Slight</u>	<u>Moderate</u>	<u>Severe</u>
6. Salivation increasing	<u>None</u>	<u>Slight</u>	<u>Moderate</u>	<u>Severe</u>
7. Sweating	<u>None</u>	<u>Slight</u>	<u>Moderate</u>	<u>Severe</u>
8. Nausea	<u>None</u>	<u>Slight</u>	<u>Moderate</u>	<u>Severe</u>
9. Difficulty concentrating	<u>None</u>	<u>Slight</u>	<u>Moderate</u>	<u>Severe</u>
10. « Fullness of the Head »	<u>None</u>	<u>Slight</u>	<u>Moderate</u>	<u>Severe</u>
11. Blurred vision	<u>None</u>	<u>Slight</u>	<u>Moderate</u>	<u>Severe</u>
12. Dizziness with eyes open	<u>None</u>	<u>Slight</u>	<u>Moderate</u>	<u>Severe</u>
13. Dizziness with eyes closed	<u>None</u>	<u>Slight</u>	<u>Moderate</u>	<u>Severe</u>
14. *Vertigo	<u>None</u>	<u>Slight</u>	<u>Moderate</u>	<u>Severe</u>
15. **Stomach awareness	<u>None</u>	<u>Slight</u>	<u>Moderate</u>	<u>Severe</u>
16. Burping	<u>None</u>	<u>Slight</u>	<u>Moderate</u>	<u>Severe</u>

* Vertigo is experienced as loss of orientation with respect to vertical upright.
** Stomach awareness is usually used to indicate a feeling of discomfort which is just short of nausea.

Appendix D

Example of a short presence questionnaire (adopted from PQ questionnaire) used in wheel change task in VR with augmented cues (VEC) study (Chapter 4).

Short presence questionnaire

Participant number________ Sensory cues________________________

Please mark X close to your agreement with the statement.

1. The interaction with the environment seems natural.

0________________________________100
Strongly Disagree Strongly Agree

2. I felt that sensory cues helped me in completing the assigned task.

0________________________________100
Strongly Disagree Strongly Agree

3. The sense of immersion inside the virtual environment was compelling.

0________________________________100
Strongly Disagree Strongly Agree

4. The virtual environment felt real.

0________________________________100
Strongly Disagree Strongly Agree

5. The virtual environment was responsive to my actions.

0________________________________100
Strongly Disagree Strongly Agree

6. Doing the task felt enjoyable.

0________________________________100
Strongly Disagree Strongly Agree

7. I experienced general discomfort.

0________________________________100
Strongly Disagree Strongly Agree

If **Agree** please state why:
Nausea sweating stomach awareness dizziness difficulty concentrating
Vertigo fullness in the head fatigue headache difficulty focusing

Appendix E

Presence Questionnaire (PQ)

(Witmer & Singer, 2005)

Name:_________________ _ Age:_____ **SEX:** male / female

Characterize your experience in the environment, by marking an "X" in the appropriate box. Please consider the entire scale when making your responses, as the intermediate levels may apply. Answer the questions independently in the order that they appear. Do not skip questions or return to a previous question to change your answer.

WITH REGARD TO THE EXPERIENCED ENVIRONMENT

1. How much were you able to control events?
|________|________|__________|_________|_________|_________|_________|
NOT AT ALL SOMEWHAT COMPLETELY

2. How responsive was the environment to actions that you initiated (or performed)?
|________|________|__________|_________|_________|_________|_________|
NOT MODERATELY COMPLETELY
RESPONSIVE RESPONSIVE RESPONSIVE

3. How natural did your interactions with the environment seem?
|________|________|__________|_________|_________|_________|_________|
EXTREMELY BORDERLINE COMPLETELY
ARTIFICIAL NATURAL

4. How much did the visual aspects of the environment involve you?
|________|________|__________|_________|_________|_________|_________|
NOT AT ALL SOMEWHAT COMPLETELY

5. How much did the auditory aspects of the environment involve you?
|________|________|__________|_________|_________|_________|_________|
NOT AT ALL SOMEWHAT COMPLETELY

6. How natural was the mechanism which controlled movement through the environment?
|________|________|__________|_________|_________|_________|_________|
EXTREMELY BORDERLINE COMPLETELY
ARTIFICIAL NATURAL

7. How compelling was your sense of objects moving through space?
|________|________|__________|_________|_________|_________|_________|
NOT MODERATELY VERY
CONSISTENT CONSISTENT CONSISTENT

8. How much did your experiences in the virtual environment seem consistent with your real world experiences?
|________|________|__________|_________|_________|_________|_________|

NOT MODERATELY VERY
CONSISTENT CONSISTENT CONSISTENT

9. Were you able to anticipate what would happen next in response to the actions that you performed?
|________|________|__________|_________|_________|_________|_________|
NOT AT ALL SOMEWHAT COMPLETELY

10. How completely were you able to actively survey or search the environment using vision?
|________|________|__________|_________|_________|_________|_________|
NOT AT ALL SOMEWHAT COMPLETELY

11. How well could you identify sounds?
|________|________|__________|_________|_________|_________|_________|
NOT AT ALL PRETTY VERY
CLOSELY CLOSELY

12. How well could you localize sounds?
|________|________|__________|_________|_________|_________|_________|
NOT AT ALL PRETTY VERY
CLOSELY CLOSELY

13. How well could you actively survey or search the virtual environment using touch?
|________|________|__________|_________|_________|_________|_________|
NOT AT ALL PRETTY VERY
CLOSELY CLOSELY

14. How compelling was your sense of moving around inside the virtual environment?
|________|________|__________|_________|_________|_________|_________|
NOT MODERATELY VERY
COMPELLING COMPELLING COMPELLING

15. How closely were you able to examine objects?
|________|________|__________|_________|_________|_________|_________|
NOT AT ALL PRETTY VERY
CLOSELY CLOSELY

16. How well could you examine objects from multiple viewpoints?
|________|________|__________|_________|_________|_________|_________|
NOT AT ALL SOMEWHAT EXTENSIVELY

17. How well could you move or manipulate objects in the virtual environment?
|________|________|__________|_________|_________|_________|_________|
NOT AT ALL SOMEWHAT EXTENSIVELY

18. How involved were you in the virtual environment experience?
|________|________|__________|_________|_________|_________|_________|
NOT MILDLY COMPLETELY
INVOLVED INVOLVED ENGROSSED

19. How much delay did you experience between your actions and expected outcomes?

|________|________|__________|_________|_________|_________|_________|
NO DELAYS MODERATE LONG
DELAYS DELAYS

20. How quickly did you adjust to the virtual environment experience?

|________|________|__________|_________|_________|_________|_________|
NOT AT ALL SLOWLY LESS THAN
ONE MINUTE

21. How proficient in moving and interacting with the virtual environment did you feel at the end of the experience?

|________|________|__________|_________|_________|_________|_________|
NOT REASONABLY VERY
PROFICIENT PROFICIENT PROFICIENT

22. How much did the visual display quality interfere or distract you from performing assigned tasks or required activities?

|________|________|__________|_________|_________|_________|_________|
NOT AT ALL INTERFERED PREVENTED
SOMEWHAT TASK PERFORMANCE

23. How much did the control devices interfere with the performance of assigned tasks or with other activities?

|________|________|__________|_________|_________|_________|_________|
NOT AT ALL INTERFERED INTERFERED
SOMEWHAT GREATLY

24. How well could you concentrate on the assigned tasks or required activities rather than on the mechanisms used to perform those tasks or activities?

|________|________|__________|_________|_________|_________|_________|
NOT AT ALL SOMEWHAT COMPLETELY

25. How completely were your senses engaged in this experience?

|________|________|__________|_________|_________|_________|_________|
NOT AT ALL SOMEWHAT COMPLETELY

26. To what extent did events occurring outside the virtual environment distract from your experience in the virtual environment?

|________|________|__________|_________|_________|_________|_________|
NOT AT ALL SOMEWHAT COMPLETELY
DISTRACTED DISTRACTED

27. Overall, how much did you focus on using the display and control devices instead of the virtual experience and experimental tasks?

|________|________|__________|_________|_________|_________|_________|
NOT AT ALL SOMEWHAT COMPLETELY

28. Were you involved in the experimental task to the extent that you lost track of time?

|________|________|__________|_________|_________|_________|_________|

NOT INVOLVED	MILDLY INVOLVED	COMPLETELY ENGROSSED

29. How easy was it to identify objects through physical interaction, like touching an object, walking over a surface, or bumping into a wall or object?

|________|________|__________|_________|_________|_________|_________|

NOT AT ALL	SOMEWHAT	COMPLETELY

30. Were there moments during the virtual environment experience when you felt completely focused on the task or environment?

|________|________|__________|_________|_________|_________|_________|

NOT AT ALL	SOMEWHAT	COMPLETELY

31. How easily did you adjust to the control devices used to interact with the virtual environment?

|________|________|__________|_________|_________|_________|_________|

NOT AT ALL	SOMEWHAT EASILY	COMPLETELY

32. Was the information provided through different senses in the virtual environment (e.g., vision, hearing, touch) consistent?

|________|________|__________|_________|_________|_________|_________|

NOT AT ALL	SOMEWHAT CONSISTENT	COMPLETELY CONSISTENT

Appendix F

NASA TXL workload rating scale

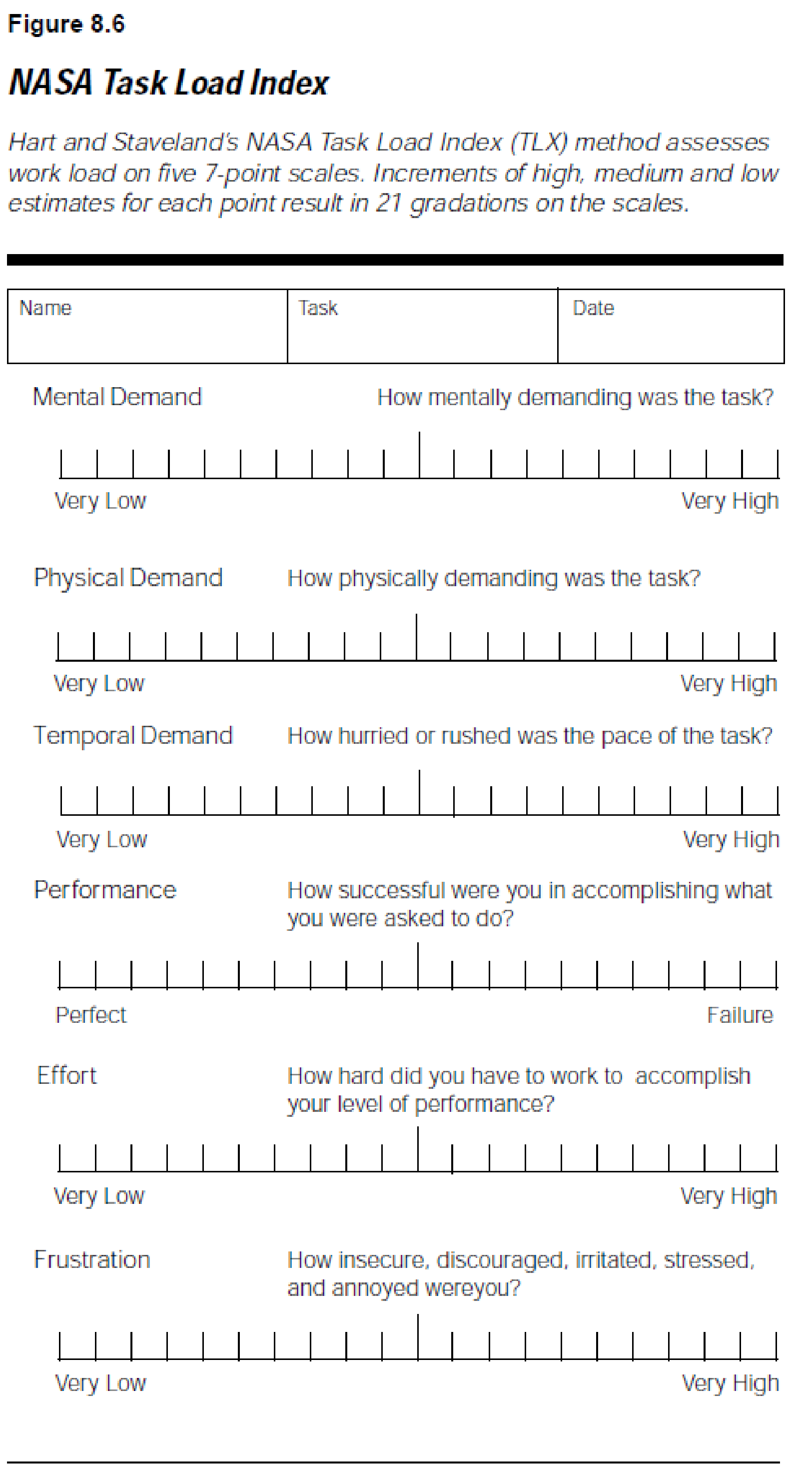


Figure 8.6

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date

Mental Demand How mentally demanding was the task?

Very Low Very High

Physical Demand How physically demanding was the task?

Very Low Very High

Temporal Demand How hurried or rushed was the pace of the task?

Very Low Very High

Performance How successful were you in accomplishing what you were asked to do?

Perfect Failure

Effort How hard did you have to work to accomplish your level of performance?

Very Low Very High

Frustration How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low Very High

Appendix G

Dexterity manual task

Please, **sort** and **arrange** the bolts and blocks from this….

…to this **as fast as you can**!!!