Est. 1841 | YORK ST JOHN UNIVERSITY

Glenn, Callum Peter (2025) Cross-Modal Temporal Integration and Its Impact on Presence and Simulator Sickness in Virtual Reality. Doctoral thesis, York St John University.

Downloaded from: https://ray.yorksj.ac.uk/id/eprint/12549/

Research at York St John (RaY) is an institutional repository. It supports the principles of open access by making the research outputs of the University available in digital form. Copyright of the items stored in RaY reside with the authors and/or other copyright owners. Users may access full text items free of charge, and may download a copy for private study or non-commercial research. For further reuse terms, see licence terms governing individual outputs. Institutional Repository Policy Statement

RaY

Research at the University of York St John

For more information please contact RaY at ray@yorksj.ac.uk

Cross-Modal Temporal Integration and Its Impact on Presence and Simulator Sickness in Virtual Reality

Callum Peter Glenn

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

York St John University

School of Education, Language and Psychology

February 2025

Intellectual Property and Publication Statements

The candidate confirms that the work submitted is their own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Chapter 5: Study 3 is based on work from jointly authored publications: **C. Glenn** and M. Coxon. (2024). Individual differences in processing multisensory information predicts presence in different virtual reality environments. *Virtual Reality*. All research was conceptualised, conducted, analysed and written by the candidate. The other author provided revisions to the written manuscript.

This copy has been supplied on the understanding that it is copyright material. Any reuse must comply with the Copyright, Designs and Patents Act 1988 and any license under which this copy is released.

© 2025 York St John University and Callum Glenn

The right of Callum Glenn to be identified as Author of this work has been asserted by them in accordance with the Copyright, Designs and Patents Act 1988.

Acknowledgements

In the summer of 2018, I took a leap of faith, leaving my life trajectory in retail management for a career in academia. Over the last seven years, countless moments have reaffirmed that this was the right decision. Yet, completing a PhD part-time, alongside the many other commitments in my life, has been both incredibly rewarding and profoundly challenging. This thesis would not have been possible without the support, guidance, and encouragement of many people to whom I owe a great debt of gratitude.

First and foremost, I would like to express my heartfelt thanks to my supervisor, Dr. Matthew Coxon, for their invaluable guidance, insightful feedback, and unwavering support throughout this project. Without their encouragement, I would never have embarked on this journey, let alone reached its conclusion. Their expertise has been instrumental in shaping the direction of this research and my academic career, offering opportunities that have arisen only because of our close relationship and shared outlook. Thank you.

I am also deeply grateful to York St John University for providing the resources, facilities, and funding necessary for this research. Special thanks go to Charlotte Elwell for their technical assistance and administrative support. I am sure they will be glad to no longer see me knocking at their door with a roulette of questions. Charlotte, you have often been my sounding board, helping me make sense of the most complex analyses and coding challenges. Thank you.

I have been fortunate to work alongside a talented group of colleagues in the Psychology department, whose discussions and camaraderie have enriched this journey. I am especially grateful to Dr. Nicola Savill and Dr. Scott Cole, whose guidance and insights have directly influenced the direction and scope of various elements of this thesis. Thank you.

To my friends and family, whose unwavering belief in me has been my greatest source of strength, thank you for your patience, love, and support. A special thank you to Ross Scott, Jake Hourd, and Jack Lowe. Ross, your guidance on readability from a non-scientist perspective has been invaluable, both within and beyond this thesis. Jake and Jack, thank you for uprooting your lives in Leicestershire to rent with me in York when commuting became too difficult a burden.

Lastly, but far from least, my deepest gratitude goes to my fiancée, Emily Hall. When I began this journey, you told me that few have the opportunity to chase their dreams, and fewer still are brave enough to seize them in the face of uncertainty. The truth is, you gave me the courage to be brave and the unwavering support to pursue this dream at every step. From always believing in me to relocating halfway across the country to be by my side, through endless hours of reading and guidance, you have been my anchor. And despite all of this, you've still agreed to be by my side for the rest of our lives. Your support has meant more than words can express.

To everyone who has played a role in this journey, whether mentioned here or not, I extend my deepest gratitude. This achievement is as much yours as it is mine.

Abstract

Presence is a phenomenon characterised by the sensation of being situated within an environment which provides sensory information to an individual. In virtual reality (VR), this feeling of "being there" can vary from person to person. One possible explanation for these variations is how individuals process the technology's multisensory information. Specifically, people differ in their ability to integrate or separate asynchronous and incongruent multisensory cues. This thesis investigates whether variations in the size of an individual's temporal binding window (TBW) affect their user experience in virtual reality environments (VRE) using head-mounted displays (HMDs). Study 1 found that wider TBWs correlated with lower self-reported presence in a joystick-controlled VR experience. Study 2 extended this by examining the link between TBWs and simulator sickness, revealing that wider TBWs were associated with less presence due to conflicting cues. Study 3 explored the effects of removing conflicting cues in a minimally mobile, room-scale VR environment, finding that without conflicts, wider TBWs increased feelings of presence, while in a unisensory setting, the opposite occurred. Finally, Study 4 utilised a within-subject design to verify whether conflicting cues diminished the sense of presence in individuals with wider TBWs. Participants experienced two variations of the same VR simulation: one using natural locomotion and another using joystick control. A four-way multivariate analysis confirmed a significant difference between the two conditions, showing that wider TBWs contributed to increased simulator sickness and reduced presence in the environment with conflicting cues. The findings from these studies are synthesised into a set of principles encompassing three proposed sensory conditions, how these conditions are processed by different-sized TBWs, and the resulting impact on the user experience in VR.

Contents

Intellectual Property and Publication Statements i

Acknowledgements ii

Abstract iv

List of Tables xii

List of Figures xiii

List of Abbreviations xv

Chapter 1: Introduction 1

Chapter 2: Literature Review 5

2.1 Presence in Virtual Environments 5

2.1.1 Models of Presence 6

2.1.2 Cognitive aspects: Attention, Perception and Sensory cues 7

2.1.3 Mental Models 10

2.1.4 Hypothesis Testing 14

2.1.5 Place Illusion vs Plausibility 17

2.2 The Temporal Binding Window 21

2.2.1 Theoretical Models of the TBW 25

2.2.2 Theoretical Underpinnings of the Reaction Time Paradigm 27

2.2.3 Theoretical Underpinnings of the Estimation and Judgement Paradigm 31

2.2.4 The TWIN Model 36

2.3 Presence and The Temporal Binding Window 40

Chapter 3: Exploratory Analysis of Correlation between TBW and Presence (Study 1) 43

3.1 Introduction 43

3.1.1 Measures of Presence 48

3.1.2 Measures of the Temporal Binding Window (TBW) 56

3.1.3 Current Study 62

3.2 Methods 63

3.2.1 Participants 63
3.2.2 Materials 64
3.2.2.1 Temporal Binding Window Tasks. 64
3.2.2.2 VR Hardware and Simulation. 65
3.2.2.3 Questionnaires. 67
3.2.3 Procedure 68
3.2.4 Ethics 69
3.3 Calculating the Window Widths 70
3.3.1 Computing the TBW from the Simultaneity Judgement Task 71
3.3.2 Computing the TBW from the RT Task 72
3.3.3 Computing the TBW from the TOJ task 75
3.4 Results 77
3.4.1 Descriptive Statistics for the Questionnaires 77
3.4.2 Checking the Validity of the Scales 79
3.4.3 Correlations Between Presence Scores, the TBW Measures, and Control Variables 82
3.4.4 Conducting the Regression 82
3.4.5 Results Summary 83
3.5 Discussion 83
3.5.1 The Relationship between TBW and Presence 84
3.5.2 SPES as a Valid Measure of Presence 89
3.5.3 Influence of Confounding Variables on Presence 90
3.5.4 Limitations 91
3.5.5 Summary and Conclusions 93
Chapter 4: Exploring the Relationship Between TBW, Presence and Simulator Sickness (Study 2) 95
4.1 Introduction 95
4.1.1 Simulator Sickness 96

4.1.2 Simulator Sickness and Presence 104
4.1.3 Simulator Sickness and Temporal Binding Windows (TBW) 105
4.2 Study 2a: Pilot of Online Study 107
4.2.1 Rationale and Adjustments for an Online Study 107
4.2.2 Aims and Hypotheses 109
4.2.3 Methods 111
4.2.3.1 Participants 111
4.2.3.2 Materials 111
4.2.3.3 Procedure 115
4.2.3.4 Ethics 116
4.2.4 Analysis 117
4.2.4.1 Participant Data Checks 117
4.2.4.2 Calculating the TBWs 117
4.2.4.3 Descriptive Statistics for the Questionnaires 118
4.2.4.4 Construct Validity of the Presence Scales 120
4.2.4.5 Comparing Online Pilot Results to In-person 120
4.2.4.6 Assessing the Relationship between Presence, Simulator Sickness and TBW 121
4.2.4.7 Results Summary 122
4.2.5 Discussion and Amendments 123
4.3 Study 2b: Online Investigation of the Relationship between Temporal Binding Windows, Simulator Sickness and Presence 126
4.3.1 Participants 126
4.3.2 Materials 126
4.3.2.1 Pre-screening Survey 126
4.3.2.2 Simulation and Hardware 127
4.4 Analysis 128
4.4.1 Participants Data Checks 128

4.4.2 Calculating the TBWs 128
4.4.3 Descriptive Statistics for the Questionnaires 129
4.4.4 Handling Missing Data 129
4.4.5 Correlations Between Predictors, Outcomes and Control Measures 130
4.4.6 Regression Analysis 131
4.4.7 Study 2b Discussion 132
4.5 Discussion 133
4.5.1 TOJ TBW, Presence, and SS 135
4.5.2 Limitations and Issues with Online Testing 137
4.5.3 Summary and Conclusions 138
Chapter 5: Moderating Temporal Binding within Uni v Multisensory Environments (Study 3) 141
5.1 Introduction 141
5.1.1 Presence 144
5.1.2 Simulator Sickness 145
5.1.3 Current Study: Aims and Hypothesis 146
5.2 Methods 148
5.2.1 Participants 148
5.2.2 Design 149
5.2.3 Materials 149
5.2.3.1 Temporal Binding Tasks 149
5.2.3.2 Questionnaires 150
5.2.3.3 VR Hardware and Simulation 151
5.2.4 Procedure 153
5.2.5 Ethics 154
5.3 Analysis 155
5.3.1 Calculating the TBW's 155
5.3.2 Descriptive Statistics 156

5.3.3 Between-group Homogeneity 157
5.3.4 Control Variables and Simulator Sickness 157
5.3.5 Moderated Multiple Regressions 159
5.3.5.1 Spatial Presence Experience Scale 159
5.3.5.2 Slater, Usoh and Steed Presence Questionnaire 162
5.3.5.3 SPES Subscales: Self-Location and Perceived Possible Actions 164
5.4 Discussion 168
5.4.1 TBW and Presence in Multisensory Room-scale VREs 169
5.4.2 Moderating the relationship between TBWs and Presence 172
5.4.3 Towards a Dual Role of the TBW in Presence and Simulator Sickness 173
5.4.4 Limitations of the Current Study 175
5.4.5 Summary and Conclusion 177
Chapter 6: Individual Differences in Temporal Binding Windows Moderate the Impact of Conflicting Vestibular Cues on Presence and Simulator Sickness (Study 4) 180
6.1 Introduction 180
6.1.1 Non-Euclidean VR and Impossible Spaces 182
6.1.2 Impossible Spaces, Presence and Related Concepts 184
6.1.3 Beyond Self-Location and Perceived Possible Actions 186
6.1.4 Rationale and Aims of the Current Study 188
6.2 Methods 190
6.2.1 Participants 190
6.2.2 Design 190
6.2.3 Materials 191
6.2.3.1 VR Hardware and Simulation 191
6.2.3.2 Questionnaires 192
6.2.4 Procedure 194
6.2.5 Ethics 195
6.3 Analysis 196

6.3.1 Calculating the Window Widths 196
6.3.2 Descriptive Statistics for the Main Measures 197
6.3.3 One-way Within Subject Multivariate Analyses 198
6.3.4 Post-Hoc Analyses 202
6.4 Discussion 204
6.4.1 Multisensory Facilitation is More Important than Temporal Judgement in the Formation of Presence 206
6.4.2 Individuals' Ability to Handle Sensory Conflict Influences Presence and Simulator Sickness 208
6.4.3 TBWs Influence Self-Location and Higher Cognitive Involvement 210
6.4.4 The Dual Role of TBW in VR User Experience 212
6.4.5 Limitations and Future Research Directions 214
6.4.6 Summary and Concluding Remarks 217
Chapter 7: General Discussion and Summary of Conclusions 218
7.1 Overview of Findings 218
7.1.1 Chapter 3: 220
7.1.2 Chapter 4: 220
7.1.3 Chapter 5: 221
7.1.4 Chapter 6: 222
7.2 Rationale for Establishing New Principles 224
7.3 The Finalised Set of Principals 225
7.4 Understanding Findings through the Sensory Principles 227
7.4.1 The Principles Applied to Studies 1-4 227
7.4.2 The Principles in the Context of Recent Literature 230
7.5 Current Limitations of the Proposed Principles and Future Directions 233
7.6 Conclusion 236
References 239
Appendices 284

Appendix 1: Gaming Experience Questionnaire Adapted from Lemmens et al., (2011) .284

Appendix 2: Important Health and Safety Information.....285

Appendix 3: Adapted MATLAB Scripts from Notter, Murray & Micah (2017).....286

Appendix 4: MATLAB Scripts from Alcalá-Quintana and García-Pérez Independent Channels Model 298

Appendix 5: MATLAB Scripts for Mapping Data to the Extended Twin Model.....307

Appendix 6: Moderation of Presence Subscales by RT TBW Across Conditions in Study 4 315

List of Tables

Table 1 *Boundaries for fmincon Optimisation of the Twin Model*.....74

Table 2 *Descriptive Statistics for the TBW Tasks in Study 1* 77

Table 3 *Descriptive Statistics for Experience and Presence Questionnaires in Study 1* 78

Table 4 *Inter-Item Correlation Between SPES and IPQ Subscales in Study 1*.....80

Table 5 *Correlation between Confounding Variables, Presence Questionnaires and TBWs in Study1*.....82

Table 6 *Descriptive Statistics for the TBW Tasks in Study 2a* 118

Table 7 *Descriptive Statistics for Questionnaires in Study 2a*.....119

Table 8 *Correlation Between Simulator Sickness, TBW and Presence Measures in Study 2a* 122

Table 9 *Descriptive Statistics for the TBW Tasks in Study 2b* 129

Table 10 *Descriptive Statistics for the Questionnaires in Study 2b*.....129

Table 11 *Correlations Between Predictors, Outcomes and Controls for Study 2b (N=54)* .131

Table 12 *Descriptive Statistics for Presence, SS and TBW by Sound Condition in Study 3*.156

Table 13 *Correlations Between Control Variables, TBW and Presence Measures in Study 3* 158

Table 14 *Correlations Between Simulator Sickness, TBW and Presence Measures in Study 3* 158

Table 15 *Descriptive Statistics for the TBW tasks in Study 4* 197

Table 16 *Descriptive Statistics for Presence and SS by Navigation Condition in Study 4*...197

Table 17 *Summary of Research Questions, Answers and Future Directions from the Present Thesis* 219

List of Figures

Figure 1 *Visual representation of a temporal binding window task trial sequence* 64

Figure 2 *Oculus Rift headset, handheld controllers, and tracking cameras used in study 1*..66

Figure 3 *Screenshot of a virtual reality environment generated using seed code 'LISTRISK' in Minecraft VR version 1.7* 67

Figure 4 *Example of a participant's simultaneity judgement temporal binding window at the 50% and 70% cut-offs*.....72

Figure 5 *Example of a participant's reaction time data fitted to the twin model* 75

Figure 6 *Example of a participant's temporal order judgement temporal binding window* .76

Figure 7 *Screenshot of a single trial from an online temporal order judgment (TOJ) task* 112

Figure 8 *Oculus Quest Headset and handheld controllers used in Study 3*.....151

Figure 9 *Image taken from 'Office Worker' in Job Simulator: The 2050 Archives illustrating the office cubicle VRE*.....153

Figure 10 *Interaction effect of the presence of sound on the relationship between individuals' temporal binding window and self-reported presence score on the spatial presence experience scale* 160

Figure 11 *Interaction effect of the presence of sound on the relationship between individuals' temporal binding window and self-reported presence score on the spatial presence experience scale* 161

Figure 12 *Interaction effect of the presence of sound on the relationship between individuals' temporal binding window and self-reported presence score on the Slater, Usoh, and Steed presence questionnaire* 163

Figure 13 *Interaction effect of the presence of sound on the relationship between individuals' temporal binding window and self-reported score on the self-location subscale.* 165

Figure 14 *Interaction effect of the presence of sound on the relationship between individuals' temporal binding window and self-reported score on the perceived possible action subscale* 166

Figure 15 *A screenshot of an interactive lever to access "loot" within the video game Tea for god, developed by Void Room*.....192

Figure 16 *Interaction effect of the VR condition on the relationship between individuals' redundant target temporal binding window and self-reported score on the MEC-SPQ* 199

Figure 17 *Interaction effect of the VR condition on the relationship between individuals' redundant target temporal binding window and change in score on the SSQ* 199

Figure 18 *Interaction effect of the VR condition on the relationship between individuals' temporal order judgement temporal binding window and self-reported score on the MEC-SPQ* 200

Figure 19 *Interaction effect of the VR condition on the relationship between individuals' temporal order judgement temporal binding window and change in score on the SSQ* 201

Figure 20 *A Model For The Dual Role Of Temporal Binding Windows In The Formation Of Presence And Simulator Sickness* 213

List of Abbreviations

2AFC	Two Alternative Forced Choice
2IFC	Two Interval Forced Choice
BIP	Break in Presence
BPS	British Psychological Society
CI	Confidence Interval
CLCC	Capacity Limited Cognitive Constructionist
CRT	Cathode Ray Tubes
CV1	Commercial Version 1
HMD	Head-Mounted Display
IPQ	Igroup Presence Questionnaire
LCD	Liquid Crystal Display
MAP	Maximum A-Posteriori Estimate
MAR	Missing at Random
MCAR	Missing Completely at Random
MCD	Multisensory Correlation Detector
MEC-SPQ	Measurements, Effects, Conditions, Spatial Presence Questionnaire
MLE	Maximum Likelihood Estimate
MNAR	Missing Not at Random
MSQ	Motion Sickness Questionnaire
OLED	Organic Light Emitting Diode
PERF	Primary Egocentric Reference Frame
PF	Perceptual Fusion
PPA	Perceived Possible Actions

PQ	Presence Questionnaire
RDW	Redirected Walking
RT	Redundant Target
SD	Standard Deviation
SEM	Structural Equational Modelling
SJ	Simultaneity Judgement
SL	Self-Location
SOA	Stimulus Onset Asynchrony
SPES	Spatial Presence Experience Scale
SSQ	Simulator Sickness Questionnaire
SUS	Slater, Usoh and Steed (Presence Questionnaire)
TBW	Temporal Binding Window
TOJ	Temporal Order Judgement
TWIN	Time Window of Integration
VIMS	Visually Induced Motion Sickness
VR	Virtual Reality
VRE	Virtual Reality Environment

Chapter 1: Introduction

Some malicious demon of the utmost power and cunning has employed all his energies in order to deceive me. I shall think that the sky, the air, the earth, colours, shapes, sounds and all external things are merely the delusions of dreams which he has devised to ensnare my judgement... I am like a prisoner who is enjoying an imaginary freedom while asleep; as he begins to suspect that he is asleep, he dreads being woken up, and goes along with the pleasant illusion as long as he can. (Descartes, 1641/2017, First Meditation)

Descartes's reflection on the reliability of sensory perception and the nature of reality parallels the ambitions of virtual reality (VR). In VR, researchers, developers, and engineers play the role of the '*demon*'. However, instead of '*malicious'* deception, the aim is to transport individuals into immersive alternate worlds. The '*energies in order to deceive*' are the technologies we use to achieve such goals. These include head-mounted displays (HMDs), haptic feedback gloves, and trackpads. When we employ these technologies, they have the power to deceive the user's senses. This ability to '*ensnare ... judgement*' underpins the psychological concepts of immersion and presence. Whilst immersion refers to the technological ability to provide 'external stimuli', presence is the outcome of the deception of the senses, which leads the user to feel like they are in another place "*like a prisoner who is enjoying an imaginary freedom*".

The current understanding of presence originates from Minsky's concept of 'telepresence' (1980), which refers to the ability to operate machinery remotely while feeling as if one is physically present at the operational site. While VR is related to telepresence, it goes beyond merely enabling remote operations (Lanier, 2017). VR offers users a rich and immersive experience, allowing the user to engage with virtual worlds, from training

simulations to entertainment. This immersive quality enables individuals to feel like they are located within the environments being displayed (Cummings & Bailenson, 2016).

The concept of being located in a different place, or "presence by transportation", as defined by Lombard and Ditton (1997), is fundamental to the idea of presence in VR. This idea is more simply expressed by Heeter (1992) as "being there," and it has more recently been referred to as the "place illusion" (Slater, 2009; Slater et al., 2022). This concept has primarily served as an indicator of the effectiveness or success of immersive technology on a subjective level. Exploring the extent to which individuals feel they are truly within the virtual world rather than being aware of their actual presence in the real world.

Theories have been proposed to explain how this sense of 'being there' develops within the human mind (e.g. Biocca, 2003; Nunez, 2007; Schubert et al., 1999b; Slater & Steed, 2000; Slater et al., 2022; Wirth et al., 2007). These theories agree on the broad principles of how presence is formed, typically beginning with attentional processes and ending in mental models of the virtual world and the acceptance of being located within them (e.g. Nunez, 2007; Wirth et al., 2007). One key element that underlines all of these stages is the processing of sensory information, specifically the congruency of multisensory information within VR. Chapter 2 explores the concept of presence through a literature review that outlines some of the most relevant theories, highlighting commonalities and the significance of sensory processes.

When multisensory information related to the same event is produced congruently, it is bound into a single percept (Stein & Meredith, 1993). This process is known as multisensory integration (Calvert & Thesen, 2004). One key component of multisensory integration is related to the time discrepancy between the production and processing of the information. This component is known as temporal binding and has been demonstrated to differ between individuals (e.g. Wallace & Stevenson, 2014). Although it has rarely been linked to presence

research before, temporal binding overlaps with individual differences in attentional processes (e.g. Talsma et al., 2007), mental models (e.g. Nardini et al., 2018) and the perception of congruency (e.g. Wallace & Stevenson, 2014).

In the second section of Chapter 2, theoretical models are, therefore, presented on how individuals integrate temporally congruent multisensory information. This section also explores the potential connection between the concepts of presence and temporal binding. While temporal binding has been extensively studied in contexts such as audio-visual integration (Colonius & Diederich, 2018), its influence on the formation of presence in VR has yet to be investigated.

This thesis aims to bridge this gap by examining how individual differences in temporal binding affect the experience of presence. By exploring this relationship, the research seeks to provide new insights into the cognitive and sensory foundations of VR experiences.

Chapters three to six outline four studies used to examine this connection and offer a new perspective on the impact of individual differences in forming a sense of presence. Chapter 3 presents the first exploratory study, investigating the relationship between temporal binding and presence. In Chapter 4, this relationship is further explored in an online study that also examines the potential confounding variable of simulator sickness. Chapter 5 examines how multisensory stimuli may moderate the link between temporal binding and presence while controlling for visual-vestibular conflict. Chapter 6 presents a final study that isolates the impact of visual-vestibular conflict on temporal binding, simulator sickness, and the development of presence, synthesising and confirming insights from the previous experiments.

Finally, the general discussion in Chapter 7 presents a set of principles that encapsulate the role of temporal binding in the development of presence. These principles are based on empirical evidence from the thesis as well as broader theoretical frameworks. The implications

of these principles are examined in relation to the studies included in this thesis and within a wider context, providing recommendations for researchers and developers to improve user experience in virtual reality.

Together, the chapters of this thesis demonstrate the complex relationship between temporal binding windows, presence, and simulator sickness in VR. The findings highlight how individual differences in TBW width shape user experiences, influencing how sensory cues are integrated within a virtual environment. Specifically, the findings advocate for the design of vestibular-congruent multisensory virtual environments that account for variations in TBWs to optimise presence while minimising simulator sickness. By establishing a set of principles, this research offers both theoretical insights and practical recommendations for improving immersive VR experiences. Through this research, the thesis aims to advance the understanding of how individual differences in cross-modal temporal binding impact VR user experience, laying the groundwork for future research and practical application in the field.

Chapter 2: Literature Review

2.1 Presence in Virtual Environments

When individuals experience books, films, or immersive technology, they receive sensory information about the medium's environment. One outcome of individuals receiving this sensory information is their potential to feel like they have been transported to the presented environment. This feeling of 'being there' within that environment is known as presence (Heeter, 1992). The concept of presence dates back long before modern immersive technology existed (Biocca et al., 1995; Biocca & Levy, 1995). However, the current definition stems from Minsky's (1980) description of ‘telepresence’. This definition is in relation to individuals feeling like they are “*actually there*" within a virtual environment. The concept of telepresence was expanded in the mid-nineties. Draper separated the cybernetic side of presence (closely linked to immersion in the current literature) and the experiential side (Draper, 1995). Whilst the cybernetic side relates to the technology used to invoke presence, experiential telepresence relates to the mental state of presence. The sense of ‘being there’. A concept Lombard and Ditton (1997) considered to be 'presence as transportation' but has more recently become known as 'spatial presence', which is the focus of this thesis. Therefore, going forward, the term presence will be discussed interchangeably with this notion of spatial presence.

The degree of presence in a virtual environment (not the 'real world') is known to vary. Not only does it vary between different environments, but also between individuals in the same environment. For example, two people may enter the same virtual reality game using the same equipment. However, one person may feel as though they have been transported to the virtual environment, whereas the other may still feel located in the 'real world'. These individual differences in presence are largely accepted to occur throughout the academic literature (e.g.,

Coxon et al., 2016; Jurnet et al., 2005; Iachini et al., 2019). However, models of presence, albeit not restricting individual differences, rarely focus on the individual nature of presence. Instead, they take a more overarching position on the concept. Therefore, this section will discuss relevant models of presence, turning a lens onto shared components of each model that may contribute to the various levels of presence individuals may experience.

2.1.1 *Models of Presence*

The most simplistic model of presence is underpinned by one fundamental idea. This idea is that users either feel present in the ‘real world’ or a virtual world. In addition to feeling present in one world or the other, users can switch between the two. Slater and Steed (2000) use the term ‘environmental selection’ to describe this basis of presence, whereas Biocca (2003) captures the phenomena with the term 'two-pole' model of presence. The two-pole model outlines that internal and external factors impact a user moving from 'not present' to 'present' within an environment. External factors relate to elements of the VRE itself. The environment should be sensorially rich, consistent, interactive, and have a self-representation and causality (Slater et al., 1994; Slater & Usoh, 1993). Internal factors are those which are subjective to the users. They explain how different people can respond differently to the same environment and are thought to be mediated through people's differing abilities to process perceptions and form mental models (Slater & Steed, 2000). Both these abilities involve the processing of multisensory information. The focus of the next section will be outlining some of these internal factors which models of presence frequently contain.

Models of presence have mainly been developed in three ways. Firstly, models have been developed to contrast other models, such as Slater and Steed's (2000) environmental selection model versus Biocca's (1997) two-pole model. Secondly, models can be developed to improve a current model. For example, Biocca (2003) expanded the previous model to

include an additional pole, expanding their model to include more imaginary stimuli such as dreams and books. Finally, models can be developed to fill gaps in the literature or introduce a new perspective. There are many examples of this, such as the embodied presence model (Schubert et al., 1999b), which included a framework for embodied cognition in presence (Glenberg, 1997). Another example, which overlaps most with the other models and hence forms the basis for the subsections of the chapter, is Wirth and colleagues' Process Model (Wirth et al., 2007). The process model looks at the steps involved from an individual receiving sensory information about an immersive technology's environment to the user feeling spatially present in that environment. It involves multiple stages, but the focus will be on those it shares with other models, such as the cognitive aspects of attention and perception, sensory cues, mental models, and hypothesis testing.

2.1.2 *Cognitive aspects: Attention, Perception and Sensory cues*

Almost all models of presence agree that for presence to form in a virtual environment, the individual must pay attention to the information the environment provides. However, Slater and Steed (2000) highlight that individuals in a virtual environment still receive sensory information from the 'real world'. This information, such as room temperature and gravity, can sometimes clash with the sensory information of the virtual environment. Therefore, virtual reality is more successful in generating a sense of presence if it attracts more attention than the 'real world' (Schuemie et al., 2001; Slater & Steed, 2000; Wirth et al., 2007). Biocca's (2003) three-pole model introduced the idea of spatial attention to capture this shift in presence. Spatial attention can be both 'top-down' and automatic. Top-down refers to the individual's goals and actively paying attention to the environment's sensory information. At the same time, automatic spatial attention is regarding sudden environmental changes and those aspects of the environment which draw in an individual's attention. For example, a virtual character moving in the corner might draw some attention, yet if the character were also to make a sound, it

would likely draw more attention. This is due to multisensory cues increasing the speed at which attention is focused (Koelewijn et al., 2010; Talsma & Woldorff, 2005) compared to single modality cues. The increase in attention is attributed to the integration of multisensory stimuli, reaching a response criterion faster (Talsma & Woldorff, 2005). This is supported by literature that indicates virtual environments which use more modalities, lead to an increased sense of presence (e.g. Dinh et al., 1999; Slater et al., 1995). Models, therefore, have been proposed to account for the different attentional processes involved in developing presence.

One such model of presence was put forward by Nunez (2007). The capacity-limited cognitive constructionist model (CLCC) captured attentional processes involved in presence regarding the constraints of human information processing. In the CLCC, information about the virtual environment is processed by an individual's stimuli attenuator. Compared to the top-down/automatic process proposed by Biocca (2003), the attenuator relates more to maintaining an entire perspective of the environment. Therefore, it captures both top-down and automatic attentional processes and attends to stimuli that are highly congruent with the individual's perspective of what the environment should be. In Nunez's (2007) model, the attenuator is modulated by higher-order processes, such as the working memory module. As working memory is known to have a limited capacity, the CLCC provides one explanation as to why multiple modalities can increase an individual's sense of presence (Dinh et al., 1999; Hecht et al., 2005). Sensory information which is meaningfully related becomes 'chunked' together. Information from two modalities related to the same percept are bound into one 'chunk', allowing more information to be attended to and subsequently processed. In addition, information that cannot be chunked together, such as information from competing environments, requires more processing and subsequently reduces an individual's feeling of presence (Nunez, 2007).

Like the CLCC, the process model (Wirth et al., 2007) highlights the importance of attention to sensory information. However, the process model differs from the CLCC in one primary aspect. Rather than one 'attenuator' maintaining the entire perspective of the environment, the process model splits attention back into controlled and involuntary attention allocation. This is an important difference between the cognitive-focused models. In splitting the attention allocation between automatic and controlled processes, Wirth and colleagues' model (2007) allows more careful consideration of which factors impact attention. The key point is that this dichotomy of attention allocation is similar to, yet more developed than, Biocca's (2003) top-down and automatic spatial attention. For example, user-directed top-down attention is somewhat the same as controlled attention allocation. In the sense that controlled attention allocation is partially directed by user goals, among other user-directed factors. In addition, a comparison between automatic spatial attention and automatic attention allocation can be made. They both focus on the role of the environment, drawing attention. However, the process model develops on this idea, going beyond capturing attention to also maintaining attention.

This maintenance of attention somewhat overlaps with the concept of the working memory's role in presence from the CLCC model (Nunez, 2007), in the sense that the process model highlights the importance of maintaining both a 'breadth' and 'depth' of sensory vividness (Steuer, 1992). These two concepts relate to the number of sensory modalities the environment provides and the amount of information each modality provides. Similar to the idea of chunking in the CLCC (Nunez, 2007), the process model highlights the importance of these sensory cues in sustaining involuntary attention allocation (Wirth et al., 2007). However, the process model also warns of overly complex stimuli leading to fatigue (similar to multisensory inhibition), which may diminish an individual’s feeling of presence.

In both the CLCC and process model (Nunez, 2007; Wirth et al., 2007), attention to sensory cues is the first step toward developing a sense of presence. These sensory cues must be coherent with other sensory cues provided; for example, a visual display of someone talking must match the audio of someone talking. This provides an experience which is neither offset nor incongruent. If the sensory cues are incongruent, the overall sense of presence diminishes (Wirth et al., 2007; Witmer & Singer, 1998), regardless of whether the incoherent cues are provided by the virtual environment itself or between environments.

Although the consistency of sensory cues is important for both the formation and maintenance of the feeling of presence, different sensory modalities have varying impacts on this experience (Dinh et al., 1999). Visual and auditory information generally influence perception most (Hutmacher, 2019). Meanwhile, haptic and olfactory sensory cues impact presence to varying extents (e.g. Dinh et al., 1999; Gibbs et al., 2022). Not only does the influence of the sensory cues differ between modalities, but individuals also display differing abilities to maintain a unified percept when the modalities are offset (e.g. Wallace & Stevenson, 2014). Although there has been little to no investigation of this effect regarding virtual reality (Glenn & Coxon, 2024), it seems logical to assume that these individual differences in the ability to integrate offset stimuli may (in part) attribute to individuals differing feelings of presence. Furthermore, this hypothesis may be true in maintaining attention to the sensory cues and the subsequent stage of feeling present in virtual reality: building a mental model.

2.1.3 *Mental Models*

In contrast to the previous section, there is no consensus on mental models in the current literature. This lack of agreement arises mainly because mental models underpin an author's model of presence, influencing its formation, maintenance, and implications. Mental models are defined as mental representations of the depicted space (Glenberg, 1997). Individuals can

simultaneously form mental models of the 'real world' and the virtual environment (Biocca, 2003; Slater & Usoh, 1993). However, in most models of presence, individuals can only accept one mental model as being the one in which they are located. Accepting only one mental model links back to Slater and Steed's (2000) suggestion that individuals can only respond and act within one environment at a given time. However, Biocca's (2003) tri-pole model proposed that all 'environments' can be integrated to form a coherent presence experience. This integration can only happen if the information from all three poles (real, virtual, and imaginary) is coherent enough that the mental model encompasses each pole in a singular experience. Biocca (2003) refers to this as spatial updating, which, although it can encompass all three of Biocca's poles, largely shifts the feeling of presence between each pole as the mental model is changed through sensory information and unconscious processes. Although Biocca (2003) disagrees with Slater and Steed (2000) on the dynamic nature of presence, both have the concept of a mental model at the centre of their explanation of how presence forms. However, the process of how this mental model forms is contested within the literature.

Schubert and colleagues put forward one explanation for how the mental model of a virtual environment is formed (Schubert et al., 1999b). The embodied presence model was based on Glenberg's (1997) embodied cognition, which proposes that the meaning of an object or situation is created by the meshed patterns of actions possible within that situation. As a result, Schubert and colleagues' model focused on perceived possible actions as a core concept, such as navigation and body movement within the virtual environment. This model argues that individuals construct a representation of the virtual environment through 'meshing' these possible patterns of action with those stored in memories for that situation. This comparative representation of perceived possible actions helps develop and maintain the mental model of the virtual environment. The mental model is formed if the comparison is accurate (the actions

mirror memory). If the perceived possible actions do not mirror memory, then the mental model collapses.

The process model of presence also considers this idea of perceived possible actions (Wirth et al., 2007). Wirth and colleagues (2007) highlight that this 'embodied cognition' is a fundamental requirement of spatial presence. In their process model, perceived possible actions represent one of two key supporting factors for the mental model. The second factor is the sensation of 'being there' within the virtual environment, which Wirth and colleagues refer to as self-location. The model claims that individuals experiencing spatial presence will only perceive actions possible in the environment that they feel located within. Furthermore, whilst feeling located in the mediated environment, individuals will be less aware of possible actions in the 'real world'. This is a key difference between the embodied presence model and the process model, as the prior identifies the perceived possible actions as both a formative and mediating factor of the mental model of the environment (Schubert et al., 1999b). Meanwhile, the process model regards perceived possible actions as only a mediator of the mental model, relying more on sensory cues for development.

The mental model in the process model is regarded as the spatial situational model (Wirth et al., 2007). As mentioned in the previous section, the process model of presence begins with individuals attending to sensory cues about an environment. These sensory cues contribute to the development of the spatial situational model, with both controlled and involuntary attentional pathways feeding information into this evolving model of the mediated environment. The cues trigger a series of spatial cognitions related to the current situation and prior knowledge of spatial scenarios. As the virtual environment continuously produces the sensory cues, the spatial situational model is evaluated for congruency with the sensory cues. Any discrepancies between the mental model and the sensory information leads to a

restructuring of the model. This restructuring may be an adaptation or a rejection in favour of another model (Wirth et al., 2007).

Whilst the development of the mental model in the process model is more advanced than other models discussed, it is not alone in the assumption that multiple cognitive processes are involved in its formation. The CLCC (Nunez, 2007) puts forth that after the sensory information has been attended to, the information is passed to two separate folk modules. These folk modules relate to past information stored in memory. The first is a physics module. This module accesses mental models about the environment from prior similar situations. The other folk module is the psychological module, which infers mental states from the sensory information (Nunez, 2007). Information from the two folk modules passes through several memory modules, which provide episodic knowledge to the stimulus attenuator as a form of feedback. This feedback acts as a confirmatory process to check the temporary structure and semantic basis of the mediated environment. Once the feedback has been confirmed as correct, the information is passed to the motor control centre with procedural memories for possible actions available for the current scenario.

Much like the attentional processes involved in the process model and the CLCC, the mental model serves a similar role in both. One argument may be made in favour of Wirth and colleagues' approach not relying on several black box mechanisms. The 'folk modules' highlighted by Nunez (2007) provide an easy escape from highlighting the actual processes occurring. In contrast, the process model clearly distinguishes between the 'bottom-up' and 'top-down' effects on the mental model. With coherent sensory inputs forming the mental model, and a combination of perceived possible actions and self-location maintaining the model. Despite the criticism of the CLCC, both models share the same fundamental premise: Individuals attend to stimuli. Once these stimuli are constructed to represent the current

scenario, mental models are developed. Through a series of checks with memories, these mental models may be accepted, adjusted, or rejected in favour of another mental model.

The formation of a mental model relies heavily on an individual's ability to process multiple sources of sensory information, which plays a crucial role in establishing spatial presence. When sensory information is inconsistent, it can lead to the development of multiple mental models, representing both the ‘real world’ and the virtual environment. For a mental model of the virtual environment to be stable, the sensory information must be processed in a congruent manner (Wirth et al., 2007). However, a mental model is not the final stage of achieving spatial presence. As mentioned earlier, individuals may accept, adjust, or reject a mental model (Nunez, 2007). When multiple mental models are formed from sensory cues taken from both the real world and the virtual environment, individuals must decide which mental model to accept in order to feel spatially present in a particular environment. This decision-making process can be described through the concept of hypothesis testing.

2.1.4 *Hypothesis Testing*

Once individuals have paid attention to the sensory information of the virtual environment and subsequently developed a mental model of the environment, they need a process to select it. This process is known as hypothesis testing, which essentially asks the question 'Does this mental model accurately represent the environment I am located within?'. However, much like the other aspects discussed so far, the hypothesis testing procedure varies based on the model being used.

The expansion of the environmental model (Slater, 2002) put forward the idea that environments are selected moment-by-moment. This moment-by-moment selection is directly linked to the idea of continuous perceptual hypothesis testing (Wirth et al., 2007). To describe this process, Slater (2002) uses the example of a virtual pit room. A virtual pit room is one

where the individual is asked to look through a hole in the floor of the virtual environment. In the hole, there is another room quite some distance below. This sensory information creates the mental model that the individual is standing at quite some height. At this point, Slater argues that the individual's hypothesis testing can become apparent to an onlooker or researcher. On one side of the hypothesis testing, there is the 'real world' mental model. Slater referred to this as 'abstract knowledge' (Slater, 2002). The individual knows that they are indeed in the 'real world' in an immersive technology, yet their perception indicates otherwise. This perceptual interpretation forms the opposing hypothesis. In this case, the individual either accepts the mental model that they are standing high above the ground or rejects it for the feeling that they are simply in a headset within a laboratory. Either outcome of the hypothesis testing leads to confirmatory actions. For example, if the pit room mental model is accepted over the ‘real world’, the individual displays physiological responses much like they would if suffering from vertigo.

Although Slater's (2002) explanation of hypothesis testing captures the basis of presence developing from a mental model, it somewhat oversimplifies the distinction between mental models. Firstly, the all-or-nothing nature of abstract versus perceptual mental models ignores the potential for multiple mental models to exist within both. Biocca's (2003) three-pole model highlights this distinction, with spatial attention and updating working in unison. Spatial updating serves as the process of hypothesis testing. At the same time, spatial attention shifts individuals between real and imaginary spaces. The combination of both components working in tangent creates a dynamic process of presence, enabling individuals to accept information from both 'abstract' and 'perceptual' mental models. When individuals accept a mental model formed from information belonging to only one environment, their presence is undivided (Biocca, 2003). In contrast, individuals who attend to multiple environments and

adjust the mental model accordingly have a divided presence. Nonetheless, they still maintain a sense of presence through this hypothesis-testing process.

The disagreement between Biocca (2003) and Slater (2002) is perhaps resolved in the process models handling of hypothesis testing (Wirth et al., 2007). Unlike the other two theories, this model introduces an additional stage between the mental representation of an environment and the final experience of feeling present. Wirth and colleagues explain that just because individuals develop a coherent mental model of the environment does not mean that they necessarily feel located within it. The model explains this distinction as a difference between mental representation and experiential state—the bridge between the two lies in primary egocentric reference frames (PERF). Egocentric reference frames are a mental model of the surrounding world from a first-person perspective (Riecke & Von der Heyde, 2003). Individuals sustain this egocentric reference frame through constant spatial updating, which allows them to feel located within it (Wirth et al., 2007).

All sensory information contributes towards constructing this reference frame, including perceptions of one's body. However, as all sensory information generates these egocentric reference frames, it is possible for multiple to exist. Information from different sensory modalities can create multiple egocentric reference frames, with individuals attempting to focus on one to avoid confusion and effectively navigate a single environment. They tend to choose the most congruent environment in which to act (Wirth et al., 2007). This selected environment becomes the PERF. Challenges arise in virtual environments when the sensory information presented to an individual creates a conflict between these reference frames. This conflict occurs due to uncertainty about whether the 'real world' or the virtual environment serves as the primary reference frame.

The perceptual hypothesis testing involved in confirming which environment is the PERF includes three stages: expectations from prior experience, information input, and confirmation or disproof of the hypothesis (Wirth et al., 2007). Stronger expectations in the initial stage require less environmental information to confirm the PERF. Conversely, when expectations are strong, more contradictory information is needed to disprove them. In the context of virtual reality, initial expectations are typically not as strong as those formed about the real world. This leads to the conclusion that a strong mental model, which demands attention through coherent sensory information, is needed for the virtual environment to be considered the PERF (Wirth et al., 2007). When this occurs, the feeling of presence is shifted into the virtual environment. However, if sensory information is inconsistent, it creates a conflict in the mental model. Given that perceptual expectations for virtual environments are weaker than those for the real world, this inconsistency can diminish the sense of spatial presence (Wirth et al., 2007).

This provides a dual explanation for the individual differences observed in virtual reality research. On one hand, individuals possess varying expectations of virtual environments used within perceptual hypothesis testing. For example, some individuals have more experience with VR than others. On the other hand, individuals also differ in their ability to handle inconsistent sensory information, which is influenced by the size of their temporal binding window (see Section 2.2). Slater (2009) captured this dichotomy of perceptual expectations and bottom-up sensory information in a relevant, albeit slightly different, manner. Comparing a perceptually driven place illusion, to a cognitively driven plausibility illusion.

2.1.5 *Place Illusion vs Plausibility*

In 2009, Slater proposed a theory that describes presence as involving two interrelated concepts. The first concept pertains to the sense of 'being there', known as place illusion.

Similar to the models discussed above, this concept refers to the individual feeling as if they have been transported to a virtual environment. However, unlike the other models, Slater proposes that individuals are always aware that they are not in a virtual environment. He defines this Place illusion as "the illusion of being in a place in spite of the sure knowledge that you are not there". Comparing the illusion of the surrounding virtual environment to a nexus cube. The lines on a piece of paper form the perceptual illusion that the 2D object becomes a 3D cube, although the individual is sure that the cube is just a drawing on a piece of paper. In virtual reality, the individual knows they are within immersive technology, yet they accept the illusion regardless. That is unless the illusion is broken.

The place illusion is broken when immersive technology stops providing synchronous correlations between the individual's actions and the updating of the environment. This can be due to the actions reaching the limit of the immersive technology's capabilities. For example, the individual in the virtual environment becomes acutely aware of the pixels making up objects rather than perceiving the objects as a whole. Alternatively, this could be due to a technical glitch. Slater regards this break-in illusion as a BIP (break-in presence) (Slater & Steed, 2000). Importantly, a BIP does not permanently diminish the place illusion. Once the cause of the BIP is rectified, the illusion is resumed. However, this is not the same for the second component Slater (2009) theorised.

The second concept Slater put forth was a plausibility illusion (Slater, 2009; Slater et al., 2022). Whereas the place illusion regards the sense of ‘being there’, the plausibility illusion relates to what is happening there. Slater defines the term as “the illusion that what is apparently happening is really happening (even though you know for sure that it is not)” (2009). Skarbez et al. (2017) use the example of a virtual library. Whereas the place illusion would be the bookshelves and setting the individual could see, the plausibility illusion would be individuals in the library being quiet and acting as the individual would expect within the setting. If the

virtual characters within the library suddenly start running around and shouting, the illusion of plausibility is likely to be shattered. In this situation, the plausibility illusion is diminished, and Slater argues that the illusion is not likely to recover. This is due to the plausibility illusion being more cognitive than perceptual. Once the reality is broken, it will not likely be reformed for that scenario.

In contrast, Skarbez et al. (2017) provided an extension to the plausibility illusion that allows for such discrepancies to Slater's original proposal. Whereas in the place illusion, the limits of the immersive technology provide the boundaries of the illusion, Skarbez (2016) argued there was a counterpart for the plausibility illusion. The introduction of coherence provides this boundary for the plausibility illusion. Coherence refers to the number of reasonable circumstances the virtual environment can demonstrate before introducing unreasonable circumstances. The key to this boundary is that it is set by prior knowledge. For example, if an individual were informed that they would be experiencing a virtual walkthrough of New York City, individuals would expect to see people walking around a city filled with skyscrapers and bustling streets filled with taxis. Individuals would not expect to see the sky filled with spaceships and individuals flying, as this would break the plausibility of the virtual environment. However, if individuals are informed that they will be experiencing a fantasy alien invasion of New York, then they will have updated their expectations of the virtual environment. The plausibility illusion is not likely to be broken because the environment is coherent with individuals' expectations of the experience.

The idea that prior knowledge influences the experience of feeling present aligns with the concept of perceptual hypothesis testing discussed earlier (Wirth et al., 2007). Expectations affect both the confirmation of shifting a mental model to achieve a sense of presence and the illusion of plausibility. These concepts help explain how individual differences can impact the experience of presence. If individuals have varying expectations about an environment, one

person might have a coherent experience of plausibility, while another might not. The individual who maintains their sense of plausibility is likely to feel present, whereas the other may not.

While differences in past experience and prior knowledge adequately explain the varying levels of presence individuals experience, they do not capture the bottom-up differences highlighted above (e.g. Nunez, 2007). Instead, these sensory-driven differences are better captured by the concept of place illusion. The place illusion overlaps nicely with the idea of a mental model, as the mental model of an environment provides the basis of the illusion. Slater (2009) highlights that multiple place illusions can exist from multiple modalities, provided they do not contradict each other. This is similar to the formation of multiple mental models based on sensory information from different modalities, reinforcing the notion that an individual's ability to process sensory information offers a secondary explanation for the individual differences observed in the experience of presence.

Additionally, when considering both metal models and the place illusion, if individuals can maintain a unified perception of multisensory stimuli, they are less likely to separate the information into two distinct mental models/place illusions. Conversely, if the multisensory information is presented too far apart, it is only processed within its own modality, resulting in a mental model/illusion for each. A problem arises if the multisensory information is too far apart to be considered a single stream of information, yet too close together to be considered distinct. Individuals are likely to notice the inconsistent information which breaks the illusion (similar to a badly dubbed film). Nonetheless, individuals differ in how far apart this information can be before splitting into two separate streams. This multisensory window may explain some variance of the individual differences in presence from a bottom-up perspective. The following section outlines theories concerning how this multisensory window might occur, focusing on temporally congruent information.

2.2 The Temporal Binding Window

When immersed in a film, it rarely occurs that the actor on the screen is not actually talking to us. In a movie theatre, the screen produces an image of an actor talking, and a surround sound system provides the audience with the narrative. However, this voice is rarely attributed to the sound speaker positioned behind us. That is, until there is a significant enough delay between when the actor's lips mouth the sentence and when the speaker produces the sound. This delay causes a break in the illusion that the voice is coming from the screen (Donohue et al., 2010). This phenomenon occurs due to a mental process known as the temporal binding window (Stein & Meredith, 1993). The temporal binding window (TBW) for integrating multisensory information refers to the window of time, either on or offset from an event, in which multiple sensory information can be perceived as co-occurring (Colonius & Diederich, 2004; Diederich & Colonius, 2009; Stevenson et al., 2013; van Wassenhove et al., 2007; Wallace & Stevenson, 2014). When the audio-visual stimuli in a film are separated by an interval that exceeds this integration window, the illusion no longer occurs. Both in movies and everyday interactions with people and objects, our ability to accurately integrate these multisensory stimuli is essential for quickly detecting and responding to multisensory events.

The ability to integrate asynchronous stimuli suggests a cognitive mechanism responsible for estimating the temporal relatedness of events (Colonius & Diederich, 2004; Meredith et al., 1987; van Wassenhove at al., 2007). Central to this mechanism is the temporal binding window (TBW); the interval during which sensory signals from different modalities are likely to be perceptually combined into a single event (Stein & Meredith, 1993). Rather than acting as a strict boundary, the TBW functions as a probabilistic threshold for simultaneity, representing the likelihood that multiple sensory signals are interpreted as arising

from a common source (Wallace, 2014). For example, participants exposed to audio-visual stimuli at varying stimulus onset asynchronies (SOAs) frequently report a sense of simultaneity even when the delays exceed 100 milliseconds (Wallace & Stevenson, 2014). However, very few individuals perceive these events as simultaneous when the delays approach or exceed 1000 milliseconds (e.g. Stecker, 2018). In the example above, minimal delays between a character's lip movements and their corresponding speech often go unrecognised, provided these sensory signals remain confined within the TBW. Similarly, in virtual reality environments, where an immersive experience depends upon the real-time synchronisation of visual, auditory, and occasionally haptic feedback, even minor discrepancies between sensory modalities can disrupt the sense of presence if they extend beyond this temporal window (Witmer & Singer, 1998). The brain's tolerance for such asynchronies underscores the flexibility of the TBW, permitting disparate sensory inputs to be integrated into a singular perceptual event. This adaptability is crucial for maintaining coherent multisensory experiences in both real and simulated environments.

It is argued that the TBW exists to account for differences in the transmission and processing speeds of sound and light (Sugita & Suzuki, 2003; Wallace & Stevenson, 2014). Light travels at 299,792,458 m/s, whereas sound travels at 343 m/s. As a result, the light emitted from an event which simultaneously produces visual and auditory stimuli will reach its receptor in a shorter time. In addition, the neural pathways which carry either audio or visual information differ in transduction speeds, with the latter being slower (Poppel, 1988, pp. 27-28). This discrepancy—faster transmission of light coupled with slower processing of visual information compared to auditory information—leads to a relatively wide window in which multisensory stimuli from a single event can be perceived simultaneously and perceptually bound (Calvert & Thesen, 2004). In fact, Stein and Meredith (1993, p. 136) estimated that the TBW is typically 100-250ms wide and expands either side of the point of simultaneity.

The width of the TBW, in which multisensory information is integrated, is considered to be important for at least two reasons.

Firstly, behavioural and neurological studies suggest that when stimuli fall outside of this TBW, the integration of multisensory information into a single percept is unlikely to occur (Spence, 2007; Stein & Stanford, 2008). For example, in behavioural studies, as the offset between multisensory stimuli increases, the percentage of perceived simultaneity decreases (e.g. Fujisaki et al., 2004; Hillock et al., 2011; Powers et al., 2009; Stevenson & Wallace, 2013; Stevenson et al., 2013). Typically, when the offset between multisensory stimuli reaches 300ms, it is no longer considered simultaneous by most participants. This ability to separate information with large offsets is important; without accurate perceptual binding of multisensory information, individuals struggle to maintain a unified and coherent perception of their environment (Wallace & Stevenson, 2014). Stein and Meredith (1993) highlighted that the process of accurately binding multisensory information is essential for helping individuals orient themselves within their surroundings.

The second reason is that the width of the TBW allows integration to be highly adaptive. For example, auditory and visual information can still be bound into a single percept even if the event occurs over 30 meters away. Sugita and Suzuki (2003) suggest that the brain can account for sound velocity up to 40 meters away from the event. In addition to adapting to various depth perceptions, the TBW is also adaptive in terms of differing window sizes for different stimuli. Simple audio-visual flash-beeps typically demonstrate the smallest windows, and complex multisensory stimuli, such as speech, produce the largest windows (Wallace & Stevenson, 2014). This is believed to be due to the additional processing time that complex stimuli require within a single modality before being integrated across different modalities (Stevenson & Wallace, 2013).

In addition to adapting to differences in distance and stimuli type, the TBW also shows a degree of adaptability between individuals. Stevenson, Zemtsov, and Wallace (2012) identified a relationship between individuals' window width and their ability to perform on a McGurk effect illusion task, with participants showing a wide range of window widths. More specifically, they identified that those individuals with a larger window regarding visually leading stimuli (where the visual information is presented before the audio) reported fewer McGurk fusions. Alongside these differences between individuals, the TBW has been shown to develop over time, with children and adolescents having a wider window than adults (e.g. Hillock et al., 2011; Hillock-Dunn & Wallace, 2012). This age-related reduction in window width is attributed to a narrowing of a larger audio-leading aspect of the TBW during childhood, resulting in more accurate perceptions in adults.

Adaptability and exclusion work together to maintain an accurate perception of the world around us. The ability to exclude stimuli from integration if the offset is too great increases the chance that the multisensory perception is genuine and accurate. Conversely, adaptability allows a greater degree of accuracy based on the location of the stimuli and what multisensory object produces it.

This TBW is not only important in maintaining a coherent and accurate perception of the world around us, but also comes with behavioural advantages. One such advantage is the ability to identify and respond to paired audio-visual stimuli in a quicker and more accurate fashion (Gielen et al., 1983). Additionally, evidence suggests that bimodal integration helps with the reflective orientation of attention during perceptually challenging tasks (e.g. Santangelo & Spence, 2007). Indeed, as Wallace and Stevenson (2014) highlight, the temporal window is important in some lower-order processes such as faster reaction times (e.g., Colonius & Diederich, 2004; Todd, 1912; Van der Burg et al., 2008), and higher-order processes such as attention (e.g. Koelewijn et al., 2010; Santangelo & Spence, 2007; Talsma et al., 2007;

Talsma et al., 2010) which are dependent on the accuracy of sensory information (Stevenson & Wallace, 2013; Wallace, 2004). The relationship between the TBW and these processes can be examined through various experimental paradigms, which typically align with theoretical models explaining how temporal binding occurs.

2.2.1 *Theoretical Models of the TBW*

Underlying theories of the temporal binding window (TBW) are built on hypothetical models of neuronal activity (Calvert & Thesen, 2004; Meredith et al., 1987; Meredith & Stein, 1983; Stein & Meredith, 1993; Stein & Stanford, 2008). One central assumption is that the TBW depends on convergence areas in the brain being temporally sensitive to multisensory overlap. The binding only occurs, providing periods of neuron activity coincide within this window width (Stein & Stanford, 2008).

Cognitive psychologists have theorised several computational models to explain how this TBW leads to integration. These computational models overlap with the two main experimental paradigms used to measure the window's width: The Reaction Time Paradigm and the Estimation and Judgement Paradigm. Although both paradigms are thought to follow the same perceptual detection process known as the canonical model (Gibbon & Rutschmann, 1969). The reaction time paradigm measures multisensory integration in terms of a facilitated response (e.g., reacting as quickly as possible to the stimuli presented). In comparison, estimation and judgement tasks involve a forced comparison between two outcomes (e.g., whether two stimuli are presented together or apart).

Specifically, the reaction time paradigm investigates multisensory integration by measuring participants' response speeds to stimuli presented across multiple sensory modalities (e.g., visual and auditory). Experiments typically manipulate the timing or stimulus onset asynchronies (SOAs) between these sensory inputs to determine which temporal intervals

produce the fastest combined multisensory responses (Colonius & Diederich, 2004; Miller, 1982, 1986; Stevenson & Wallace, 2013). Within this framework, two main theoretical models prevail. First, race models (Raab, 1962) suggest that each sensory modality independently races toward a neuronal response threshold, with faster multisensory responses occurring simply because the presence of multiple stimuli increases the likelihood that one reaches the threshold quickly. In contrast, coactivation models (Miller, 1982, 1986) advocate for neural integration, proposing that multiple sensory signals are combined, resulting in quicker and more efficient responses than what unisensory inputs can provide.

Alternatively, the estimation and judgment paradigm measures multisensory integration through explicit perceptual evaluations, focusing on participants' judgements about whether sensory events are simultaneous (Simultaneity Judgement; SJ) or determining the sequence in which they occur (Temporal Order Judgement; TOJ) (Spence et al., 2003; Stevenson & Wallace, 2013; Van der Burg et al., 2008). Unlike reaction-time tasks, these paradigms prioritise perceptual accuracy rather than speed. Consequently, the theoretical models linked to estimation and judgment tasks focus on how the brain combines sensory cues based on their reliability and consistency. Optimal cue integration models (Oruç et al., 2003; Ernst, 2012) propose that sensory cues are weighted according to their reliability, achieving statistically optimal integration. More advanced models, such as the Bayesian coupling prior model (Ernst, 2012), incorporate prior knowledge or expectations to resolve discrepancies between sensory modalities, distinguishing between sensory reliability and perceptual biases. Additionally, correlation detector models (Parise & Ernst, 2016) focus on the temporal aspects of integration, detailing how the brain assesses correlations and temporal offsets between sensory signals.

The following sections discuss the different computational models which map onto each of these experimental paradigms. Results obtained from experiments within the reaction

time paradigm are explained via race models (Raab, 1962) and coactivation models (Miller, 1982). Meanwhile, cue combination models and correlation detector models (Oruç et al., 2003; Parise & Ernst, 2016) are based on data provided by experiments within the estimation and judgement paradigms. Due to the differences in data produced by either paradigm, they are typically theorised separately. However, models have been proposed which can provide a theoretical explanation for both paradigms (Alcalá-Quintana & García-Pérez, 2013; Diederich & Colonius, 2015).

2.2.2 *Theoretical Underpinnings of the Reaction Time Paradigm*

Early computational models of multisensory integration in reaction time paradigms focused on separate neuronal activation (Hershenson, 1962; Raab, 1962). These separate activation models were more commonly known as race models, as the models suggested that there is a race between the cross-modal neurons to reach a response. Race models suggested that a decrease in multisensory reaction time was simply a statistical facilitation of one unisensory channel reaching this response level quicker than another (Raab, 1962). In multisensory trials where 'τ' represents the stimulus onset asynchrony (SOA), some fast responses of the 'slower' modality will replace the slow responses of the 'faster' modality:

$$\Sigma RT_{VA+\tau} \equiv \Sigma \min\{RT_v, RT_{A+\tau}\} \leq \min\{\Sigma RT_v, \Sigma RT_{VA+\tau}\} \quad (1)$$

However, this interpretation is problematic in light of behavioural and neuronal evidence, which demonstrates that multisensory interactions provide faster responses than any unisensory response does alone (e.g. Driver & Spence, 2000; Gielen et al., 1983). In addition, the race model put forth by Raab (1962) assumes a level of stochastic independence (the two processes randomly occurring without relation). As there is no empirical method to test the stochastic dependency of simultaneous random audio, visual and audio-visual reaction times, a race model which is stochastic independent can exist under the assumption of context

invariance (Gondan & Minakata, 2016). This is where the upper and lower margins of audio-visual reaction times ($H_{VA}(s,t)$) are equal to the cumulative distribution of the visual ($F_V(s)$) and audio ($F_A(t)$) reactions, if spatial and temporal differences are infinitely non-negative (Colonius, 1990; Colonius & Diederich, 2018). However, the marginal distributions of $H_{VA}(s,t)$ are not observable in the reaction time paradigm. Consequently, Miller (2016) argues that if context invariance/independence is removed from the race model equation, the model becomes unfalsifiable. Unfalsifiable because the slower stimuli are free to have any distributions which would always satisfy the race model equation (Miller, 2016). In addition, Miller (1982) observed that rewriting the race model equation to remove the constant non-negative component leads to an equation of race model inequality:

$$F_V(t) + F_A(t) = F_{VA}(t) \leq \mathrm{P}(RT_V \leq t) + \mathrm{P}(RT_A \leq t). \quad (2)$$

This race model inequality can be employed to empirically investigate whether reaction times to multisensory stimuli are faster than predicted by a race mechanism (Colonius & Diederich, 2018; Mégevand et al., 2013; Miller, 1982, 1986). When the race model inequality equation is violated, the behavioural evidence points toward a level of multisensory integration occurring above the level of statistical facilitation (e.g. Driver & Spence, 2000; Gielen et al., 1983; Miller, 1986). Rather than statistical facilitation, the combined response is faster than simply one channel reaching the activation level faster than another. Miller (1982) argued that these violations of inequality provide evidence for a coactivation model rather than the separate activation of the race models. Coactivation models put forward a better account of more recent evidence for the reaction time paradigm.

Miller proposed that multiple sensory inputs are combined to satisfy a "criterion level" that initiates a neuronal response (Miller, 1982, 1986). There is a decrease in response time because the combined response reaches this criterion level quicker than a single sensory

neuronal spike (Miller, 1982). Within the coactivation models, the idea of a time window is introduced, suggesting that all information about the stimuli must be contained within a specific timeframe for the neuronal response to accumulate at the criterion level (Miller, 1986).

The coactivation approach can be divided into discrete or continuous accumulation models. Discrete counter models assume that when a stimulus is presented, it produces a discrete "counter" for each event, which is analogous to a neuron spike. These counters are internally represented for a set period of time. The event counters are summated when another stimuli event occurs, providing the event is synchronous with the first. Visual and auditory counters can be represented by $N_V(t)$ and $N_A(t)$, where bimodal responses combine to form $N_{VA}(t)$. The probability that the number of stimuli counters is equal to or greater than the fixed criterion number (C), is equivalent to the probability that the random waiting time for the criterion count (S_c) to be reached is less than or equal to the time passed:

$$P(S_c \leq t) = P(N(t) \geq C). \quad (3)$$

The Poisson counting process provides a quantitative prediction for the expected sum of S_c. It involves a constant (λ) known as the stimuli intensity parameter. This Poisson process produces an expectation of S_c, which is determined by a subjective criterion level (c) based on task dependency,

$$N_{VA}(t) = \mathbf{E}S_c = c/(\lambda_V + \lambda_A) \quad (4)$$

According to Miller (1982), the subjective criterion level (c) is raised in tasks which require a higher degree of accuracy and reduced in tasks which require faster reaction times (Colonius & Diederich, 2018; Miller, 1982, 1986; Schwarz, 1989). However, this model predicts that an increase in the stimuli intensity would lead to a persistent decrease in response times. Yet, seminal evidence reported by Meredith and Stein (1983) and consistently afterwards (e.g. Holmes, 2007; van de Rijt et al., 2019) highlights that this is not the case. On

the contrary, the rule of 'inverse effectiveness' suggests that smaller intensities lead to larger multisensory integrations. This is due to high-intensity unisensory stimuli requiring no aid from additional modalities, whereas smaller intensities benefit from the combination (Meredith & Stein, 1983). This discrete accumulation process model cannot account for this rule. In comparison, the continuous diffusion model (Diederich, 1995) can provide an explanation of the ‘inverse effectiveness’ rule.

Similar to the counter model, the diffusion model also identifies a target threshold level. However, rather than counters, the diffusion model has drift parameters (μ), which account for the stimuli drifting toward the desired threshold known as the absorption boundary (c). This absorption boundary is comparable to the criterion level in the counter models. The expected time to reach this threshold is given by the following equation,

$$\mathbf{E}T_{VA} = c/(\mu_V + \mu_A). \tag{5}$$

This diffusion model theorises that a signal always rests between a lower boundary and this upper boundary (c). The signal drifts toward the upper boundary in a random walk pattern over a continuous time. Diederich (1992, 1995) adjusted the diffusion model by calculating the drift rate through a process known as the 'Ornstein-Uhlenbeck process'. Through this process, the drift of the signal (μ) is calculated by the force (δ) directed toward the upper boundary minus the proportional magnitude of the distance to the signal origin known as the decay parameter (γx) (Diederich, 1995). The super-positioning of the multisensory stimuli allows the overall response to reach the upper boundary quicker, where it is absorbed (Colonius & Diederich, 2018). By calculating the drift rate using the Ornstein-Uhlenbeck process, the diffusion model allows for inverse effectiveness (Diederich & Colonius, 2004) when the drift rate equals

$$\mu(VA) = (\delta_V + \delta_A)[1 + (\delta^{max}{}_V - \delta_V)\,(\delta^{max}{}_A - \delta_A)]. \tag{6}$$

This calculation captures both additive properties when one stimulus is at maximum but also accommodates inverse effectiveness when both stimuli are away from the maximum. This aligns with behavioural data from reaction time tasks (e.g., Drugowisch et al., 2014). However, neither diffusion nor counter models can be utilised to explain behavioural data from estimation and judgement tasks.

2.2.3 *Theoretical Underpinnings of the Estimation and Judgement Paradigm*

Responses to estimation and judgement tasks necessitate a mental process different from reaction time responses. Since the estimation and judgement paradigm requires a decision-making process underpinned by integration, it has been theorised to involve both intramodal properties, like intensity, and supramodal properties, including locational and shape-based cues (Colonius & Diederich, 2018). Given the involvement of both intra- and supramodal properties, models of cue combination are better suited to provide the theoretical underpinnings of the estimation and judgement paradigm. The following section discusses the central principles of cue combination models with reference to the most relevant and well-developed theories.

The optimal integration model is a cue combination model which assumes the value of each cue in isolation is the true unbiased value of the property (Oruç et al., 2003). The objective of the observer is to reduce multiple cue-derived estimates (S_n) to one single estimate (S). Furthermore, these cues are assigned a non-negative weight (w), which is restricted to a sum of 1. This is known as a weighted linear combination,

$$s_w = \sum_{i=1}^{n} w_i S_i \tag{7}$$

The variance of the single estimates (S_w) is calculated based on the individual cue variances ($\sigma^2_i > 0$) and the chosen weights (w_i). The optimal weight for each cue is determined by dividing its variance by the sum of all variances. To achieve optimal multisensory

integration, these weights must also be proportional to the reliability of the cues. The reliability of each cue is considered to be the reciprocal of its variance, which means that the combined estimate will be most reliable when summed (Oruç et al., 2003). Consequently, the stimulus that provides the most reliable information has a greater influence on the final perception.

However, Oruç and colleagues (2003) point out that this model assumes the cues are both uncorrelated and follow a random Gaussian distribution. They suggest that when cues are correlated, the weights should be adjusted to account for the reliability of each cue.

$$r_i' = r_i - \rho\sqrt{r_1 r_2} \tag{8}$$

Once the reliability has been corrected for correlation, a new optimal weight choice can be calculated,

$$w = \frac{r_1'}{r_1' + r_2'} \tag{9}$$

This updated weight calculation for correlating cues provides a potential for negative weights. When the reliability of cues increases to a certain point, a smaller correlation is needed with another cue, which may lead to a negative optimum weight. This situation occurs when cues are integrated with uncorrelated sources of noise (Ernst, 2012; Oruç et al., 2003). However, an issue arises when there is a large discrepancy between the cues. Due to the optimal integration model assuming cues are presented unbiased, the models cannot account for whether discrepancies are due to noise (i.e. imprecision) or bias (i.e. inaccuracy) in the environment (Colonius & Diederich, 2018).

The ‘Bayesian coupling prior model’ provides a way to overcome this uncertainty. Ernst (2012) argues that the perceptual system has a priori knowledge about the co-occurring cues and provides a coupling prior which predicts how much one stimulus should reflect another. For instance, the visual stimuli of a hammer hitting metal should give a cue

surrounding what auditory stimuli to expect. The coupling prior involves several calculations in order to work out the maximum likelihood estimate (MLE) of the discrepancy (D^{MLE}) between two sensory estimates (S_i^{MLE}). The sensory estimates (S_i) are assumed to be made up of both real-world attributes (Sw_i) and this inaccurate bias (B_i). Additionally, sensory estimates are also corrupted by noise (ε), which has sensory variance (σ^2_i). This results in a sensory measurement (z_i), which is formed of ($S_i + \varepsilon$). Since noise is assumed to be independent; the covariance is considered to be zero, leading to a joint likelihood function with N_1 as a normalising factor,

$$p(z|S) = \frac{1}{N_1} exp\left[-\frac{1}{2}\left(\frac{(z_1-s_1)^2}{\sigma_1^2} + \frac{(z_2-s_2)^2}{\sigma_2^2}\right)\right] \quad (10)$$

Using Bayes' rule, a posteriori distribution can be calculated, which considers this a priori knowledge relating to bias ($p(S_i)$), and a measurement of noise,

$$\rho(S_i) \times p(z_i|S_i) \propto \rho(S_i|z_i) \quad (11)$$

This Bayesian calculation provides a maximum a-posteriori (MAP) estimate, enabling cue combination models to account for discrepancies that arise when not assuming unbiases. By comparing the MAP estimate to the MLE, the Bayesian model can identify whether the discrepancies are more likely due to bias or noise. Additionally, when there is no a priori knowledge regarding the stimuli, the MLE serves as the basis for updating perceptual knowledge to become the MAP (Colonius & Diederich, 2018; Ernst, 2012).

These optimal cue integration models have been found to account for behavioural data for both audio-visual cues to location (e.g. Alais & Burr, 2004) and visual-tactile cues to size (e.g. Ernst & Banks, 2002). However, the model has not fully accounted for findings from other studies, including the localisation of audio-visual cues (e.g. Battaglia et al., 2003) and visual-haptic cues (Rosas et al., 2007) of slanted textures (*for a review, see* Landy et al., 2012). Thus,

it appears that optimal cue integration models might not fully capture the computational processes of the TBW.

Alternative models that operate in the estimation and judgement task paradigm are known as correlation detector models. Much like all of the models discussed here, the basis of these models is identifying whether cross-modal information originates from a single event and should be integrated. However, the multisensory correlation detector model (MCD: Parise & Ernst, 2016) differs from these other models in how it addresses multisensory information. Unlike the other models that only calculate the correlation between stimuli, the MCD model calculates the correlation and temporal lag between the stimulus information of the multisensory modalities. This approach is based on an adaptation of the Hassenstein–Reichardt detector model (Hassenstein & Reichardt, 1956).

The MCD model (Parise & Ernst, 2016) consists of five stages. In the first stage, audio and visual sub-units receive input from spatially aligned visual and auditory receptors. Next, these signals are processed by temporal 'low-pass' filters (Burr et al., 2009), which account for specific characteristics of each modality, such as differences in transmission and transduction speeds. The temporal filters refer to the modality (*mod*) and the temporal constants of the filter (τ),

$$f_{mod}(t) = t\exp(-t/\tau_{\mathrm{mod}}) \tag{12}$$

The next stage involves the filtered inputs entering two separate mirror-symmetric subunits (u_1, u_2), both of which are tuned to the temporal order of the signals ($S_V(t)$, $S_A(t)$). These subunits apply a temporal shift to one of the signals through another 'low-pass' filter and then multiply both signals.

$$u_1(t) = \{[s_A(t) * f_A(t)] * f_{VA}(t)\} \times [S_V(t) * f_V(t)]$$

$$u_2(t) = [s_A(t) * f_A(t)] \times \{[S_V(t) * f_V(t)] * f_{VA}(t)\} \quad (13)$$

The outputs of the two subunits are then either multiplied to obtain a correlation detection representation or subtracted to produce a lag detection representation.

$$MCD_{Corr}(t) = u_1(t) \times u_2(t)$$

$$MCD_{Lag}(t) = -u_1(t) + u_2(t) \quad (14)$$

These figures are then turned into a summary variable by averaging out the output over a time window,

$$\overline{MCD_{Corr}} = mean[MCD_{Corr}(t)]$$

$$\overline{MCD_{Lag}} = mean\left[MCD_{Lag}(t)\right] \quad (15)$$

The $\overline{MCD_{Corr}}$ variable can be used in relation to cognitive tasks such as two alternative forced choice (2AFC) tasks, which measure synchrony judgements. This is due to the variable increasing closer to the point of simultaneity and decreasing afterwards. In comparison, the $\overline{MCD_{Lag}}$ variable is more suited to temporal order judgment (TOJ) tasks. This is because the variable follows the same trend in terms of probability values, starting at a low probability for audio-first responses when visual is presented first and increasing to a high value when audio is presented first.

Although Parise and Ernst (2016) have no explicit claim on the criterion level at which the TBW rests within their model, their model replicates the shape of standard task response data. Furthermore, Parise and Ernst (2016) tested the models to determine if they fit previously published task results. For the MCDLag variable, they compared the model's fit to data from Spence and colleagues (2003) Experiment 2, which investigated spatially non-correspondent audio-visual stimuli via a TOJ task. The $\overline{MCD_{Lag}}$ fitted directly on the experimental data (Spence et al., 2003). In addition to the TOJ task, Parise and Ernst (2016) generated model

responses to Slutsky and Recanzone's (2001) 2-AFC simultaneity judgement task. The model again was able to fit the response data, this time using the $\overline{MCD_{Corr}}$ calculation. In addition to the parsimonious fit to published data, the MCD model bridges the gap between neuronal data and behavioural data, as it can also be used to model the responses of cross-modal neurons. However, Colonius and Diederich (2018) suggest two possible shortcomings of the multisensory correlation detector model.

The first possible shortcoming of the MCD model is that the model calculates the audio-visual filter temporal constant to be 786ms. The literature largely reports that multisensory integration occurs over much smaller time scales between 100 and 300ms (e.g. Stein & Meredith, 1993; Wallace & Stevenson, 2014). Secondly, the model assumes spatial correspondence, with no calculation to account for its occurrence or lack thereof. Previous studies have identified that inputs which are not spatially correspondent can be integrated into a single percept, providing they are still temporally aligned (e.g. Van der Burg et al., 2008). The current version of the multisensory correlation model does not allow for these spatial differences to be accounted for by the individual subunits. Alternatively, Colonius and Diederich's (2004, 2015) time-window-of-integration model (TWIN) allows stimuli with a large spatial difference to remain unified, providing they co-occur within this TBW.

2.2.4 *The TWIN Model*

Developed within the reaction time paradigm, the TWIN model combined features of both race and coactivation models (Colonius & Diederich, 2004b). At its most basic level, the model contains two stages: a race stage and an integration stage. However, the model requires several important assumptions off the back of these stages.

The first stage assumption is related to the initial processes, which assume a cross-modal stimulus triggers a race between peripheral unisensory neurons. The second assumption

is that a further stage must integrate the neural inputs and prepare a motor response. While these two primary assumptions may simplify the complexities of these processes, Colonius and Diederich (2015, 2018) highlight that the model serves as a framework in which alternative models can be developed and formulated.

At the centre of the framework exists the TWIN assumption, which states that multisensory integration only occurs if the processes of the peripheral sensory pathways terminate within a given temporal interval. Similar to the work of Parise and Ernst (2016), this acts as a temporal filter, triggering the multisensory integration process. However, unlike Parise and Ernst's (2016) MCD model, the TWIN model also incorporates an assumption of spatiotemporal separability. This assumption considers the degree of interaction during the second stage of processing as a function of the spatial arrangement of the stimuli, while assuming that this interaction does not depend on their spatial asynchrony (Diederich & Colonius, 2009). This 'temporal tuning' of the interaction occurs because the probability of integration, $P(I)$, is modulated by the "TWIN filter". This filter assumes that multisensory integration (I) is modulated by stimulus-onset asynchrony (τ) and the TBW width (ω). This is achieved by assigning 'statistically independent' random non-negative values to visual (V) and audio (A) stimuli, which leads to the equation,

$$I = \{|V - (A + \tau)| < \omega\} = \{A + \tau < V < A + \tau + \omega\} \cup \{V < A + \tau < V + \omega\} \tag{16}$$

The presentation of the visual stimuli is considered an arbitrarily zero point. The numerical definition of the probability of integration, $P(I)$, can be determined based on the distribution functions of V and A. The distributions of the peripheral processing times of V and A are assumed to be exponential with parameters of λ_v and λ_a,

$$F_V(t) = 1 - \exp[-\lambda_v t]$$

$$F_A(t) = 1 - \exp[-\lambda_a t]$$

(17)

The probability of integration can then be used to calculate the mean reaction time E[RT] in the cross-modal condition where S_1 and S_2 refer to the first and second stages,

$$E\left[RT_{VA,\tau}\right] = E[S_1] + P(I)E[S_2|I] + [1 - P(I)]E[S_2|I^c] = E[S_1] + E[S_2|I^c] - P(I) \times \Delta.$$
$$= E[\min(V, A + \tau) + E[S_2|I] - P(I) \times \Delta.$$

(18)

It is noted that I^c denotes an event complementary to I, and that Δ stands for $E[S_2|I^c] - E[S_2|I]$. In addition, the term $P(I) \times \Delta$ provides a measure of the expected amount of multisensory integration, which can be positive or negative and relates to facilitation or inhibition.

Despite Colonius and Diederich (2004) highlighting that in its original form, the TWIN model is oversimplified, the model has provided a good fit for reaction time data. This includes applications of the TWIN model to saccadic reaction times for audio-visual stimuli ((e.g. Diederich & Colonius, 2008) and visual-tactile stimuli (e.g. Diederich & Colonius, 2007, 2008). Additionally, the model has been successfully applied to data from audio-visual perceptual localisation tasks (Glazebrook et al., 2016). Furthermore, Diederich and Colonius (2015) demonstrated that a simple extension of the TWIN model can account for data from both paradigms, such as responses to judgment tasks.

Diederich and Colonius (2015) extended the TWIN model using a canonical model (Gibbon & Rutschmann, 1969) to reanalyse data from Mégevand et al. (2013). In this extension of the TWIN model, the first stage, known as the race stage, replaces the canonical model's initial stage. Consequently, the idea that the initial perceptual stages for both RT and

temporal order judgement tasks are identical is linked to the processing times of the visual (*V)* and auditory (*A*) stimuli.

Additionally, it is argued that judgements of temporal order are modulated by the temporal window of integration. This is due to the assumption that when detection times for the stimuli fall within the window width, participants cannot make a judgement based on sensory evidence available and subsequently respond with a 'visual-first' bias (β). Only when stimuli do not arrive within the integration window can participants accurately respond to the temporal order of the stimuli. Temporal order judgement tasks typically produce data which is in a relative frequency format $\psi(\tau)$ of responses to a range of SOAs (τ) where,

$$\psi(\tau) = \Pr(\text{visual first}|\ visual\ stimulus\ presented\ \tau\ ms\ before\ acoustic)$$

(19)

When considering the arrival time of stimuli with the assumption of visual first bias, the TWIN calculation for temporal order judgement tasks becomes,

$$\begin{aligned}\psi(\tau) &= \Pr(I) \times \beta + \Pr(V + \omega < A + \tau) \\ &= \{\Pr(A + \tau < V < A + \tau + \omega) + \Pr(V < A + \tau < V + \omega)\} \times \beta \\ &\quad + \Pr(V + \omega < A + \tau)\end{aligned}$$

(20)

However, when there is no prior visual bias, such as when β=0.5, positive and negative SOAs must be considered separately,

$$\psi(\tau) = \begin{cases}\psi^{+}(\tau), & if\ \tau \geq 0; \\ \psi^{-}(\tau), & if\ \tau < 0.\end{cases}$$

(21)

The window width acts as a model parameter, controlling both the reaction time and the probability of judging the temporal order of audio-visual stimuli.

Overall, the TWIN model is well suited for estimating individual differences in TBWs using data from reaction time and perceptual judgment tasks. By accounting for both lower- and higher-order processes involved in multisensory integration, the model provides a flexible theoretical framework for understanding how sensory inputs are processed and integrated. It highlights that multisensory integration depends on several key factors, including the temporal offset of sensory information, the processing times of each modality, and the width of an individual's integration window.

This ability to estimate TBW size across two distinct paradigms—reflecting different levels of sensory processing—makes the TWIN model particularly suitable for investigating perceptual phenomena like presence in virtual environments. By utilising these paradigms, the model provides insights into how temporal binding operates across both basic reaction processes and higher-order perceptual judgments, even though the precise role of temporal binding in presence remains unclear.

By applying the TWIN model to estimate individual TBWs, we can explore how these differences influence the coherence of multisensory experiences in VR. This line of inquiry not only sheds light on individual variability in presence but also contributes to expanding theoretical frameworks by integrating temporal binding as a key factor in presence theory.

2.3 Presence and The Temporal Binding Window

Theoretical models of presence encompass a variety of sensory and cognitive factors that lead to an individual feeling situated within an environment. While sensory input alone does not fully explain why individuals experience virtual environments differently, it is clear that sensory factors play a significant role, as presence is primarily a perceptual phenomenon.

Research has begun to explore how the integration of multiple senses relates to presence (Kim & Lee, 2022; Marucci et al., 2021), with some studies linking this integration to phenomena like cybersickness (Sadiq, 2019; Sadiq & Barnett-Cowan, 2022), often considered the antithesis of presence (Weech et al., 2019). Despite these advances, there remains a notable gap in understanding how individual differences in temporal binding—specifically, the temporal binding window (TBW)—may contribute to varying levels of presence in virtual reality (VR).

Virtual environments present information that is significantly different from the real world. In the real world, events are typically perceived from considerable distances, whereas in VR, sensory stimuli are presented mere centimetres away from their sensory receptors. Moreover, multisensory information produced by VR systems is often slightly offset, a phenomenon referred to as temporal asynchrony (Tachibana & Matsumiya, 2022). The temporal misalignment of these sensory cues, alongside variations in sensory processing times, may disrupt the integration process for some individuals, leading to difficulty in binding multisensory information into a coherent mental representation of the environment. When multisensory cues are not effectively integrated, individuals may experience a fragmented representation of the virtual world, which could negatively affect their sense of presence (Wirth et al., 2007).

Considering the importance of temporal binding in multisensory integration, individual differences in the width of one's TBW may play a critical role in shaping the experience of presence. Individuals with a smaller TBW may struggle to integrate sensory information that arrives slightly out of sync, which may hinder the formation of a unified perception of the virtual environment. In contrast, those with a broader TBW may be better equipped to process asynchronous sensory cues, facilitating a more seamless integration of visual and auditory stimuli and thereby enhancing the feeling of being "present" in the virtual world. This

variability in the ability to integrate sensory information could explain why some individuals experience higher levels of presence in VR, while others feel disconnected or disoriented.

Despite these insights, the question of how individual differences in temporal binding influence presence remains largely unexplored. Given the potential for TBWs to modulate multisensory integration and, consequently, the sense of presence in VR, this research aims to answer the following question: Do individual differences in cross-modal integration, as determined by the width of one's TBW, impact the feeling of presence within virtual reality? Furthermore, if such a relationship exists, are there other moderating factors—such as cybersickness—that might influence how temporal binding affects presence in VR?

Chapter 3: Exploratory Analysis of Correlation between TBW and Presence (Study 1)

3.1 Introduction

Processing sensory information is a fundamental part of how individuals experience the world. Individuals constantly receive multiple sensory cues regarding objects and events within their environment (Hillis et al., 2002). Some of these objects provide unisensory cues. For example, when viewed from a distance, stationary and silent objects (e.g., walls) provide only visual cues. Other objects/events within an individual's environment provide sensory cues to multiple modalities. When information regarding the environment is multisensory, individuals benefit from correctly identifying whether sensory information originated from the same object or event (Stein & Meredith, 1990). There are two main benefits for individuals. Firstly, multisensory information relating to a singular object/event increases the speed and accuracy of recognising the object/event within the environment (e.g. Gielen et al., 1983; Santangelo & Spence, 2007). Faster and more accurate recognition occurs because multisensory information is processed more rapidly than unisensory information (Todd, 1912). The second benefit to individuals is the ability to orient oneself within the environment (Stein & Meredith, 1993). When sensory information about objects and events is multisensory, individuals can orient themselves more quickly, as the mental model of the environment, known as a reference frame, is established faster than in unisensory environments (Nardini et al., 2008).

In order to identify whether multisensory information originated from a single object/event, individuals are thought to utilise a specific cognitive process. This cognitive process is known as the temporal binding window (Colonius & Diederich, 2004). The temporal binding window accounts for the physical differences between sensory information produced from a singular event/object (Sugita & Suzuki, 2003; Wallace & Stevenson, 2014). Specifically, sensory information from different modalities has varying transmission and

transduction speeds (Poppel, 1988, pp. 27-28). The length of time the sensory information takes to reach a sensory receptor from its point of origin is known as the transmission speed. On the other hand, the transduction speed is the length of time for the information to pass from the sensory receptor to the region of the brain responsible for processing (e.g., eye to visual cortex; Marzvanyan & Alhawaj, 2023). When an object or event contains audio and visual information, such as a person talking, one of the modalities may be processed faster because of differences in both transmission and transduction speeds across different senses (Calvert & Thesen, 2004). Without a cognitive process to adjust for this, the information may be attributed to two separate events. One event representing the visual information (e.g., lips moving), and another representing the auditory information (e.g., voice). The temporal binding window accounts for the difference between the faster and slower modalities (Stein & Meredith, 1993). Providing that the information from both modalities is processed within a close enough time frame to one another (e.g. the temporal binding window), the information is combined into a unified percept (Meredith et al., 1987). The width of the temporal binding window automatically determines whether sensory information relates to a singular event based on the time offset (Stein & Meredith, 1993).

The estimated size of an individual's temporal binding window differs based on several factors (Wallace & Stevenson, 2014). The first factor is the task used to estimate the size of an individual's binding window (Stevenson & Wallace, 2013). Reaction time tasks require individuals to respond as quickly as possible to the first stimuli detected (Todd, 1912). Tasks involving these simple reaction times to offset multisensory information typically report the smallest window widths (Diederich & Colonius, 2015). This is due to the measurable speed of an individual’s reaction time decreasing due to a process known as multisensory facilitation (Miller, 2016). The response becomes quicker when information from two or more senses is processed together (Miller, 2016). Temporal order judgement tasks differ from these reaction

time tasks in how they ask individuals to respond. Rather than measuring a simple reaction time, individuals must decide which modality was presented first (e.g. Bald et al., 1942). The additional process of making a judgement leads to larger estimated window widths (Megevand et al., 2013).

The type of stimuli used within tasks also impacts the estimated size of an individual's window width (Stevenson & Wallace, 2013). Flash-bleep stimuli, where a brief visual flash is presented simultaneously or very close in time with an auditory beep, are the easiest to process (Van Eijk et al., 2008). These stimuli typically produce the smallest estimates of the width of an individual's binding window (Stevenson & Wallace, 2013). This is because they usually contain small, noncomplex information regarding an item, such as a checkerboard (Powers et al., 2009).

Speech stimuli (e.g. McGurk effect; MacDonald, 2018), which consist of visual mouth movements and paired auditory words, are much more complex (Vatakis et al., 2008). Moving visual stimuli provide more information regarding an object than stationary visual stimuli (Stevenson & Wallace, 2013). In addition, the auditory information from speech is richer than bleep stimuli as speech provides varying tones and pitches from words (Vatakis et al., 2008). As the complexity of multisensory information from an object or event increases, the estimated size of the temporal binding window also increases.

In addition to the differences between task demands and stimulus complexity, individuals naturally differ in the size of their temporal binding window (Stevenson et al., 2012). Some individuals have narrower windows than others. When individuals have a narrow window width, multisensory information is perceived to be incongruent at shorter offsets from a simultaneous presentation (Colonius & Diederich, 2018). These narrower windows make it less likely for offset information from multiple sources to be combined into a single percept

(Stein & Meredith, 1993). Integrating less offset information, because of a narrow window, can lead to more accurate representations of the environment (De Gelder & Bertelson, 2003). These representations are accurate because they reduce the chance of combining two unrelated events into a singular percept (e.g., the McGurk effect; McGurk & MacDonald, 1976). The different window widths lead to behavioural and perceptual differences between individuals. For example, when viewing a film where multimodal information is presented simultaneously, both individuals with a narrow or wide window width would perceive the multisensory event as a singular percept. If the auditory information presented from the technology became offset by 100ms, such as in a badly dubbed film, individuals with a broader window would not likely notice this offset. The multisensory information would still be processed together as a single percept. On the other hand, individuals with a narrow window would no longer combine the modalities as a single percept and process the visual information and the auditory as two separate events/objects.

These differences between individuals should be an important consideration when presenting individuals with information about their environment in an unusual way, for example, the presentation of a virtual environment through a head-mounted display (HMD). A VR headset displaying a virtual environment contains visual and auditory information, which originates much closer to the sensory receptors than a ‘real world’ environment (Rempel et al., 2007). Despite the information's reduced transmission time, inconsistencies between multisensory events/objects may arise. The inconsistencies are due to the technology updating the visual information rather than relying on the environment producing the information (Tachibana & Matsumiyu, 2022). VR headsets may suffer from delays in processing/updating this information, known as ‘latency’ (Raaen & Kjellmo, 2015). A large latency could offset the visual and auditory information by several milliseconds (Raaen & Kjellmo, 2015). Some

individuals may recognise the latency, whereas others may not. This ability could be due to an individual's tolerance for offset information, i.e., the size of their temporal binding window.

Efficient processing of sensory information is important for individuals to feel present within a virtual environment (Steuer, 1992; Wirth et al., 2007; Witmer & Singer, 1998). Witmer and Singer were amongst the first to highlight that sensory information should be consistent within a multisensory virtual environment (1998). When presenting multisensory information about an object or event in a virtual environment, it should be done simultaneously. This concept is based on Steuer's (1992) dimensions that determine 'telepresence'. According to these dimensions, the consistency of the information enhances an individual's sense of presence, which is influenced by both the breadth and depth of the information presented (Steuer, 1992). Depth relates to the richness of the information, while, the breadth of sensory information relates to the number of modalities stimulated. In support of this, it has been reported that virtual environments which present information to multiple modalities increase individuals' sense of presence (e.g. Dinh et al., 1999). The increased presence in the virtual environment is likely due to it capturing more of the individual's attention. As individuals focus more on the information presented, they begin to orient themselves within this virtual space. This orientation has been linked to the additional sensory information, which helps in developing more robust mental models of the virtual environment (Wirth et al., 2007).

Virtual environments which present information to multiple modalities also pose a risk of diminishing presence if the information presented is not congruent (Slater, 2002). Incongruent information is not normally combined into a single percept. The process model of presence hypothesised that incongruent streams of information create multiple egocentric reference frames (Wirth et al., 2007). These are mental representations of the environment (Riecke & Von der Heyde, 2003). The most congruent of these representations is known as the primary egocentric reference frame (PERF). When information from the virtual environment

is incongruent, the PERF will likely be the mental representation of the natural world (Wirth et al., 2007), due to the real world containing more congruent sensory information (e.g., temperature, gravity, tactile feedback from the floor). If this were the case, an individual's sense of presence within the virtual environment would be diminished (Wirth et al., 2007).

Nevertheless, individuals report varying levels of presence whilst exploring the same virtual environment and using the same VR technology. These individual differences might be caused by the ability to maintain a single egocentric reference frame of the virtual world. Individuals with different-sized temporal binding windows may have different propensities to handle the multisensory information presented in virtual reality, leading to different experiences of presence. Individual differences in temporal binding widths may be related to individual differences in the presence experience. Yet little research has been conducted on whether individual differences in reported presence are somewhat accounted for by individuals' ability to integrate multisensory stimuli. This chapter is dedicated to an exploratory study investigating if individuals' temporal binding windows and feelings of presence are related.

3.1.1 *Measures of Presence*

Much like most psychological phenomena, presence can be measured subjectively or objectively. Sheridan (1992) argued that because presence is a subjective mental manifestation, objective measures may not easily identify the phenomena. Specifically, objective measures may not easily identify presence because they focus more on physiological/behavioural components. For example, researchers might monitor how an individual reacts if something is perceived to be thrown at them within the virtual world. This type of objectively measurable response is different from the subjective mental models that are also formed. The mental models, and individuals feeling situated within them, are best assessed via subjective reports.

Sheridan suggested that the two measures could be used together within the correct environment (1992), allowing for both the subjective reporting of presence and a convenient objective measure. The OmniPres compendium (Baren & IJsselsteijn, 2004) outlines how both measures may best be utilised. Objective measures are split into two subcategories: behavioural observations and psychophysiological measures.

When measuring presence, behavioural observations refer to actions individuals perform while experiencing a medium, such as VR. The most common of which include facial expressions (Huang & Alessi, 1999), 'nulling' (Parker & Prothero, 2003), conflicting sensory cues (Slater et al., 1995), automatic social behaviours (Bailenson et al., 2003; Ijsselsteijn et al., 1990; Sheridan,1992), postural responses (Freeman et al., 2000) and reflex responses (Held & Durlach, 1992). These measures present certain challenges that merit consideration. Firstly, researchers can only use behaviour observations in a limited number of situations. For example, 'nulling' requires counterbalancing of self-motion cues between real and virtual worlds, such as investigating individuals' levels of presence when they are spinning in the real world and the virtual world is spinning in the same direction. Welch (2003) highlights that due to the visual feedback delay in VR, this may be inappropriate in most circumstances because information needs to be presented precisely at the right time to counterbalance real-world sensory cues. Besides 'nulling', one issue shared across behavioural observation measures is the difficulty in interpreting the behaviour as simply a correlate of presence (Insko, 2003; Laarni et al., 2015). A posture adjustment, for instance, might be misunderstood as a behaviour influenced by the virtual environment when, in reality, it could simply be a random adjustment made by the individual. The responsibility for accurate interpretation falls on the observer; thus, behavioural observations may not be objective.

Psychophysiological measures involve equipment, rather than observers, to detect changes within an individual. The four main categories of measurement previously used to

explore presence are cardiovascular (e.g., Meehan et al., 2002), skin measures (e.g., Meehan et al., 2002; Ravaja et al., 2004), ocular measures (e.g., Laarni et al., 2015), and electromyography (e.g., Aymerich-Franch, 2010). These psychophysiological measures are more objective than behavioural observations as they detect measurable changes within an individual's physiology (Riva et al., 2003). In order to detect these changes, the virtual environment must be stimulating and arousing, such as environments which cause a fearful response during an individual's virtual experience (e.g. Meehan et al., 2002). The need to generate arousing environments produces limitations on when these measures can be utilised since not all environments include aspects that induce a sizeable physiological response (Riva et al., 2003). Thus, both objective measures outlined here are limited in the range of virtual environments in which they can assess presence. In order to capture individuals' experience of presence across a range of virtual environments, subjective measures may be more reliable (Laarni et al., 2015).

Several subjective tools have been developed to measure presence, including continuous assessment and subjective corroboration (Baren & IJsselsteijn, 2004). However, the most commonly used form of subjective measurement is presence questionnaires (Hein et al., 2018). Their popularity stems from administering them after an individual has experienced the virtual environment, unlike the other measurement methods (Grassini & Laumann, 2020). Completing the questionnaire post-experience provides a measure that does not interfere with the user's experience of VR (Laarni et al., 2015). The following section discusses some of the most prominent presence questionnaires.

The Presence Questionnaire (PQ) (Witmer & Singer, 1994) was one of the first questionnaires to be developed to assess individuals' experience within a virtual environment. Not only was it one of the earliest questionnaires of its kind, but it also remains a widely used and prevalent presence questionnaire (Grassini & Laumann, 2020; Hein et al., 2018). It is widely used due to its inclusion of both 'involvement' and 'immersion' within the assessment

of presence. Involvement relates to the attentional processes of an individual within the virtual reality (Witmer & Singer, 1998). In comparison, Witmer and Singer regarded immersion as a psychological state which occurs when an individual perceives themself to be ‘enveloped by, included in, and interacting with’ the virtual environment (1998). Both of these concepts are considered to contribute toward a sense of presence. The PQ was developed from four factors believed to influence both of these elements of presence. The first factors, *control* and *distraction*, relate to individuals’ actions and selective attention. Meanwhile, *sensory* and *realism* factors relate to the richness of sensory input and the realness of the virtual environment. Whilst the questionnaire has been adjusted over time (Witmer & Singer, 1998; Witmer et al., 2005), these underlying principles have remained the same.

Despite the continued use of the PQ (Grassini & Laumann, 2020), some have raised concerns about its effectiveness in assessing presence. The main argument against using the PQ is its potential lack of validity. This concern stems partly from the questionnaire's inability to separate the concepts of presence and immersion (Slater, 1998). Slater argues that the questions invoke a subjective response to the immersive qualities of the VR system. Therefore, it becomes less a measure of immersion and more a measure of an individual's way of making sense of the world, which does not add to the assessment of immersion or presence (Slater, 1998). Whilst these claims are refuted as an oversimplification (Singer & Witmer, 1998), a questionnaire focusing explicitly on presence may be preferred (e.g. SUS; Slater et al., 1994; SPES; Hartmann et al., 2016).

The SUS (Slater, Usoh and Steed Presence Questionnaire) is a presence questionnaire which focuses on the sense of ‘being there’ within a virtual environment. It was initially developed as a single-factor, 3-item questionnaire assessing spatial presence in virtual reality (Slater et al., 1994). Participants respond to the items using a 7-point scale that ranges from "not at all" to "very much." The original scoring strategy was based on a model (Slater & Usoh,

1993;1994) which considered presence a binary concept. Either an individual felt located within the virtual environment or did not. Only scores of 6 or 7 on the 7-point Likert scale were considered to indicate a sense of 'being there' within the virtual environment. Since its conception, the SUS has been used as a generic measure of presence (e.g. Schubert et al., 2001). The items are scored either in binary form or continuously (Slater et al., 1995; Usoh et al., 1999). A continuous representation of presence indicates that individuals can feel more or less present within the virtual environment (Hartmann et al., 2015). Due to the ability to score the questionnaire items based on different conceptualisations of presence (binary/continuous), the SUS has been integrated into other presence questionnaires. For example, the Igroup Presence Questionnaire (IPQ), which presented a solution to separate the concepts of presence and immersion by introducing the idea that cognitive processes mediate the relationship between the two (Schubert et al., 1999a, 2001).

The IPQ was developed in line with Schubert et al.'s (1999b) embodied presence framework. This framework suggests that presence is formed from a meshed pattern of perceived possible actions within an environment (Glenberg, 1997; Schubert et al., 1999b). As such, an individual's ability to act within their surroundings is considered central to the sense of presence. In order to develop a questionnaire which assessed this central concept, a series of studies were conducted (Schubert et al., 2001). Seventy-five items relating to presence were either newly generated or sourced from other questionnaires, such as the PQ and the SUS (Slater et al., 1994; Witmer & Singer, 1998). Data was collected from 246 and 296 participants across the two studies who experienced a VRE in either an HMD or via a monitor. The 75 items were assessed through principal component and confirmatory factor analyses over two studies. The results of which led to a 14-item questionnaire. The questionnaire was developed by loading the items into three factors: spatial presence, involvement, and realism. A final item was included in the questionnaire as a generic presence factor (an SUS item). This rigorous

approach to development has led to high reliability of the measure (Regenbrecht, 2002; Youngblut, 2003). Alongside the measure explicitly being developed for VR, unlike other measures developed at the time, which are cross-medium. For example, the ITC-Sense of Presence Inventory (Lessiter et al., 2001) and the Temple Presence Inventory (Lombard et al., 2000).

The presence questionnaires highlighted above were developed without strict concordance with a theoretical model (Laarni et al., 2015). Rather than developing questions based on theoretical elements of a model, items were refined to represent an overall concept (e.g. Schubert et al., 2001). For example, the spatial presence scale of the IPQ. Hartmann and colleagues (2016) refer to this as an inductive method of questionnaire development. This is considered inductive because factor analyses generate the components of presence rather than the components stemming from a developed theory. Meaning that mapping participant's responses to the underlying theoretical principles can be tenuous. Wilson and Nichols (2002) highlight a preference for presence questionnaires to be developed using a deductive method. In a deductive approach, the items included in a presence questionnaire are derived directly from an established presence theory (Wilson & Nichols, 2002).

The theoretical model of presence used within this thesis is the process model developed by Wirth and Colleagues (2007). This theoretical model was selected because it separates multiple stages in which multisensory processes can directly impact the formation of presence. This separation allows for an easier comparison of how individuals' temporal binding windows might affect their sense of presence. From this theoretical model, two questionnaires have been developed in a deductive manner: Measurement, Effects, Conditions, Spatial Presence Questionnaire (MEC-SPQ) (Vorderer et al., 2004a) and the Spatial Presence Experience Scale (Hartmann et al., 2016).

The Measurement, Effects, Conditions, Spatial Presence Questionnaire (MEC-SPQ) was developed from 103 items representing the eight constructs of the process model: Attention Allocation, Spatial Situational Model, Self-location, Possible Actions, Higher Cognitive Involvement, Suspension of Disbelief, Domain-Specific Interests, and Visual Spatial Imagery (Vorderer et al., 2004a). The validity and reliability of the items were tested in a multi-media study, investigating the questionnaire's applicability to multiple mediums across 290 participants (Text, Hypertext, Film, and VR; Vorderer et al., 2004b). The results highlighted high internal consistency for 64 items split between the subscales (α=0.78-0.93; Vorderer et al., 2004b). Based on the results of the validation study, a revised final version of the MEC-SPQ was compiled. This final version can be administered in three sizes. The full version contains eight items per construct. The two smaller versions contain six items and four items per construct, respectively (Vorderer et al., 2004a). Despite the difference in size, the items are administered in the same way for each version. Items related to the *domain-specific interest* are administered in a single block. *Visual-spatial imagery* and *higher cognitive involvement* items are presented in a randomised block. The final items relating to the other concepts are also presented in a separate randomised block. Due to this process, the administration of the questionnaire can be time-consuming. An alternate questionnaire, the Spatial Presence Experience Scale (SPES), was created to resolve this issue (Hartmann et al., 2016).

The SPES was created as a "shortened and fine-tuned version of the two spatial presence-related scales of the MEC-SPQ that assess self-location and perceived possible actions" (Hartmann et al., 2016, p. 4). This refined questionnaire is quick to deploy because it contains a total of 8 items. The items were chosen based on 2 studies which employed a deductive approach to the questionnaire construction. Initially, data from 290 participants' responses to 18 items relating to both subscales underwent a principal component analysis. The items that loaded onto the two-factor structure were checked for internal consistency. Items

which had very high internal consistency (e.g., too similarly worded) were excluded along with items with low internal consistency. The remaining eight items are split into two 4-item subscales. A second study using 395 participants' responses to these eight items confirmed the measure's validity using a two-factor confirmatory factor analysis. The study also highlighted that both subscales had good internal consistency (self-location, α=0.91, possible actions, α=0.84). Furthermore, the overall total of the SPES was considered internally consistent (α=0.89).

As shown by Hartmann et al. (2016), the SPES offers a theoretically grounded and empirically validated measure of spatial presence. Building on this foundation, the present study utilises the SPES to assess individuals' sense of presence in a virtual environment. There are two primary reasons for this choice. Firstly, the SPES was deductively constructed in relation to its underpinning theoretical model (Hartmann et al., 2016; Wirth et al., 2007). The way this questionnaire was developed allows for conclusions to be drawn more accurately regarding the cognitive processes involved. For example, individuals' ratings on the questionnaire sub-scales (Self Location and Possible Actions) directly relate to the hypothetical processes occurring for individuals to feel present (Hartmann et al., 2016). The SPES was also selected because of the speed at which the questionnaire can be administered. While it is a validated measure of presence (α=0.89), it is also quick to administer (Hartmann et al., 2016). The short questionnaire can be completed within a few minutes, helping to reduce the overall fatigue effects of completing multiple tasks/questionnaires (Hopstaken et al., 2015). Reducing fatigue helps ensure genuine self-report responses (Iwaniec, 2019).

Alongside demographic questionnaires and a task to stimulate a sense of presence, an additional measure of presence was included to account for the fact that the SPES had not yet been validated beyond its initial development studies. Hartmann et al. (p.12, 2016) also recommended using a second questionnaire to assess the convergent validity of the SPES

further. A high correlation between the two questionnaires would provide evidence for the SPES's ability to measure Spatial Presence. The IPQ was selected as the secondary presence measure (Schubert et al., 2001) as it provides a useful comparison to the SPES. This is because the process model underlying the SPES builds upon Schubert's embodied presence framework (1999a), leading to overlaps between the self-location and possible actions scales in the SPES and the spatial presence and involvement scales in the IPQ (Hartmann et al., 2016; Schubert et al., 2001). In addition to assessing the convergent validity of the SPES, the IPQ also captures an aspect of presence not measured by the SPES: realism. This could provide additional insight into the relationship between an individual's temporal binding window and presence.

3.1.2 *Measures of the Temporal Binding Window (TBW)*

Experiments exploring the temporal binding window commonly involve some form of audio-visual stimuli, such as flash-beeps (e.g., Donohue et al., 2010; Navarra et al., 2005; Stevenson & Wallace, 2013), videos of dynamic tools in action (e.g. *hammers hitting an object*, Stevenson & Wallace, 2013), videos of a speaker uttering single syllable speech (e.g. Stevenson & Wallace, 2013; van Wassenhove et al., 2007; Vatakis & Spence, 2007), or multi-syllable speech (e.g. Navarra et al., 2005; Vatakis & Spence, 2007), and video recordings of a hand playing piano (Navarra et al., 2005). Whilst the majority of studies have focused on audio-visual interaction, studies have also looked at visual-tactile interaction (e.g. Beauchamp, 2005; Diederich & Colonius, 2004; Pasalar et al., 2010), audio-tactile (Foxe, 2009; Navarra et al., 2007) and olfactory interaction (Leonard & Masek, 2014; Thesen et al., 2004). Despite the ability to investigate these other interactions, visual and auditory interactions are known to dominate the senses (Colavita, 1974; Hecht et al., 2005). Sight and hearing account for approximately 90% of the information we perceive about reality (Man & Olchawa, 2018). In addition, these modalities are the primary senses involved in virtual reality research. For these

two reasons, the focus here is on the audio-visual stimuli used in experiments exploring individuals' temporal binding windows.

In order to assess the temporal binding of audio-visual information, stimuli pairs are presented offset from one another. These stimulus onset asynchronies (SOAs) can be audio or visual leading. The stimuli of one modality are presented in a range of intervals before the other. These intervals usually range between ±300ms. Larger intervals than 300ms produce floor effects in terms of response frequencies (Stevenson & Wallace, 2013). After a delay of more than 300ms, individuals detect the two stimuli as occurring separately (e.g. Fujisaki et al., 2004; Hillock et al., 2011; Powers et al., 2009; Stevenson & Wallace, 2013)). As the offset between stimuli reduces closer to the point of true simultaneity (0ms), the frequency of temporally bound response increases (e.g. Stevenson et al., 2013). Once the frequency reaches a certain threshold, it is considered the point at which individuals perceive the two stimuli as co-occurring. The operationalisation of this temporal binding depends upon which cognitive task is utilised to assess the temporal binding window (Colonius & Diederich, 2018). These tasks differ based on which overarching experimental paradigm they exist within: 'Estimation and Judgement' or 'Reaction Time'.

Experiments within the estimation and judgement paradigms arise from the idea that there is a point of simultaneity which is subjective to each individual. As the distance (in milliseconds) from this point of simultaneity increases, the likelihood of stimuli being integrated into a single percept decreases. Beyond a certain point, there is only a limited chance of integration, which marks the boundary of the temporal binding window. Typically, within this paradigm, behavioural studies use four types of cognitive tasks to provide a measure of the temporal binding window:

(1) *Perceptual-Fusion tasks (PF)*; require participants to respond to whether or not the stimuli were presented as a single event. Responses produce a measurement of the *percentage of perceived fusion*, which is based on the number of responses participants report as a single event for each SOA interval (e.g. Stevenson & Wallace, 2013).

(2) *Temporal Order Judgement tasks (TOJ)*; involve participants judging whether the audio or visual stimulus was displayed first. This task produces a measure of the temporal binding window known as the *percentage of visual leading responses*, which is operationalised by the number of responses at each SOA interval in which an individual reports that the visual stimulus was presented first (e.g. Bald et al., 1942).

(3) *Two-Interval Forced Choice tasks (2IFC);* is a Simultaneity Judgement (SJ) task that requires participants to determine which of two audio-visual stimuli pairs are presented simultaneously, identifying which of the two presentations occurred at the same time. Responses are measured in terms of the *percentage of correct responses* operationalised by correctly identifying which presentation was simultaneous. The other presentation has a varying level of SOAs (e.g. Grant & Seitz, 2002; Marill, 1956).

(4) *Two-Alternative Forced-Choice tasks (2AFC);* are also an SJ task and involve participants simply choosing whether or not the stimuli were presented simultaneously. This leads to the *percentage of perceived synchrony* measure, based on the number of responses that identify the stimuli as simultaneous at each SOA interval (e.g. Levitin et al., 1999).

Experiments within the reaction time paradigm arise from the concepts of facilitation and inhibition (Chapter 2, Section 2.2.2). When multisensory information about a single event is provided within the temporal binding window, reaction times decrease, a phenomenon known as facilitation (Raab, 1962; Miller, 1982). Conversely, if conflicting cues are presented

within the window, reaction times increase, which is referred to as inhibition (Colavita, 1974; Hecht & Reiner, 2009). In order to assess each aspect, two distinct categories of cognitive tasks are typically used:

(1) *Divided attention tasks (redundant target);* involve participants responding as quickly as possible to the first stimuli presented. The task typically involves two or more stimuli from different modalities, presented simultaneously or at a range of offset asynchronies. The reaction times are recorded at each SOA, identifying the point at which facilitation occurs (e.g., Todd, 1912).

(2) *Focused attention tasks:* involve participants only responding to a target stimulus, such as a flash or a sound. The task is typically the same as the divided attention tasks outlined above in all ways except the instruction given. The reaction times are recorded at each SOA, which in this task identifies the point at which inhibition occurs (e.g., Diederich & Colonius, 2004)

Demand characteristics are most prevalent in the 2IFC task (Stevenson & Wallace, 2013). This is due to the structure of the other tasks only eliciting simultaneous/speeded responses when the stimuli fall within a participant's TBW. In comparison, the experimental design of the 2IFC task presents asynchronous/synchronous pairs together, back-to-back. The presentation of the pairs leads to participants being more likely to develop a response strategy (Nidiffer et al., 2018). This is usually overcome by introducing catch trials, which involve including synchronous/synchronous pairs or vice versa (e.g., Nidiffer et al., 2018; Powers, Hillock, & Wallace, 2009). Despite these catch trials reducing the chance of demand characteristics. Adding extra trials to the tasks also increases completion time and potential fatigue.

Fatigue and its effect on attention generally affect the accuracy of the results through participants' response bias (Powers et al., 2009), where individuals are more prone to select one response over another. However, this within-session fatigue can be offset with breaks and controlled through analysis (Hilllock et al., 2011; Hillock-Dunn & Wallace, 2012; Nidiffer et al., 2018; Powers et al., 2009; Stevenson et al., 2013). For example, data can be selected for exclusion/manipulation based on unreliability (whether a response is too fast or repetitive). Whilst it is important to account for within-session fatigue regardless of the cognitive task, different task manipulations can naturally lead to dissimilar levels of fatigue (e.g., Hopstaken et al., 2015). For example, the 2IFC will involve extra catch trials over the same number of intervals as the 2AFC. In some cases, this can double the time to complete the task (e.g. Powers et al., 2016). In comparison, the TOJ task contains one less interval because it does not contain a simultaneous (0ms) trial. Consequently, the TOJ task is the shortest to administer over the same range of SOAs (Stevenson & Wallace, 2013). The shorter task length slightly reduces the effects of fatigue on the participants' responses (Simon et al., 2020).

Although individuals are known to differ in terms of their window width, task averages have also been shown to differ within participants (e.g., Mégevand et al., 2013). The PF task typically displays the largest window width, followed by the 2AFC task. The TOJ and 2IFC tasks both produce narrower windows (e.g., Mégevand et al., 2013; Stevenson & Wallace, 2013), with reaction time tasks demonstrating the narrowest windows (Diederich & Colonius, 2015). These differences in overall window width between the tasks are attributed to the level of processing each task requires. For instance, PF tasks are believed to invoke feedback from higher-order perceptual representations as they require participants to judge the likelihood of same-event derivation rather than more straightforward temporal judgments (Stevenson & Wallace, 2013). This increase in the length of time to form a response, in turn, leads to greater accuracy across a broader range of SOAs and, hence, a wider window width. The wider

windows recorded in PF tasks suggest that the measurement includes more than just the temporal binding window (Stevenson & Wallace, 2013). With the added time for retrieving feedback from the higher-order representations not usually being accounted for within the calculations to estimate window width.

Multiple studies have used two or more of these cognitive tasks to assess individual temporal windows to different stimuli (e.g., Powers et al., 2009; Stevenson & Wallace, 2013; Van Eijk et al., 2008; Vatakis et al., 2008). Stevenson and Wallace (2013) conducted a study using multiple stimuli and task types to investigate the relationship between different variables. The stimulus types consisted of videos of speech, a simple flash beep, and dynamic tool stimulus. In the study, the four estimation and judgement tasks were used. Their analysis showed a wider TBW for speech stimuli than for non-speech stimuli. Moreover, the analysis indicated that although each task might relate to a different cognitive operation, they share a common underlying process which correlates strongly with the whole temporal binding window. Considering both Stevenson and Wallace's (2013) analysis and the issues highlighted in the section above, it was decided to use at least two different types of tasks when investigating the temporal window to compensate for differences between the measures.

In order to decide which combination of tasks measures the TBW the most effectively, it is worthwhile to consider differences between the tasks previously discussed. Despite the similarities between the 2AFC task and the PF task, it is argued that the 2AFC task is a more accurate measure of an individual’s TBW (De Niear, 2016) because it does not require feedback from higher-order representations to form a response (Stevenson & Wallace, 2013). In addition, the aspects of fatigue and inattention should be accounted for when deciding which task to employ. This is especially true when using multiple tasks, which extends the participation time. The 2IFC task involves the most trials over the same range of SOAs. The increase in trials is due to the need to include catch trials, as highlighted above (e.g., Nidiffer

et al., 2018). TOJ tasks involve fewer trials because they do not have a synchronous condition (0ms SOA). As such, the TOJ task is preferred over the 2IFC task, due to the additional length of time the 2IFC would take. Research from Van Eijk et al. (2008) and Stevenson and Wallace (2013) suggest flash beep stimuli provide the smallest measure of the temporal window as they are less complex than videos of speech or dynamic objects. These stimuli are common within experiments using the Redundant Target (RT) tasks. A RT task also provides a comparative measure of the TBW from a different paradigm. These three tasks (2AFC, TOJ and RT) are distinct enough from each other to provide a series of estimations for an individual's TBW width. At the same time, they share an underlying cognitive process that relates to the TBW as a whole (Basharat et al., 2018; Colonius & Diederich, 2018). Each task was therefore included in the first study.

3.1.3 *Current Study*

In this first study, participants' sense of presence within an interactive VR environment was measured through two questionnaires (*SPES*: Hartmann et al., 2016; *IPQ*: Schubert et al., 2001). These two questionnaires were chosen to test the convergent validity of the SPES alongside providing a separate measure of presence, which also included realism. The participant's temporal binding window size was calculated based on the responses to the three separate tasks outlined above (2AFC, TOJ and RT). The relationship between the size of the participant's TBW and the self-reported presence scores was of key interest. Information was also collected relating to potential confounding variables which have been previously demonstrated to relate to presence and multisensory integration: age (e.g., Dieuleveult et al., 2017; Dilanchian et al., 2021; Dresel et al., 2023; Lorenz et al., 2023), gender (e.g., Felnhofer et al., 2012; Traunmüller & Öhrström, 2007), gaming (e.g., De Leo et al., 2014; Di Luzio et al., 2021; Donohue et al., 2010; Pöhlmann et al., 2022), and VR experience (e.g., De Leo et al., 2014; Kim et al., 2021), in order to identify any potential additional individual differences.

As the SPES is regarded as a shortened and fine-tuned version of the MEC-SPQ (Vorderer et al., 2004) with a focus on spatial presence, it is hypothesised that the SPES will significantly correlate with both the IPQ total score and the IPQ self-location subscale.

Since this study is the first to investigate the relationship between individuals' temporal binding window and feelings of presence, the direction of this relationship and the differences between the TBW measures are unknown. The hypotheses of this exploratory study are therefore non-directional. However, as the ability to maintain consistent sensory information is considered necessary for presence to arise (Wirth et al., 2007), it is hypothesised that individuals' TBWs will be a significant predictor of their sense of presence.

Additionally, there are likely confounding variables that could correlate with either participants' levels of presence or the size of their TBW. Among these variables, it is anticipated that prior gaming and virtual reality experience will show the strongest relationship, as previous research indicates that experience has both a long-term impact on individuals' feelings of presence (De Leo et al., 2014) and their ability to integrate multisensory information (Di Luzio et al., 2021; Donohue et al., 2010).

3.2 Methods

3.2.1 *Participants*

Twenty-nine participants (20 Males, 9 Females) aged 18 – 32 (M = 21.5, SD = 3.03), took part in this study. Participants were all students or staff at English universities, recruited through opportunity sampling (e.g., physical/email advertisement). Participants were either entered into a prize draw for monetary gift cards or received course credit for participation. All participants had normal hearing, normal or corrected-to-normal vision, and were fluent in English.

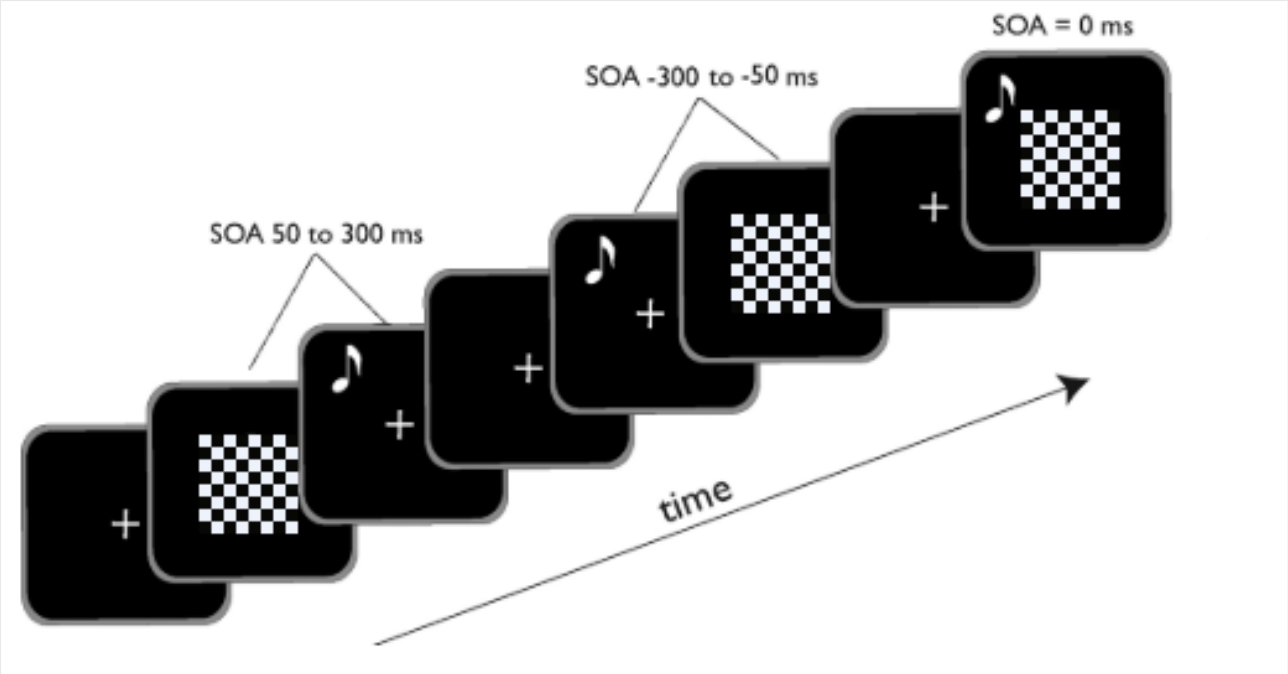
3.2.2 *Materials*

3.2.2.1 **Temporal Binding Window Tasks.**

Participants' temporal binding windows were calculated from responses to three flash-beep tasks, all coded in Psychopy2 (Peirce, 2007). Each task used the same stimuli but differed in design and response criteria. The visual stimuli consisted of a 5x5 (cm) black and white checkerboard, which appeared in front of a fixation cross on a black background (see Figure 1). The checkerboard was presented approximately 60cm in front of the participants on a Lenovo Legion Y25f-10 monitor, which had a 62.23cm screen size and a 144Hz refresh rate. The auditory stimuli were a 440HZ pure tone with a 33ms duration presented binaurally via Sony MDR-ZX110 headphones.

Figure 1

Visual representation of a temporal binding window task trial sequence



2AFC Simultaneity Judgement task: Participants were asked to judge whether an auditory bleep was presented at the same time as a visual checkerboard flash. The stimuli were offset in intervals of 50ms, ranging from 300ms audio leading (bleep presented before the

checkerboard) to 300ms visually leading (checkerboard presented before the bleep) (Figure 1). Participants responded by pressing the ‘Z’ key on a keyboard if they believed the stimuli were presented together or the ‘M’ key if they thought they were not. Each stimulus offset asynchrony was repeated randomly 12 times, leading to 156 trials over the 13 intervals. Participants completed these blocks twice on either side of a two-minute break to improve accuracy.

Redundant Target (RT) task: This task followed the same structure as the simultaneity judgement task detailed above. However, participants were asked to respond as quickly as possible to whichever stimulus was presented first by pressing the spacebar. Much like the simultaneity judgement task, a total block of 156 trials was repeated twice on either side of a two-minute break.

Temporal Order Judgement (TOJ) task: This task also used the same stimuli as the other temporal binding window tasks. However, participants were asked to respond with ‘Z’ if they thought the auditory stimulus was presented first or 'M' if they believed the visual was presented first. As the task required participants to judge whether auditory stimuli or visual stimuli were presented first, there was no 0ms SOA condition. This led to 144 trials over 12 intervals. Like the other tasks, the block of 144 items was repeated twice, separated by a two-minute break.

3.2.2.2 **VR Hardware and Simulation.**

The equipment used to elicit a sense of presence in this study was the commercial version 1 (CV1) of the Oculus Rift (Figure 2), which was the most current version of the headset at the time of the study. The CV1 is a head-mounted display (HMD) with a resolution of 1200 x 1080 pixels per eye, a 110˚ field of view, and a 90Hz refresh rate. The HMD provided both head and position tracking through infrared LEDs. The HMD was connected to an Alienware X51 desktop with an Intel i5-6400 processor and an 8GB dual channel DDR4

2133MHZ and NVIDIA GeForce GTX 970 with 4GB GDDR5: VR-Ready graphics card. Participants could control movement and interactions using the Oculus Rift handheld controllers in the virtual environment.

Figure 2

Oculus Rift headset, handheld controllers, and tracking cameras used in study 1

The virtual reality environment (VRE) was Minecraft© VR version 1.7, a sandbox video game released by Mojang. Minecraft VR allows participants to interact with and traverse through a block-based virtual world. All participants explored an identical environment, using the world generator seed ‘LISTRISK’ (see Figure 3). The world placed participants between two villages in different biomes, one desert and the other forest. The participants were instructed not to climb up high places or go underwater. These warnings were provided to help individuals avoid feelings of vertigo or claustrophobia, and no other restrictions were communicated.

Figure 3

Screenshot of a virtual reality environment generated using seed code 'LISTRISK' in Minecraft VR version 1.7

3.2.2.3 **Questionnaires.**

Two self-report questionnaires were used to assess individuals' reported feelings of presence within the virtual environment.

Spatial Presence Experience Scale (SPES): An eight-item questionnaire adapted from a larger presence questionnaire (MEC-SPQ; Vorderer et al., 2004) by Hartmann et al. (2016). The questionnaire consists of four items relating to self-location (SL) and four items relating to perceived possible actions (PPA). Self-location refers to participants' feelings of 'being there' within the environment. In comparison, perceived possible actions refer to individuals' ability to manipulate and interact with the environment. Each item is rated on a five-point scale, where one is 'I do not agree at all', and five is 'I fully agree'. In the few studies which have analysed the questionnaire's Cronbach alpha, both scales have reported good reliability (α = >0.8) (Barreda-Ángeles et al., 2020; Eckstein et al., 2018; Weigelt & Wiemeyer, 2012). In addition, the scales have shown good construct validity (Hartmann et al., 2016).

Igroup Presence Questionnaire (IPQ): A fourteen-item questionnaire which was originally developed by Schubert et al. (2001). Five questions relate to spatial presence, four relate to involvement, and three relate to realness. A final question, 'I had a sense of being there…' which was taken from the Slater, Usoh and Steed (SUS) presence questionnaire (Slater

et al., 1994), forms a general scale. Each item is rated on a 7-point scale ranging from -3 to +3. Three of the involvement items are negatively scored alongside one of the spatial presence items. The questionnaire was chosen due to its similarity to the SPES. The spatial presence questions relate to both perceived location and action. The other two scales provide additional information about awareness and realism. Unlike the SPES, the validity and reliability of the IPQ has been confirmed by studies in multiple languages alongside its English version (German; (Schubert & Friedmann, 2001), Persian; (Panahi-Shahri et al., 2009), Portuguese; (Vasconcelos-Raposo et al., 2016), Italian; (Iachini et al., 2019), Dutch; (Krijn et al., 2004), and Japanese; (Hyun et al., 2010).

Previous Gaming Experience: A measure of previous gaming experience was also administered in addition to the presence questionnaires. This measure was used to control for the potentially heightened multisensory temporal processing abilities of participants who play more video games (Donohue et al., 2010). The measure was taken from Lemmens et al. (2011a, 2011b) 'time spent on games' questionnaire and adapted to include previous virtual reality experience. The measure, therefore, involved four items. The first and second items were adapted to include mobile phones, which asked how many days per week the participant ‘gamed’ and how many hours they ‘gamed’ on an average day, respectively. The two answers were multiplied to provide a total weekly gaming score. The third and fourth items relate to past virtual reality experiences, which ask how often a participant has used VR equipment and the average time per session. Again, the two answers were multiplied to produce a total VR experience score. These two scores of previous experience were used as control variables. A copy of the questionnaire can be found in the appendices (Appendix 1).

3.2.3 *Procedure*

Each participant was assigned to a counterbalanced group (RT First, TOJ First, SJ First) and completed the study in a single session. Participants were seated at a computer and asked to complete a consent form and answer the demographic questions (e.g., date of birth, gender, participant number). Based on their assigned group, participants completed each temporal binding window task in one of three orders (RT-TOJ-SJ, TOJ-SJ-RT, SJ-RT-TOJ). The researcher left the room whilst the participants were completing the tasks.

The virtual reality simulation followed the temporal binding window tasks. Participants adorned the HMD and adjusted the straps and lenses until they were in an optimum position. Participants were then given up to ten minutes as a familiarity session. This familiarity session was timed to be used as a measure of gaming experience, and participants were given strict instructions to notify the researcher when they felt familiar with the simulation and the controls. The Oculus handheld controllers operated the virtual simulation. Once participants notified the researcher that they felt familiar, or ten minutes was reached, participants had up to a ten-minute break before re-entering the VR.

The participants re-entered the same virtual environment they were in during the familiarity period. During the familiarity session, the researcher was permitted to talk to the participants to guide them with the controls and answer any queries. However, once the main session had begun, the researcher stuck to simple responses only when necessary, unless participants required help. Participants were informed that once they re-entered the VR, they had up to 30 minutes to explore the virtual environment, but if they wished to remove the HMD at any time, they were permitted to do so. The length of time the participants were within the VR environment was recorded. The virtual reality simulation was followed by the three paper questionnaires: SPES, IPQ and the gaming experience measure.

3.2.4 *Ethics*

The research followed the BPS guidelines for ethical practice, seeking written consent from participants and informing participants of their right to withdraw from the study up to three weeks after data collection had finished. Participant's safety was considered a priority, and multiple steps were taken to keep them from harm. Individuals had to pass certain safety checks to participate in the study, such as not being intoxicated or pregnant (Appendix 2). In order to minimise the simulator sickness/discomfort from the experience, instructions were given on best practices whilst using the VR equipment. Participants were advised to avoid climbing heights or going underwater within the virtual simulation, as both situations might cause discomfort. Whilst in the simulation, participants were asked to remain seated, which served to both reduce discomfort and risk of injury. In addition, the researcher remained present for the entire duration of the participants' time in the virtual environment. During the participant's time in the virtual experience, the researcher was vigilant in checking that they were ok. If participants began to show signs of simulator sickness, such as increased sweating, heavy breathing, or nausea, the researcher would ask if they were ok to continue, if they would like to take a break, or if they would like to come out of the simulation completely. No participants required intervention. Precautionary measures were also taken to mitigate the effects of simulator sickness, such as making drinking water readily available. Participants were also advised not to operate heavy machinery after using the VR equipment. In addition to the advice given before entering the VR equipment, after the session, participants were given contact information for help if they later suffered from any simulator sickness or if it increased. The York St John University Research Ethics Committee approved the study.

3.3 Calculating the Window Widths

To analyse the relationship between the temporal binding windows and the presence scores, the windows had to be generated from the raw data for each temporal binding task. Although the three tasks all relate to temporal binding, they tap into slightly different processes

(Basharat et al., 2018; Colonius & Diederich, 2018). Furthermore, models developed to analyse the window width from each task rarely consider a computation for that of another task paradigm. The exception to this is the TWIN models' canonical adaptation (Diederich & Colonius, 2015), which provides a method for both TOJ and RT analysis but not for simultaneity judgement. Therefore, three separate models were used to analyse the three tasks. Each model was initially developed to analyse each task specifically. Using three separate models reduces the likelihood that manipulation, such as applying a canonical adaption, would lead to high multicollinearity in the subsequent regression.

3.3.1 *Computing the TBW from the Simultaneity Judgement Task*

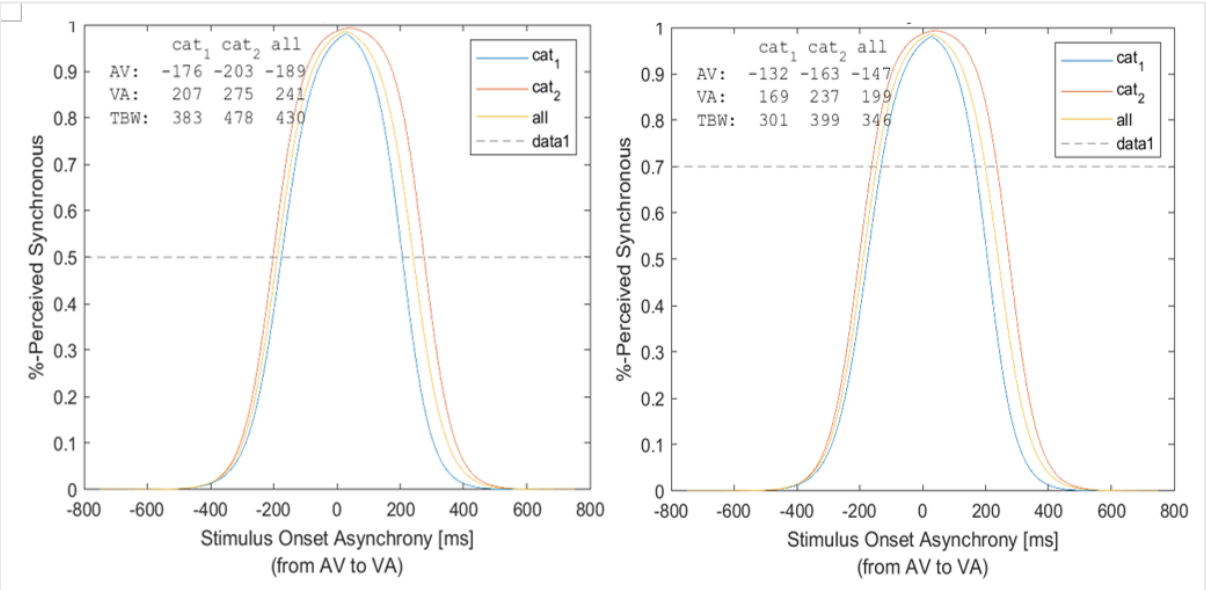
The size of the participant's TBW from the simultaneity judgement task data was calculated in line with Stephenson's and Colleagues' (2012) analysis. First, the raw data was cleaned. Null responses and those with a response time of over 3000ms were removed as neither can be attributed to occurring as a perceptually bound response (Pöppel, 2009). The proportion of responses regarded as simultaneous were then calculated for each SOA [Total Perceived Simultaneous/Total Responses], producing 13 results for each of the two blocks. A MATLAB script was used to plot two sigmoid functions needed to calculate the window size.

The script was adapted from Notter, Michael, and Murray (2017) and can be found in the appendices (Appendix 3). This script differed from Stevenson's (Stevenson et al., 2012) in how it computed the two sigmoid functions. Firstly, Notter et al.'s (2017) script split the data into two equally large parts to create audio-leading and visually-leading conditions. Once the sigmoid curves were computed for both groups, Stephenson resized the curves to reach 100%. In contrast, when using Notter et al.'s (2017) script, the sigmoid functions were not resized, thus avoiding needless data manipulation. The two sigmoid functions were then recombined to form a temporal binding curve.

Two heights of interest were selected in line with a proposal from Stevenson et al. (2012) to investigate and report the differences between both commonly used cut-offs (50% and 70% simultaneity). The width of the participant's temporal binding window is considered to be the point where this height of interest intersects both sides of the temporal binding curve (e.g., Figure 4). Where there is no intersection, the window width was either considered a maximum of 1500 or regarded as a 0, depending on whether the sigmoid curves were entirely above or below the height. These results were then later removed from further analysis. Descriptive statistics for the simultaneity window widths are reported below (Table 2).

Figure 4

Example of a participant's simultaneity judgement temporal binding window at the 50% and 70% cut-offs



3.3.2 *Computing the TBW from the RT Task*

Similar to the simultaneity judgement task, the raw data from the redundant target (RT) task was cleaned. In line with the literature, extreme responses (≤30ms and ≥3000ms) were

removed (Pöppel, 1997, 2009). Once these responses were removed, participants' data was checked for any remaining outliers, which could skew the mean reaction time for each SOA. Any responses over or under three standard deviations from the mean were removed.

The cleaned data was then entered into MATLAB. The scripts used to calculate the windows from the RT task were adapted from Diederich and Colonius's (2015) twin model (see Appendix 3). The scripts use an optimisation function to solve the objective function at the heart of the twin model:

$$\sum_{\tau}\left(\frac{mean\left[RT_{VA,\tau}\right]-E\left[\widehat{RT_{VA,\tau}}\right]}{SE\left(mean\left[RT_{VA,\tau}\right]\right)}\right)^{2}$$

The calculation is identical to the one proposed by Diederich and Colonius (2015), where mean[$RT_{VA,\tau}$] relates to the mean reaction times for each SOA (τ) in the RT task and $E[\widehat{RT_{VA,\tau}}]$ are the predicted values for each SOA. Therefore, the objective function calculates the deviation between measured and estimated reaction time means, returning the fit with the smallest deviation. The predicted values arise from minimising a set of parameters which are passed through the twin models calculations (see Chapter 2, Section 2.2.4). The five parameters are: visual processing time (λ_V); auditory processing time (λ_A); the common processing stage (μ); the TBW for the RT task (ω_{RT}); and multisensory temporal gain (Δ). The main difference between the original author's minimisation and the one conducted here was the choice of optimisation solver within MATLAB. Diederich and Colonius (2015) used *fminsearch* as the solver in the minimisation process. While this choice of optimisation solver adequately mapped the TWIN script to Mégevand and colleagues' (2013) data in their paper, it is not without its problems. Firstly, because *fminsearch* is an unbounded optimisation solver, the parameters which theoretically could be returned are anywhere between ±Infinity (Coleman et al., 1999). Whilst it is acknowledged that the calculations within the twin model itself would not allow an

extremely large value for any of these parameters, there is still the possibility of the solver returning impossible values (such as negative processing times) (Kandil et al., 2014). Secondly, because of the flexibility of *fminsearch,* the results returned may not be as accurate as those of other optimisation solvers. Therefore, whilst all other aspects of minimising the objective function remained the same, the optimisation solver used here was *fmincon.* This problem solver provides more accurate solutions and allows boundaries to be set on the parameters. The boundaries for each parameter in the *fmincon* optimisation are reported in Table 1 below.

Table 1
Boundaries for fmincon Optimisation of the Twin Model

	Visual Processing Time (λ_V)	Auditory Processing Time (λ_A)	Common Processing Stage (μ)	RT TBW (ω_{RT})	Multisensory Temporal Gain (Δ)
Lower Bounds	20	20	0	0	0
Upper Bounds	10000	10000	10000	10000	10000

The scripts return the minimised parameters for each participant and a graph plotting the twin model's predicted values over the measured mean reaction times for each SOA (e.g., Figure 5). The estimated window widths from the RT task (ω_{RT}) were checked for extreme values in SPSS; only one participant's TBW was exceptionally large (≥3000ms) and, therefore, removed from any further analysis. Descriptive statistics for the RT windows are reported below (Table 2).

Figure 5

Example of a participant's reaction time data fitted to the twin model

3.3.3 *Computing the TBW from the TOJ task*

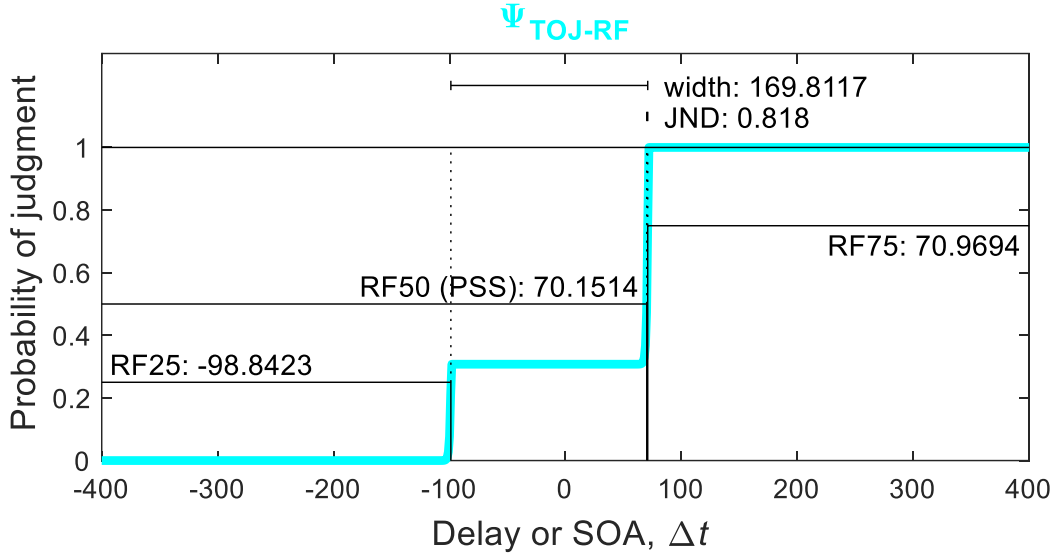
As the temporal order judgement task is more closely related to the accuracy of perception than the other two tasks, data was not removed based on reaction time to stimuli. Instead, participants' responses were reviewed, and any identified to be using a response strategy were removed from subsequent analysis (e.g., responses such as ZZZMMM or ZMZMZM). MATLAB scripts generated by (Alcalá-Quintana & García-Pérez, 2013) were used to avoid fitting arbitrary Gaussian or logistic functions to TOJ data because data from TOJ tasks typically plateau midway along the range of SOAs. As Alcalá-Quintana and García-Pérez (2013) explain, any irregularities average out when comparing group means. However, gaussian or logistic functions cannot do the same for individual data, which was the focus of

this study. The scripts based on Alcalá-Quintana and García-Pérez's independent-channels model can be found in the appendices (Appendix 4).

The model requires a set amount of input parameters alongside data from each individual. The set-up was almost identical to Alcalá-Quintana and García-Pérez (2013). Inputs include boundaries for responses to which stimuli were presented first (LamBounds), alongside boundaries for processing delays (TauBounds) and resolution of small differences (DeltaBounds). The values for each variable are as follows; LamBounds = [1200 1/3]; TauBounds = [−Inf Inf]; DeltaBounds = [0 Inf]; LamTStart = [1/70 1/10]; LamRStart = [1/70 1/10]; TauStart = [−70 70]; DeltaStart = [20 150]; ErrStart = [0.05]; BiasStart = [0.5]; Model = 1; Plot = true; Disp = true. These MATLAB scripts generate a graph of participants' responses (e.g., Figure 6); unlike the typical Gaussian fits, it is possible to identify the window in which the responses plateau and, therefore, make assumptions of the window width directly from individuals' data.

Figure 6

Example of a participant's temporal order judgement temporal binding window



Each individual's estimated window width was input into SPSS for further analysis. Data did not need to be removed at this stage, as all data cleaning checks outlined above had been conducted within the scripts. Descriptive statistics for the temporal order judgement windows and all the other temporal binding window widths are included in the table below (Table 2). Based on the descriptive statistics for the 70% simultaneity (0.7) being closer to the other two windows' width than the 50% simultaneity (0.5), this cut-off level was chosen for further analyses.

Table 2

Descriptive Statistics for the TBW Tasks in Study 1

Task Type (N)	Mean (ms)	Min (ms)	Max (ms)
0.5 Simultaneity (27)	600.33	266.00	1398.00
0.7 Simultaneity (27)	340.56	42.00	728.00
Redundant Target (26)	26.39	7.09	181.19
Temporal Order (29)	161.53	4.16	394.73

Note: 0.5 and 0.7 refer to the 50% and 70% height of interest cut-offs in the window width calculation of the simultaneity judgement window.

3.4 Results

The results section is structured as follows: the first section reports the descriptive statistics of all the questionnaires used in this study, and the second section addresses the first hypothesis by assessing the convergent validity of the SPES and the IPQ. The final section focuses on the relationships between the temporal binding windows and presence measures. Specifically, the third hypothesis is explored through correlational analyses to check for confounding variables before the main hypothesis is tested through multiple regression analyses examining the relationship between the temporal binding windows and each of the presence questionnaires.

3.4.1 *Descriptive Statistics for the Questionnaires*

The raw questionnaire data were converted to total scores in the following ways. The two SPES subscales were calculated by adding up the total of the first four and last four items to form the self-location (SL) and the perceived possible actions (PPA) scales, respectively. The scores for these subscales ranged from 4 to a maximum of 20. The SPES questionnaire total was then calculated by adding up the scores of the SL and PPA subscales to form a score between 8 and 40.

The IPQ questionnaire consists of 4 subscales and involves reverse scoring. The spatial presence subscale was calculated by adding the scores of items 3, 6, 9, and 10 and the inverse of item 13. Each item can be scored -3 to 3, generating a total range of -15 to 15. The Involvement subscale is calculated by adding items 1, 7, 14 and the inverse of item 11, which forms a total ranging from -12 to 12. Item 8 formed its own subscale, 'general presence', ranging from -3 to 3. Realism was the final subscale, calculated from the response of items 4, 5, and 12 and the inverse of item 2, again with a range of scores from -12 to 12. The IPQ total was calculated by adding all subscales, generating a score between -42 and 42.

Participants' previous gaming experience was calculated by multiplying the average daily usage (Hours) with the average amount of days spent playing a week, providing a continuous score. Previous VR experience was calculated comparably by multiplying the average time spent within virtual reality per session (Minutes) by the number of sessions. The descriptive statistics for all the questionnaires and subscales are displayed below (Table 3).

Table 3
Descriptive Statistics for Experience and Presence Questionnaires in Study 1

Questionnaire (N)	Mean	Standard Deviation	Minimum	Maximum
VR Experience (23)	20.109	31.537	0	100
Gaming Experience (29)	16.231	18.374	0	63
SPES SL (28)	13.214	3.370	6	19
SPES PPA (29)	16.138	2.850	11	20
SPES total (28)	29.536	5.351	20	39
IPQ SP (27)	7.074	4.103	-3	14
IPQ Realism (29)	-3.552	3.269	-11	1

IPQ Involvement (29)	1.690	5.211	-10	10
IPQ General (27)	1.778	.892	0	3
IPQ Total (27)	8.185	8.513	-10	23

Note: VR= Virtual Reality, SPES = Spatial Presence Experience Scale, SL = Self-Location, PPA = Perceived Possible Actions, IPQ = Igroup Presence Questionnaire, SP = Spatial Presence.

3.4.2 *Checking the Validity of the Scales*

Bivariate Pearson's correlations were conducted for individual items of both SPES and IPQ scales to identify if any item did not correlate with the overall measure. All of the SPES items correlated with the SPES Total and their subscales. Item 12 of the IPQ ("the virtual world seemed more realistic than the real world") was the only item not to correlate with the total score. However, it did correlate with its subscale (Realism). The correlation between each item, subscale and total questionnaire score is reported below (Table 4). Correlations were also conducted between total scores and subscales of each presence questionnaire to assess if the SPES had convergent validity with the IPQ. The SPES Total significantly correlated with: IPQ Total ($r(27)=.595$, $p<.01$), IPQ Spatial Presence ($r(27)=.555$, $p<.01$), IPQ Involvement ($r(27)=.519$, $p<.01$), and IPQ General ($r(27)=.588$, $p<.01$) , see table below (Table 4).

Table 4

Inter-Item Correlation Between SPES and IPQ Subscales in Study 1

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.
1. SPES Total															
2. SPES SL	**.924****														
3. SPES PPA	**.865****	**.607****													
4. SPES Q1	**.854****	**.881****	**.619****												
5. SPES Q2	**.805****	**.685****	**.772****	**.661****											
6. SPES Q3	**.617****	**.803****	0.230	**.560****	0.247										
7. SPES Q4	**.739****	**.872****	**.392***	**.690****	**.365***	**.738****									
8. SPES Q5	**.732****	**.519****	**.837****	**.402***	**.754****	0.214	**.331***								
9. SPES Q6	**.801****	**.635****	**.830****	**.568****	**.769****	0.303	**.445****	**.753****							
10. SPES Q7	**.605****	**.492****	**.610****	**.585****	**.314***	**.370***	**.358***	0.211	**.498****						
11. SPES Q8	**.471****	0.211	**.701****	**.340***	**.457****	-0.135	0.080	**.480****	0.281	0.233					
12. IPQ Total	**.595****	**.687****	**.336***	**.622****	**.468****	**.536****	**.604****	**.444****	**.374***	0.195	0.001				
13. IPQ SP	**.555****	**.631****	**.326***	**.646****	**.500****	**.378***	**.546****	**.403***	**.369***	0.178	0.036	**.892****			
14. IPQ Realism	0.269	**.338***	0.116	0.305	0.253	0.296	0.245	0.235	0.090	0.010	-0.005	**.689****	**.476****		
15. IPQ Involvement	**.519****	**.605****	0.284	**.438****	0.305	**.612****	**.577****	**.388***	**.365***	0.198	-0.076	**.820****	**.569****	**.398***	
16. IPQ General	**.588****	**.631****	**.393***	**.674****	**.490****	**.365***	**.547****	**.404***	**.331***	0.287	0.158	**.797****	**.797****	**.511****	**.485****
17. IPQ Q1	0.142	0.310	-0.112	0.184	-0.046	**.536****	0.291	0.051	0.113	0.021	**-.464****	**.581****	**.370***	**.368***	**.710****
18. IPQ Q2 Transformed	-0.043	0.034	-0.133	0.050	-0.061	0.097	0.021	-0.092	-0.179	0.004	-0.126	**.422***	0.235	**.678****	0.235
19. IPQ Q3	0.249	0.306	0.116	0.223	0.204	0.234	**.320***	0.310	0.187	-0.163	-0.018	**.548****	**.641****	0.154	**.457****
20. IPQ Q4	0.255	0.296	0.141	0.236	0.220	0.265	0.235	0.144	0.157	0.109	0.022	**.626****	**.544****	**.680****	**.415***
21. IPQ Q5	**.325***	0.308	0.270	**.335***	**.366***	0.216	0.107	**.379***	0.233	0.086	0.089	**.388***	0.294	**.674****	0.099
22. IPQ Q6	**.474****	**.476****	**.361***	**.491****	**.519****	0.155	**.409***	**.403***	**.315***	0.174	0.176	**.716****	**.822****	**.505****	**.346***
23. IPQ Q7	**.448****	**.560****	0.197	.354*	0.246	**.588****	**.586****	**.404***	0.270	0.054	-0.138	**.575****	**.390***	0.223	**.747****
24. IPQ Q8	**.588****	**.631****	**.393***	**.674****	**.490****	**.365***	**.547****	**.404***	**.331***	0.287	0.158	**.797****	**.797****	**.511****	**.485****
25. IPQ Q9	**.635****	**.733****	**.358***	**.812****	**.522****	**.518****	**.562****	**.361***	**.461****	**.334***	-0.035	**.785****	**.749****	**.434****	**.626****
26. IPQ Q10	**.641****	**.677****	**.444****	**.733****	**.571****	**.402***	**.528****	**.464****	**.527****	0.290	0.075	**.771****	**.841****	**.486****	**.445****
27. IPQ Q11 Transformed	0.204	0.246	0.102	0.110	0.024	0.265	**.364***	0.102	0.083	0.092	0.032	**.509****	0.281	0.244	**.722****
28. IPQ Q12	0.044	0.127	-0.074	0.064	0.012	0.102	0.216	0.084	-0.076	-0.245	-0.027	0.165	-0.023	**.357***	0.191
29. IPQ Q13 Transformed	0.219	0.299	0.062	0.311	0.184	0.184	0.301	0.083	0.042	0.104	-0.034	**.583****	**.723****	0.276	**.313***
30. IPQ Q14	**.691****	**.581****	**.672****	**.617****	**.701****	0.283	**.329***	**.557****	**.603****	**.432****	**.424***	**.610****	**.564****	0.274	**.558****

Note: SL = Self-Location; PPA = Perceived Possible Actions; SP = Spatial Presence; *. Correlation is significant at the 0.05 level (1-tailed); **. Correlation is significant at the 0.01 level (1-tailed).

Table 4: Continued

	16.	17.	18.	19.	20.	21.	22.	23.	24.	25.	26.	27.	28.	29.	30.
1. SPES Total															
2. SPES SL															
3. SPES PPA															
4. SPES Q1															
5. SPES Q2															
6. SPES Q3															
7. SPES Q4															
8. SPES Q5															
9. SPES Q6															
10. SPES Q7															
11. SPES Q8															
12. IPQ Total															
13. IPQ SP															
14. IPQ Realism															
15. IPQ Involvement															
16. IPQ General															
17. IPQ Q1	0.276														
18. IPQ Q2 Transformed	**.425***	0.268													
19. IPQ Q3	**.312***	**.461****	-0.059												
20. IPQ Q4	**.338***	0.290	**.420***	0.291											
21. IPQ Q5	**.388***	0.126	0.225	0.063	0.186										
22. IPQ Q6	**.685****	0.110	0.305	**.519****	**.561****	0.246									
23. IPQ Q7	**.339***	**.497****	0.014	0.248	0.025	0.138	0.131								
24. IPQ Q8	**1.000****	0.276	**.425***	**.312***	**.338***	**.388***	**.685****	**.339***							
25. IPQ Q9	**.743****	**.502****	0.208	**.345***	0.264	**.385***	**.466****	**.558****	**.743****						
26. IPQ Q10	**.785****	0.269	0.207	**.332***	**.466****	**.429***	**.669****	**.368***	**.785****	**.816****					
27. IPQ Q11 Transformed	0.231	**.355***	0.265	**.348***	**.510****	-0.152	0.236	0.260	0.231	0.149	0.058				
28. IPQ Q12	0.056	0.211	0.031	0.010	-0.141	0.169	0.025	**.423***	0.056	0.169	-0.026	-0.105			
29. IPQ Q13 Transformed	**.542****	0.105	0.225	0.224	**.450****	0.086	**.480****	0.218	**.542****	**.347***	**.485****	0.217	-0.191		
30. IPQ Q14	**.528****	0.062	0.097	0.189	**.328***	0.193	**.528****	0.249	**.528****	**.551****	**.590****	**.329***	-0.026	**.341***	

Note: SL = Self-Location; PPA = Perceived Possible Actions; SP = Spatial Presence; *. Correlation is significant at the 0.05 level (1-tailed); **. Correlation is significant at the 0.01 level (1-tailed).

3.4.3 *Correlations Between Presence Scores, the TBW Measures, and Control Variables*

Nonparametric correlations (Spearman's rank order) were conducted between the control variables (age, gender, condition, gaming experience, VR experience and total gametime within VR (seconds)), both the presence measures and all TBWs, to identify any potential confounding variables. A significant relationship was identified between IPQ Total and the amount of time individuals were in the VR simulation (r(28) = .459 *p*=.014). However, age, gender, condition, total game experience and total VR experience did not correlate with any measures and, therefore, were not included as controls in further analysis (see Table 5).

Table 5
Correlation between Confounding Variables, Presence Questionnaires and TBWs in Study1

	Age	Gender	Condition	Gametime Seconds	Gaming Experience	VR Experience
IPQ Total	-0.327	-0.133	-0.199	**.459***	0.187	0.009
SPES Total	-0.113	0.033	-0.182	0.367	0.266	-0.076
Redundant Target TBW	-0.092	0.289	-0.176	-0.326	-0.126	0.088
Temporal Order Judgement TBW	0.115	-0.062	0.226	-0.087	-0.113	0.223
0.7 Simultaneity TBW	-0.355	-0.118	0.215	-0.096	0.075	-0.041

**. Correlation is significant at the 0.05 level (1-tailed); **. Correlation is significant at the 0.01 level (1-tailed).

3.4.4 *Conducting the Regression*

Two multiple regressions were conducted to analyse the relationship between the TBW and self-reported presence. The first regression used a backward entry method to predict the scores of the SPES based on all the TBW measures (RT, Simultaneity 0.7, TOJ TBW) and the control variable, which was previously identified to correlate with at least one measure (Gametime). A backward entry regression was used due to the exploratory nature of the present study (Field, 2013). A significant regression was found (F(1,24)= 5.383, p=.029) with a R^2 of .183. Participants predicted SPES score is equal to 31.232 - .073 (RT width), where RT width

is measured in milliseconds. Participants' SPES score decreased by .073 per millisecond increase in Redundant Target (RT) window width. Only RT width was a significant predictor of participants' SPES score.

The second regression also used a backwards entry method alongside the same predictors as the first. However, the IPQ total score was used as the outcome measure. A significant regression was also found ($F(2,23 = 7.633, p=.003$), with an R^2 of .399. Participants predicted IPQ score is equal to 1.735 - .145 (RT width) + 0.006 (Gametime), where RT width is measured in milliseconds and gametime is measured in seconds. Participants' IPQ score decreased by 0.145 per millisecond the RT window width increased. However, the score increased by 0.006 per each second the participants were in the VR simulation. Only the RT window width was a significant predictor of participants' IPQ score.

3.4.5 *Results Summary*

In summary, raw participant data was cleaned and temporal binding windows were created for each task. Each questionnaire was checked for convergent validity through bivariate correlations; all items correlated with either the subscale or the total score. In addition, the SPES significantly correlated with the IPQ. Nonparametric correlations only identified gametime as a confounding variable; therefore, this variable was included as a control measure in the following regressions. Two regressions were conducted, one for each presence scale. The first regression identified a negative relationship between the SPES score and the RT window width. The second regression also identified a negative relationship between the RT window width and the IPQ scores.

3.5 Discussion

This exploratory study addressed three hypotheses by exploring the relationship between individual's temporal binding windows and their self-reported feelings of presence on

two questionnaires (SPES; Hartmann et al., 2016; *IPQ*; Schubert et al., 2001). Firstly, it was hypothesised that the size of an individual's TBW would predict how present they felt within the virtual environment. This hypothesis aligns with the idea that individuals need to maintain a coherent stream of multisensory information to develop a sense of presence (Wirth et al., 2007). The opportunity was also taken to explore whether the short and refined presence questionnaire (Hartmann et al., 2016), which underpins the process model of presence (Wirth et al., 2007) could be validated. It was hypothesised that the SPES would have high convergent validity with the longer-standing IPQ. This prediction was based on the PPA scale of the SPES being largely derived from the embodied presence theory (Schubert et al., 1999b), which the IPQ was developed to assess (Schubert et al., 2001). Another opportunity was taken to investigate if there were any potential variables which might confound the relationship between an individual's TBW and their sense of presence. Paying particular attention to gaming and VR past experiences as these have previously been highlighted to have a relationship with both measures (e.g., De Leo et al., 2014; Di Luzio et al., 2021).

3.5.1 *The Relationship between TBW and Presence*

The main hypothesis predicted that the width of individuals' TBW would be a significant predictor of their self-reported presence scores. To develop a sense of presence in a virtual environment, individuals need to maintain a coherent stream of multisensory information (e.g., Wirth et al., 2007). The size of an individual's temporal binding window might affect the maintenance of multisensory information in two ways. Individuals with a narrow temporal binding window do not bind offset information (Stevenson et al., 2012). These individuals can easily identify if sensory information originated from the same or separate events. When multisensory information from the same event is presented coherently, it is maintained as a unified stream. For example, if the audio and visual information of a pickaxe hitting a stone block is presented simultaneously, it will likely be considered the same event. If the

information is offset (the audio is provided before or after the visual information), it is regarded as two separate events. In this example, the two events would be a pickaxe hitting a stone block and a separate sound of a pickaxe hitting something.

The alternative hypothesis is that individuals with a wider window would not perceive the offset information to be of two separate events (Stevenson et al., 2012) if the audio-visual information was meant to represent one event (as in this example) but was delayed. This is the same regardless of whether the delay was due to the processing times of the technology or the individual. Individuals with a wider temporal binding window would have the potential to maintain this as a unified percept.

Consequently, there are two opposing hypotheses concerning how the width of the TBW may influence an individual's sense of presence. The analysis conducted in this study found that an individual's TBW, as measured by the RT task, was a significant predictor of both presence measures. Specifically, the findings showed that individuals with a narrower window of temporal binding in the RT task also self-reported higher levels of presence within the subsequent simulation on both the IPQ and the SPES. These findings support the idea that individual differences in temporal binding play a role in developing presence.

There are at least three possible reasons which might explain the direction of this relationship. The first possible explanation for the negative relationship between presence and an individual's TBW is linked to the first step in the formation of presence. Most models agree that paying attention to the information presented in the virtual environment is an important first step toward developing a sense of presence (e.g., Biocca, 2003; Nunez, 2007; Wirth et al., 2007). This attention can be endogenous (voluntary) or exogenous (involuntary). Within the process model of spatial presence, these are called controlled and automatic attention, respectively (Wirth et al., 2007). Endogenous/controlled attention is considered purposeful

(Theeuwes, 1991), such as searching for an object within an environment (Macaluso, 2010). In contrast, Exogenous/automatic attention is stimulus-driven (Theeuwes, 1991). For example, sensory events within the real or virtual world that capture individuals' sensory processing.

Both exogenous and endogenous attention allocation have been explored in relation to the processing of multisensory information. The relationship between attention and multisensory integration is complex, but the literature agrees on a few fundamental principles (Koelewijn et al., 2010; Talsma et al., 2010). Endogenous attention is considered to choose whether to attend to a multisensory target selectively (Tang et al., 2017). Binding of multisensory information is more likely to occur if the modalities are consciously attended to (e.g., Talsma & Woldorff, 2005). Consciously attending to the stimuli is considered a top-down influence of attention on integration. On the other hand, multisensory integration can have a bottom-up effect on exogenous attention. Where integrated cues increase the allocation of attention toward a multisensory stimulus (Tang et al., 2017).

While individuals are in virtual reality, sensory information is still being presented about the natural world (e.g., gravity, temperature, olfactory cues). To feel present, the virtual environment should capture as many senses as possible (Steuer, 1992; Slater et al., 1994), with audio-visual information being considered the most important (Dinh et al., 1999). As highlighted above, individuals with a wider temporal binding window are more likely to combine offset information into a unified percept (Stevenson et al., 2012). This is also true if the information is presented close enough together but from two distinct events (e.g., McGurk Effect; MacDonald, 2018). If an individual hears an event in the real world at a similar time to a visual event in the virtual world, they might process the sensory information in a unified way. Processing the information as a unified percept may split their attention between the virtual world and cues from the real world.

In contrast, individuals with a narrower window may be better at attending to coherent information presented from one medium. If the information of that medium is presented congruently, it likely captures the individual's attention more quickly (Talsma & Woldorff, 2005). Additionally, these individuals are likely to separate the information presented from two different mediums, concentrating on the information that captures the most attention. It is argued here that individuals with a narrower window potentially have the ability to ignore incongruent/irrelevant sensory cues. More so than individuals with a wider window. This is especially relevant for virtual reality, as there are multiple irrelevant cues to ignore that are incongruent with the audio-visual information provided by the HMD. The ability to ignore these cues might explain the findings reported here, which are that narrower windows produce larger self-reports of presence in this study.

Although attention is a common first step in models of presence, there are at least two subsequent stages that may also help explain the relationship identified here. It is hypothesised that once individuals attend to the environment, they build a mental model of the world they perceive (e.g., Slater et al., 1994; Biocca, 2003). In the process model (Wirth et al., 2007), these mental models are formed from egocentric reference frames (Riecke & Von der Heyde, 2003)—representations of an environment based on sensory cues.

Individuals with narrow windows may create more egocentric reference frames due to processing offset/incongruent information as two separate events (Ehrenfeld & Butz, 2013). In contrast, individuals with a wider window may combine information into a single reference frame, even if the information is offset (Lohmann & Butz, 2017). This difference can impact the sense of presence in two ways.

Firstly, if the VR presents offset audio-visual information, individuals with a narrow window are more likely to notice the discrepancy. This could lead them to develop two separate

reference frames, which may diminish their sense of presence (Ehrenfeld & Butz, 2013). Without a singular reference frame to help build a mental model, they may struggle to feel located within the virtual environment. In contrast, those with a wider window are less likely to notice such offsets, leading them to develop only one cohesive reference frame, thereby enhancing their sense of presence (Lohmann & Butz, 2017). However, the opposite effect can occur when it comes to incongruent information.

The key distinction is that offset information relates to sensory cues from the same event occurring at different times (Temporal Incongruence; Kim & Lee, 2022). In comparison, incongruent information clashes with the expected event (Sensory Conflict; Donohue et al., 2013). For example, consider a virtual environment where an individual moves throughout the virtual world yet remains stationary in the real world, as in this study. The audio-visual representation depicts that they are moving, whereas their proprioception provides sensory cues that they are stationary. Rather than being offset, the information is incongruent. Individuals with a narrow window may be more likely to separate these senses into two representations: one of the virtual world, containing congruent audio-visual information, and one of the natural world, where they are stationary in a lab. If the reference frame of the virtual world is the most salient, they may still develop a sense of presence. In contrast. those with a wider window may bind proprioception cues with audio-visual cues. The binding of these incongruent cues would lead to an inaccurate reference frame. Combined into a single percept, the conflicting sensory information could undermine an individual’s experience of presence, leading to the lower levels of presence reported here.

The following step and final explanation for why narrow windows might lead to a higher reported presence is due to testing these reference frames. Perceptual hypothesis testing is the final step within most models of presence (Wirth et al., 2007). A cognitive process asking, "Am I really located here?". Within the process model, the most salient egocentric reference

frame is accepted as the primary egocentric reference frame (PERF) (Riecke & Von der Heyde, 2003; Wirth et al., 2007). If this is of the virtual environment, a sense of presence is thought to occur. Individuals with narrow windows may create salient representations of the virtual world because they do not bind incongruent or offset information into this representation (Kim & Lee, 2022). However, this does not necessarily mean that the virtual environment would be the most salient. If the reference frame developed from the cues of the natural world (e.g., proprioception) were more salient than the virtual environment, they would feel located within the real world (Wirth et al., 2007). This is where the dominance of audio-visual information becomes important. Providing that the audio-visual information from the virtual world is congruent, an individual with a narrow window is likely to have a clearly developed PERF of the virtual world. They would separate the conflicting cues into a non-primary reference frame. This may be a reason why individuals with a wider window might experience a diminished sense of presence because their PERF potentially includes conflicting information (e.g., Donohue et al., 2013). This is likely due to them not successfully attributing the sensory cues to two separate reference frames. Instead, the combined incongruent cues undermine the mental model of the environment.

It's important to note that if all the information presented in a virtual environment were consistent across all sensory receptors, the width of the temporal windows would likely not matter. In this case, individual differences in reported presence would not be attributed to sensory factors. However, in virtual reality environments (VREs) where individuals navigate through a virtual space without physically moving, like in this study, the absence of consistent sensory cues may explain why narrower temporal windows result in a greater sense of presence.

3.5.2 *SPES as a Valid Measure of Presence*

One of the aims of this exploratory study was to identify whether the SPES (Hartmann et al., 2016) displays convergent validity with the IPQ (Schubert et al., 2001), a presence questionnaire that has existed for a longer period. All scales of the IPQ significantly correlated with the SPES Total (.519-.595), except for the realism scale (.269). These findings demonstrated good convergent validity between the two scales overall.

These findings support two assumptions outlined by Hartmann et al. (2016). Firstly, Hartmann predicted that the self-location dimension would have a positive correlation with the spatial presence subcomponent of the IPQ ($r(27)=.631$, $p<.01$). Secondly that the SPES would have lower correlations with the realism scale due to the process model not considering realism as "an inherent dimension of spatial presence" (Hartmann et al., 2016, p.12). Overall, this supports the continued use of the SPES as a valid measure of spatial presence.

3.5.3 *Influence of Confounding Variables on Presence*

An additional aim of this exploratory study was to identify any factors that might covary with the relationship between TBWs and presence. Whilst it is expected that individuals' temporal binding windows might play a role in the formation of presence, it is unlikely to be the sole reason behind the variance between individuals. Instead, multiple factors are likely to be involved. A variety of influences have been shown to impact both factors in previous research: gender (e.g., Felnhofer et al., 2012; Lorenz et al., 2023), age (e.g., De Dieuleveult et al., 2017; Dresel et al., 2023), gaming experience (e.g., Di Luzio et al., 2021; Pöhlmann et., 2022). This is in addition to elements which have been shown to impact just presence, such as length of time within VR (e.g., Thorp et al., 2022), and those which just affect temporal binding, such as task order (e.g., Kostaki & Vatakis, 2018). However, in this study, only the length of time individuals were within the VR was found to be a covariate of presence

($r(28)=.459$, $p=.014$) and a non-significant predictor of individuals' presence scores ($\beta = .006$) on the IPQ.

The finding that length of VR time is related to presence is consistent with previous research, which has demonstrated that the duration of time individuals are in VR impacts their sense of presence. For example, Thorp and Colleagues (2022) found that individuals' sense of presence increased in increments from 1 minute to 5 minutes within VR. Unlike the study presented here, presence started to reduce after 5 minutes for females. In addition, it took longer for males to report feelings of presence. However, within this study, participants' levels of presence, as measured by the IPQ, increased steadily with the amount of time they spent within the VE. This finding was not replicated with the presence score measured by the SPES, indicating the need to investigate this and other covariates further.

3.5.4 *Limitations*

Due to the exploratory nature of this study, there are some potential limitations to highlight. Primarily, individual's TBW widths were estimated based on a particular model from within each paradigm (Alcalá-Quintana & García-Pérez, 2013; Diederich & Colonius, 2015; Stevenson et al., 2012). Although, this decision was made to avoid high multicollinearity, the use of multiple models could explain why neither the Simultaneity Judgement nor the Temporal Order Judgement windows were a predictor of presence. Each model is thought to encapsulate the slightly different processes of each task (Basharat et al., 2018). However, without a model that provides a canonical adaptation for all three tasks, the lack of relationship could be due to the variation within the cognitive models used (Colonius & Diederich, 2018) instead of the differences between individuals' TBWs on each task.

It should be noted that whilst there are no models which provide an adaptation for all three tasks, both the independent-channels model (Alcalá-Quintana & García-Pérez, 2013) and

the TWIN model (Diederich & Colonius, 2015) can be used to estimate TBWs across two tasks. The independent-channel model provides an estimation for individuals' TBWs based on data from either the SJ or the TOJ tasks. Both sit within the same estimation and judgement paradigm highlighted in Chapter 2 (Section 2.2.3). In comparison, the TWIN model was extended by Diederich and Colonius (2015) to account for data from both the redundant target (RT task) and the estimation and judgement paradigms (TOJ task). In order to negate the variance of different cognitive models' estimations of the TBW, the extended TWIN model will be used going forward (Colonius & Diederich, 2018). Despite the extended TWIN model not allowing for investigation of the binding windows estimated from an SJ task, it does continue to represent the slightly different processes occurring within each paradigm through the TOJ and RT tasks.

A second key limitation relates to the influence of confounding variables. Six potential confounds were measured here, each identified as having a potential impact from past literature: A*ge* (e.g., Lorenz et al., 2023), *Gender* (e.g., Felnhofer et al., 2012), G*aming Experience* (e.g., Di Luzio et al., 2021), *VR Experience* (e.g., Kim et al., 2021) and *Task Order* (e.g., Kostaki & Vatakis, 2018) and *Time Spent in Game* (e.g., Thorp et al., 2022). Only the latter was found to have a significant relationship with presence, as self-reported from the IPQ. Despite attempts to control for a selection of confounding variables, there are likely to be some which still need to be explored. Simulator sickness has demonstrated a complex relationship with presence in the past (Weech et al., 2019). The complexity of this stems from studies reporting positive (e.g., Witmer et al.,1996), negative (e.g., Wilson et al., 1997) and null relationships between the two constructs (e.g., Mania & Chalmers, 2001). While simulator sickness was not measured in this study, at least five participants verbalised experiencing some degree of simulator sickness after completing the study. The degree of simulator sickness might have impacted their self-reported presence scores, although the direction of this relationship is

unclear. In addition, recent research has identified a relationship between temporal order judgements and simulator sickness (Sadiq, 2019; Sadiq & Barnett-Cowan, 2022; Zilka & Bonneh, 2022). The relationship has found individuals with wider TBWs to report higher levels of simulator sickness. Based on this, simulator sickness could be a confounding variable which needs to be explored in future studies.

3.5.5 *Summary and Conclusions*

This exploratory study aimed to further the research on individual differences in feelings of presence by addressing three key questions: Does the size of an individual's TBW predict self-reported presence within a VR HMD? Is the SPES a valid measure of individuals' Spatial Presence? Are there any individual differences that confound a relationship between temporal binding and presence? Three tasks were used to assess the width of an individual's temporal binding window in line with recommendations from the literature (Stevenson et al., 2013). Afterwards, individuals experienced a virtual reality experience in the form of an interactive game (Minecraft). Then, two self-report presence questionnaires (*SPES*, Hartmann et al., 2016; *IPQ*, Schubert et al., 2001) were completed relating to their experience.

The results identified that individuals' TBW, as measured by the redundant target (RT) task, significantly predicted self-reported presence scores. As the size of an individual's TBW window increased, presence decreased. One explanation for this relationship is that individuals with a wider window may combine incongruent/irrelevant stimuli from the real and virtual worlds into a single reference frame (Lohmann & Butz, 2017). As such, the lack of consistent multisensory information (Steuer, 1992) and potentially conflicting sensory cues (Donohue et al., 2013) diminishes the mental model (Wirth et al., 2007) and subsequently presence. Secondly, self-reported presence was not predicted by either TBWs generated from the estimation and judgement tasks. It is argued that this may be due to the variance within model

estimation (Colonius & Diederich, 2018) rather than differences in individual performance between the tasks (Basharat et al., 2018). Adopting a canonical model such as the extended TWIN model (Diederich & Colonius, 2015) and then selecting the tasks it can measure (RT and TOJ) would address this issue.

The results regarding the validity of the SPES supported Hartmann and Colleagues' (2016) predictions. The self-location subscale had a positive relationship with the spatial presence component of the IPQ. In addition, the SPES total did not correlate with the IPQ Realism subcomponent. The lack of correlation is likely due to the process model not considering realism necessary for spatial presence to arise (Hartmann et al., 2016). These findings and significant correlations between the SPES total and the remaining IPQ subscales demonstrate the validity and support the continued use of the SPES.

Finally, although the results only identified a significant relationship between presence as measured on the IPQ and one confounding variable (Gametime), this was not a significant predictor of presence when included within the multiple regression. More so, some individuals who exited the VR reported feelings of simulator sickness. Simulator sickness was not measured within this study as it is known to have a complex relationship with presence (Weech et al., 2019). More recently, simulator sickness has also been shown to have a relationship with individuals' TBWs (e.g., Sadiq, 2019). The following study was developed to measure this potential confound empirically. It will also aim to improve the reliability of these exploratory results. Unlike this exploratory study, the next study aimed to recruit a much larger sample to increase the power of the empirical findings.

Chapter 4: Exploring the Relationship Between TBW, Presence and Simulator Sickness (Study 2)

4.1 Introduction

Individuals code information presented about their immediate surroundings into first-person 'schema-like' representations called egocentric reference frames (Riecke & Von de Heyde, 2003). These egocentric reference frames help individuals orient themselves within and navigate real and virtual worlds (Gramann et al., 2010; Klatzky, 1998). The most salient or prominent egocentric reference frames form the building blocks of mental models which underpin spatial presence (Wirth et al., 2007). Like spatial presence, the ability to develop coherent reference frames is known to differ between individuals (Gramann, 2013). The results from the study reported in Chapter 3 (Section 3.4) indicate that the size of an individual's binding window might partially explain why some individuals are better at developing these reference frames than others. It is hypothesised that individuals with narrower windows are better at developing coherent representations of the environment they are located within. These more coherent representations are formed by creating multiple reference frames when offset/incongruent sensory information is presented (Riecke & Von de Heyde, 2003). This is in contrast to individuals with wider windows, who are hypothesised to combine this offset/incongruent information into a single reference frame. When these reference frames are 'perceptually tested', salient reference frames lead to presence (Wirth et al., 2007). On the other hand, reference frames that lack consistency diminish the sense of presence within a virtual environment (Wirth et al., 2007).

The previous study highlighted that narrow windows can lead to increased presence levels in the presence of potentially conflicting information. For instance, when vestibular cues suggest stillness (sat stationary), while visual cues indicate movement (moving through a

virtual world). This suggests that individuals who separate conflicting information into different egocentric reference frames feel the most present (Wirth et al., 2007). Another factor that may influence individuals' sense of presence is simulator sickness (Grassini et al., 2021; Weech et al., 2019). Although the previous study did not assess simulator sickness, several participants expressed varying degrees of discomfort. Moreover, simulator sickness has a complex relationship with presence, with some studies reporting negative (e.g., Witmer & Singer, 1998), positive (e.g., Ling et al., 2013), or no relationship (e.g., Ryan & Griffin, 2016) between the two concepts. This variability points to the possibility that different factors—such as individual differences, environmental selection, and task demands—may influence the relationship between presence and simulator sickness. Additionally, simulator sickness has been identified to correlate with TBWs (Sadiq & Barnett-Cowan, 2022; Zilka & Bonneh, 2022), which, combined with the previous study's findings, suggests potential connections between the three concepts.

The studies presented within this chapter aimed to investigate the relationships between the size of an individual's temporal binding window, presence, and simulator sickness. In the following section, the concept of simulator sickness is detailed and explored in relation to presence and TBWs. The rationales for the adjustments to the study are also discussed before specific hypotheses are outlined.

4.1.1 *Simulator Sickness*

Simulator sickness, also known as cybersickness, is a sub-category of motion sickness (Johnson, 2005; Grassini et al., 2021) that occurs within virtual reality systems, such as head-mounted displays (HMDs). Symptoms of discomfort, such as dizziness, nausea, and sweating, characterise both forms of sickness. These symptoms can be assessed through physiological measures, such as galvanic skin response (Caserman et al., 2021), or subjective measures, such

as the Simulator Sickness Questionnaire (Kennedy et al., 1993). Nevertheless, despite the awareness of the negative impacts of motion and simulator sickness, there is no agreed scientific explanation for what causes either (Grassini et al., 2021; Howard & Van Zandt, 2021; Saredakis et al., 2020). Although there is no agreed-upon explanation of why either form of sickness occurs, there are two leading theories for motion sickness, which also apply to simulator sickness (Rebenitsch & Owen, 2016). These are the postural instability (Riccio & Stoffregen, 1991) and sensory mismatch theories (Reason & Brand, 1975).

The postural instability theory considers motion sickness to be caused by the lack of control over posture in an environment that induces motion (Johnson, 2005). The control of posture here relates to the ability to dynamically adjust various body elements to allow for different behaviours or goals (Riccio & Stoffregen, 1991). For example, how a person controls their stance on a swaying boat versus a stationary path. On a boat, the combination of different bodily configurations (knees, hips, torso, arms, etc.) is very different from how these would be organised on land. If an individual can control their posture, it is argued to be stable. However, if the individual does not possess, and cannot develop a way to control their posture, this leads to instability (Riccio & Stroffregen, 1988). The combination of a prolonged state of instability and an individual's lack of a sufficiently effective strategy to maintain stability in environments such as VR leads to motion sickness, or in this case, simulator sickness (Riccio & Stoffregen, 1991; Grassini et al., 2021). Studies have reported mixed evidence for the postural stability hypothesis as an explanation of simulator sickness, with a larger number of studies identifying that motion sickness still occurs in environments when an individual is stable (Dennison & D'Zmura, 2017; Littleskare, 2021; Saredakis et al., 2020). Suggesting that, at best, postural instability may be a precursor to simulator sickness and that motion sickness may stem from another cause (Lawson et al., 2021).

Sensory mismatch theory, also known as the sensory conflict theory (Kohl, 1983), proposes that motion sickness arises from a mismatch of sensory information relating to spatial orientation (Reason & Brand, 1975). The visual, vestibular, proprioceptive and somatosensory systems receive sensory information regarding motion. This sensory information is then compared to expected physical motion within the environment based on stored information regarding past motion history (Reason & Brand, 1975). It is hypothesised that when the perceived sensory input does not match the expected motion, such as in VR, sickness occurs (e.g. Grassini et al., 2021). The sensory conflict explanation has received the most support of the two leading theories of simulator sickness (e.g. Johnson, 2005). It is, therefore, considered the most commonly accepted explanation of how simulator sickness arises (Lawson et al., 2021; Warwick-Evans et al., 1998). A conflict of sensory representations may explain why some individuals reported experiencing a degree of discomfort in the previous study.

The sensory conflict which causes simulator sickness is thought to be due to two different subsets of motion sickness. The most common form of sickness in VR is thought to be visually-induced motion sickness (VIMS). VIMS occurs in environments where the visual sensory input generates an illusory sense of motion called vection (Lubeck et al., 2015; Riecke & Schulte-Pelkum, 2013). Vection typically involves dynamic visual stimuli, such as moving through a virtual world. A conflict arises when there is a mismatch between this visual perception of vection and corresponding vestibular information (Bonato et al., 2008). For instance, if a participant was seated in a HMD yet perceived the world moving around them.

On the other hand, motion sickness can also be vestibular-induced. Vestibular-induced motion sickness is more common in real-world examples, such as in moving vehicles, where the vestibular information indicates movement, but the visual information does not (Chang et al., 2020). In more recent HMDs, simulator sickness may also be vestibular-induced. Vestibular-induced simulator sickness would occur if the individual can navigate around a

given area whilst wearing the headset, but the visual information is not correctly synchronised. At the heart of either explanation lies the conflict between different sensory inputs.

Regardless of how simulator sickness may arise, it is evident when it does. Individuals express their feelings of discomfort, including nausea, eye strain and headaches. In order to capture the severity of the discomfort, multiple measures have been developed. Objective measures such as Galvanic Skin Response (e.g., Al-Ashwal et al., 2021) attempt to provide a measure that is not impacted by an individual's judgements. These measures record physiological changes in the body whilst experiencing a simulation that is likely to cause discomfort, such as conflicting cues (Islam et al., 2020). Much like the objective measures used in presence research, there are issues with using them to investigate simulator sickness. Primarily, despite the claim that the measures are objective, subjective judgements still occur (Grassini et al., 2020). For instance, isolating what a potential change in Galvanic Skin Response genuinely signifies. Instead, much like presence, subjective measures tend to be preferred (e.g. Johnson, 2005). At least five questionnaires have been developed to assess simulator sickness; the Motion Sickness Susceptibility Questionnaire (Golding, 1998), the Fast Motion Sickness Scale (Keshavarz & Hecht, 2011), the Cybersickness Questionnaire (Stone et al., 2017), and the Virtual Reality Sickness Questionnaire (Kim et al., 2018). However, the questionnaire which is considered to be the golden standard (Johnson, 2005) is the Simulator Sickness Questionnaire (SSQ; Kennedy et al., 1993).

The SSQ was derived from the Pensacola Motion Sickness Questionnaire (MSQ) (Kellogg et al., 1965) as a shortened and fine-tuned version specifically for assessing sickness within virtual reality simulators (e.g. flight simulations). Items on the MSQ were considered either irrelevant or misleading for use within simulators (Kennedy et al., 1993). A sample of 1119 pairs of pre- and post-exposure MSQs from navy flight simulators were used for the SSQ development. Comparing the pre- and post-exposure results, infrequent, no change and

misleading items were removed. Out of the original 28 items, 16 were retained and measured on a 4-point scale. These 16 items were subjected to a series of factor analyses, identifying a 3-factor solution: Oculomotor Discomfort, Disorientation, and Nausea. The unit weights on the subscales identified from the varimax rotation are then added together and multiplied by a conversion factor to get a total score for each subscale. A total severity score can be calculated by adding all the item weightings and applying its own conversion formula. The reason for using conversion formulas, rather than just the totals, is to provide similar variabilities across the subscales. A range of similar variabilities makes the SSQ a valuable tool for investigating the total severity of discomfort caused by a simulation and the simulation's comparable impact on different elements of simulator sickness (Balk et al., 2013).

With the SSQ being considered the gold standard of simulation sickness measurement (Johnson, 2005), multiple studies have used the measure to investigate factors that might affect the severity of the discomfort within VR. These factors are split down into human factors and VR-based factors. Human factors include age (e.g., Hakkinen et al., 2002; Keshavarz et al., 2018), gender (e.g., Jaeger & Mourant, 2001; Park et al., 2006; Rangelova et al., 2020), and prior VR experience (e.g., Freitag et al., 2016; Stanney et al., 2003). In comparison, the VR factors relate to duration (Liu & Uang, 2012; Stanney et al., 2003), latency (e.g., Moss et al., 2011; Stauffert et al., 2018), optical flow (e.g., Bonato et al., 2008; Keshavarz et al., 2019), frames of reference (e.g., Hwang et al., 2012; Prothero et al., 1999), field of view (e.g., Lin et al., 2002; Seay et al., 2001) and task type (e.g., Nelson et al., 2000; Weißkert et al., 2018).

Meta-analyses have identified mixed results regarding human-related factors (Chang et al., 2020; Grassini et al., 2021; Howard & Van Zandt, 2021; Saredakis et al., 2020). For example, females have been reported to be more susceptible to simulator sickness, with higher SSQ scores than males (Chang et al., 2020; Howard & Van Zandt, 2021). At least three explanations have been proposed for why females may be more susceptible than males. Firstly,

it is highlighted that the use of HMDs may be affected by the fact that females generally have a smaller interpupillary distance (Fulvio et al., 2021). This is especially true within HMDs which do not provide an adjustable lens to account for this genetic difference. Females have also been recorded to generally have an increased postural sway (Arcioni et al., 2019). Increased sway can be a precursor to postural instability and therefore, simulator sickness in virtual environments which induce feelings of motion (Curry et al., 2020). Despite the proposed explanations, inconsistencies within the literature question whether this is an inherent gender difference (Lawson et al., 2021). Lawson and colleagues (2021) outline that out of 64 studies investigating gender differences, only 31 (48%) found females more susceptible to simulator sickness. Furthermore, the studies that reported females as more susceptible controlled for less confounding variables. Instead, it is proposed that females are not any more susceptible to simulator sickness; rather the gender difference is an artefact of unsuitable HMDs and males' tendency to under-report feelings of simulator sickness (Bannigan et al., 2024; Rangelova et al., 2020).

In reference to age, Saredakis et al. (2020) highlighted that although older individuals (35+) tend to report significantly lower SSQ scores, a limited number of studies investigate a wide age range. In Contrast, there are studies in which the opposite direction is reported (Hakkinen et al., 2002; Keshavarz et al., 2018; Park et al., 2006). For example, Hakkinen and colleagues (2002) found that an older age group (25.5+) reported an increased SSQ oculomotor discomfort score after experiencing a virtual race environment in an HMD. One potential explanation for these mixed results is that it has more to do with experience than just age alone (Lawson et al., 2021). Previous experience with virtual reality (VR) has a generally negative relationship with simulator sickness. Most studies indicate that individuals who use VR more frequently tend to experience less severe symptoms of simulator sickness (Howard & Van Zandt, 2021). This suggests that increased VR usage may help reduce the intensity of simulator

sickness over time. Overall experience with VR could be one potential reason for the mixed findings above.

Unlike human factors, the influence of VR-based factors is more clearly distinguished. These VR-based factors can be split into content (duration, task type, frames of reference, and optical flow) and hardware (latency, field of view).

Research has shown that the longer someone is in VR, the more likely they are to develop symptoms of simulator sickness (Dużmańska et al., 2018; Kennedy et al., 2000; Liu & Uang, 2012; Stanney et al., 2003). This research has led to recommendations surrounding the duration individuals should use VR before taking breaks (10-15 minutes every hour: Yao et al., 2014). Alongside duration, the type of task also affects the severity of simulator sickness (Chang et al., 2020). Task types can either be Passive or Active. Passive experiences are ones where the individual cannot control the navigation within the virtual reality environment (VRE); Active experiences are the opposite. Simulator sickness is reported to occur more severely in passive tasks than in active tasks (Chen et al., 2011), which aligns closely with sensory conflict theory. The theory outlines that sickness occurs when an individual's motion, as perceived by their visual sense, does not align with their expected motion (Reason & Brand, 1975) due to their inability to control it. One way to counter the impact of this sensory conflict is to introduce fixed visual stimuli such as trees (Lin et al., 2002) or grid patterns (Hwang et al., 2012). These fixed stimuli provide a frame of reference, reducing sensory conflict and simulator sickness (Prothero et al., 1999). However, simulator sickness may still arise from the conflict between visual and vestibular information. The final content-related factor is similar to the previous two. Optical Flow relates to the perceived vection through the virtual environment caused by moving (rather than stationary) stimuli. For example, studies have used "star field" environments, which place users in a seemingly infinite space surrounded by white dots (e.g., Keshavarz et al., 2019). These dots are programmed to move past the user at varying

speeds, providing an illusory movement, which has been shown to increase simulator sickness (Bonato et al., 2008; So et al., 2001). The denser the stimuli are, and the faster these stimuli move, leads to an increase in both perceived vection and simulator sickness (Chang et al., 2020).

The other factors related to virtual reality that influence simulator sickness are associated with the hardware used. The field of view of the hardware refers to how much of the environment the individuals can perceive. It is widely reported that a field of view that is too wide is associated with increased simulator sickness (e.g. Fernandes & Feiner, 2016; Lin et al., 2002; Seay et al., 2001). The increase in simulator sickness is thought to be due to the increase in the amount of conflicting information being presented (Van Emmerik et al., 2011). Another hardware factor that impacts simulator sickness is Latency. Latency refers to the time gap between a user performing an action, such as moving their head, and the headset accurately displaying the corresponding image (Rebenitsch & Owen, 2016). A longer delay between the user's action and the visual update in the headset results in heightened sensations of simulator sickness (DiZio & Lackner, 2000). Furthermore, it has been suggested that varying levels of latency within HMDs increase simulator sickness compared to a fixed latency, regardless of delay (Draper et al., 2001). Suggesting some malleability when it comes to experiencing offset information.

Over the past twenty years, HMD developers have implemented adjustments to reduce the negative impact of simulator sickness. A review by Kourtesis and colleagues (2019) highlighted multiple technological advancements that aim to alleviate feelings of simulator sickness based on some of the impact factors above. Display screens have been upgraded from Cathode Ray Tubes (CRT) to Liquid Crystal Display (LCD) and now to Organic Light Emitting Diode (OLED). OLEDs help reduce motion blur and improve black-to-white contrast ratios (Kim et al., 2017). Both of these improvements reduce issues surrounding perceived latency.

Resolution per eye has increased from around 480x240 pixels to 960x1080 pixels, and the refresh rate has also been improved from 45hz to 75hz. Both improvements aim to reduce the delay in the visual image updating. The field of view has improved from around 40 degrees in older displays (e.g., VFX 3D: Mindflux, 2000) to around 110 degrees in the modern HMD (e.g., Oculus Rift: Oculus VR, 2016; Kourtesis et al., 2019). However, as argued above, too wide a field of view can potentially increase simulator sickness (Fernandes & Feiner, 2016), whilst too narrow a field of view impacts the immersive quality of the technology.

It is debated whether the technological advancements of HMDs have reduced the impact of simulator sickness (Saredakis et al., 2020). Despite the uncertainty around whether improvements to the technology can help alleviate simulator sickness, the technology's immersive qualities have been shown to improve (Cummings & Bailenson, 2016). Improving the technology's immersive qualities leads to improvements in individuals' sense of presence. Whilst one might assume that experiencing more of the positive elements of VR would reduce the negative, the relationship between presence and simulator sickness is known to be complex.

4.1.2 *Simulator Sickness and Presence*

While presence measures the subjective experience of immersion, simulator sickness is a subjective measure of individuals' discomfort whilst using virtual reality (Weech et al., 2019). One might consider presence to be a positive user experience of VR and simulator sickness to be a negative. However, the relationship between presence and simulator sickness is complex. Experiencing either presence or simulator sickness in VR does not always mean the other feeling is reduced. For example, individuals who have high immersive tendencies have been reported to go on to feel both more present within a virtual environment and yet also experience higher levels of simulator sickness (Ling et al., 2013). More so, even in the most sophisticated immersive systems (HMDs with body tracking), some individuals may still experience

simulator sickness (Rebenitsch & Owen, 2016). Suggesting that the technology involved in eliciting a greater sense of presence may not always successfully eliminate the impact of the simulator sickness.

In a recent review by Weech and Colleagues (2019), both negative correlations (e.g., Witmer et al.,1996) and positive correlations (e.g., Wilson et al., 1997) were highlighted between the two subjective feelings. Eleven of the twenty studies included in the review displayed a negative relationship between presence and simulator sickness. Only six displayed a positive relationship. However, most studies included in the review used HMDs which were developed before the mid-2000s (17 out of 20), which has been argued to heighten the relationship between the two concepts (Saredakis et al., 2020). There have only been a few studies measuring simulator sickness in more modern head-mounted displays (HMDs) developed after the mid-2000s. To date, no connection has been found between simulator sickness and the sense of presence reported by participants using these devices (Ryan & Griffin, 2016; Grassini et al., 2021). The technical quality of the immersive experience may not be the only factor which leads to differences between presence and simulator sickness. Other factors which might impact this relationship include the type of content generated (Saredakis et al., 2020) and the individual's susceptibility (Stanney et al., 2020; Tian et al., 2022). This indicates that alongside presence, individual differences may impact levels of simulator sickness within virtual realities (Weech et al., 2019). As such, it is important to explore the relationship between these factors to better understand how individual differences, like temporal binding windows, contribute to these experiences.

4.1.3 *Simulator Sickness and Temporal Binding Windows (TBW)*

The relationship between simulator sickness and temporal binding windows has only recently begun to be explored. The idea that the two concepts are related originates from the

sensory mismatch theory (Kohl, 1983; Oman, 1990; Reason & Brand, 1975). It highlights that symptoms of simulator sickness may arise when sensory information about the virtual environment is conflicting or desynchronised. Typically, this relates to the perceived visual stimuli and the expected physical motion (Grassini et al., 2021). If one does not match the other, this can lead to postural instability, and sickness will likely increase. Much like in motion sickness (Riccio & Stoffregen, 1991). If individuals bind dissociative vestibular-visual information, this conflicting information increases the susceptibility to experiencing sickness (Zilka & Bonneh, 2022). Vestibular information is processed uniquely compared to other sensory inputs, such as visual and audio, because it does not have a unimodal cortex (Guldin & Grusser, 1998). Instead, it has a widespread network distributed across multiple cortices (Lopez et al., 2012) and contributes to several brain functions through multisensory interactions (Ferrè et al., 2015). One of these interactions is the integration of vestibular and visual cues, which combines optical flow with perceived self-motion (Chen et al., 2010; Fetsch et al., 2009; Prsa et al., 2012). Therefore, if visual cues generate perceived vection, while the vestibular network provides no cues regarding movement, the conflicting information is likely to lead to feelings of sickness (Zilka & Bonneh, 2022). Much like audio-visual integration, there is a window of time where the information can be offset yet perceived as simultaneous, and this vestibular-visual TBW also differs between individuals (Shayman et al., 2018). These differences mean that some individuals can better handle delays between visual and vestibular information, and others are more proficient in separating them into two processes.

Delays in updating visual information from modern HMDs can be up to 20 ms (Van Waveren, 2016; Zhang & Luo, 2012). Therefore, if an individual binds the information from their physical movement with the delayed visual information displayed, it is hypothesised that they will experience simulator sickness (Sadiq, 2019). Whilst this is a vestibular-visual interaction, audio-visual temporal binding can be used to assess this relationship, as

multisensory integration has common properties across modalities (Rajagopalan et al., 2017; Zilka & Bonneh, 2022). Although there has been little research into the relationship between the two concepts, the three studies which have investigated the relationship have found that wider temporal binding windows are associated with higher susceptibility to simulator sickness (Sadiq, 2019; Sadiq & Barnett-Cowan, 2022; Zilka & Bonneh, 2022). The results from the study presented in the previous chapter indicated that narrower TBWs led to higher reported presence. However, this relationship might be confounded by individuals' levels of simulator sickness. There is the possibility that individuals with a broader window bound more conflicting information and, therefore, reported diminished feelings of presence due to experiencing sensory conflict. This aligns with some participants' verbal reports of experiencing simulator sickness during the simulation. This supports the hypothesis that the relationship between TBWs and simulator sickness arises because individuals with a wider TBW are more likely to integrate conflicting visual and vestibular information into a single perception (Shayman et al., 2018). This may reduce individuals' attention to the environment or the subsequent spatial situational model more than those who do not experience this sensory conflict (i.e., those with a narrower window). In order To measure individuals' self-reported levels of simulator sickness, the SSQ (Kennedy et al., 1993) was included in this study. The questionnaire was administered after participants had completed their VR experience.

4.2 Study 2a: Pilot of Online Study

4.2.1 *Rationale and Adjustments for an Online Study*

Smaller samples are typically used in both exploratory research and research investigating TBW (e.g. Megevend et al., 2013; Powers et al., 2009; Slater & Usoh, 2013). Despite this, a larger sample is needed to expand upon the previous findings with sufficient statistical power (Abraham & Russell, 2008; Cohen, 1992). A decision was taken to move to

online research to maximise the number of participants. The move to online research was due to two factors. Primarily, this study began data collection in the second half of 2020, when conventional methodologies were restricted due to the COVID-19 pandemic. However, the unprecedented nature of the pandemic led to improvements in online data collection software as researchers across the globe turned to these alternative methods (Anwyl-Irvine et al., 2021). This improved online software with Prolific's online participant recruitment scheme, which had access to over 100,000 vetted professional participants at the time, allowed for a much larger target population (Peer et al., 2022). A pilot study was conducted to validate that online recruitment and testing would produce results comparable to study 1.

The pilot study featured four variations from the lab-based study detailed in the previous chapter. To measure simulator sickness along with the sense of presence, the SSQ (Kennedy et al., 1993) was administered alongside the presence questionnaires. In this study, the presence questionnaire used with the SPES was the SUS rather than the IPQ as mentioned in Section 4.2.3.2. Additionally, the TBW tasks were re-coded into millisecond Inquisit 6 (2020) instead of PsychoPy 2 (Peirce et al, 2019). This change was necessary due to the reliability in timing accuracy of the respective software when used over the internet. Although the newer Psychopy 3 had the online data-collecting capability (Peirce et al., 2019), Inquisit was the only software to install a client on the participant's computer, which tracks system clock speeds (Inquisit 6, 2020). Due to this, it was considered the best online software for millisecond accuracy. Considering the tasks measure participants' responses in millisecond increments, the software's accuracy was pivotal for estimating the size of individuals' TBWs.

In addition to re-coding the tasks, the criteria for participating in the research had to be adjusted. Due to participants being recruited through Prolific, the study relied on participants needing their own VR equipment. At the time, there was no available data on how many individuals on Prolific owned an Oculus Rift headset or had access to Minecraft VR. To

accommodate this, the participation criteria was expanded to include a variety of VR headsets and simulations.

4.2.2 *Aims and Hypotheses*

The following study aimed to build upon the results of the exploratory research detailed in Chapter 3 by increasing the sample size and including a measure of simulator sickness (SSQ: Kennedy et al., 1993). To check whether the adjustments for online testing still provided similar data to the lab-based study for the TBWs and presence, a small pilot study was conducted (Study 2a). Although the online participant recruitment required individuals who owned their own VR HMDs, this was only expected to affect how much gaming and VR experience these individuals had. Data was analysed before further adjustments (detailed in Section 4.2.4) to the complete study (Study 2b) were made.

The study's aims were two-fold whilst maintaining the overarching aim of the thesis: to investigate the relationship between individuals' sense of presence and their temporal binding windows. The first aim was to assess whether online measurement of the variables identified within the first study was feasible and replicable. The suitability of online testing for exploring the relationship between presence and temporal binding (H1) may be influenced by differences in participant characteristics, particularly past gaming and VR experience (H2). These differences, inherent to the online recruitment process, could potentially impact the comparability of scores between the online pilot and lab study.

H1: If online testing is unsuitable for exploring the relationship between presence and temporal binding, significant differences will be observed between Presence, RT TBW, and TOJ TBW scores in the online pilot and lab study. Although no major differences are expected between these primary measures, minor deviations may occur due to variations in participant characteristics.

H2: Participants in the online study are hypothesised to have significantly higher past gaming and VR experience compared to lab-based participants. This is attributed to the recruitment criteria, which target individuals who own a VR headset and are therefore more likely to have prior experience with these activities..

The second aim was to investigate the confounding relationship of simulator sickness between the temporal binding window and presence. Individuals with narrower TBWs have been shown to report higher levels of presence (study 1) and lower levels of simulator sickness (e.g., Zilka and Bonneh, 2022). The connection between presence and simulator sickness is debated; however, individual differences in temporal binding may provide some insights into this complex relationship. In situations where people are presented with a virtual environment containing conflicting visual-vestibular information, the relationship between presence and simulator sickness may be influenced by the size of an individual's temporal binding window. A narrow TBW may be associated with a higher sense of presence and lower severity of simulator sickness, while a wider TBW may be associated with the opposite. If the presence or absence of conflicting cues is causally related to simulator sickness then, in environments with no conflicting information, it is hypothesised that there would be no relationship between TBW and simulator sickness, as there is no sensory mismatch. Nevertheless, in these environments, wider TBWs might lead to quicker attention and better maintenance of a spatial situational model, resulting in higher levels of presence. Therefore, the following hypothesis was investigated:

H3: It is hypothesised that individuals' TBW widths will correlate negatively with presence and positively with simulator sickness.

This is likely due to individuals with a wider TBW combining incongruent cues surrounding their environment, diminishing their sense of presence and increasing sensory conflicts involved in simulator sickness.

4.2.3 *Methods*

4.2.3.1 **Participants**

Eighteen participants (14 male, mean age = 30.17 years, SD = 9.06 years, range = 18 - 49 years) were recruited through Prolific using the following criteria: their nationality was limited to UK (English, Welsh, Scottish or Irish), they answered yes to being fluent in English, they had no self-reported literacy difficulties, and they both owned a VR headset and had used it at least once. Out of the 18 participants, 17 submitted complete data. Each participant was paid £7.69p/hr for taking part. Completion times ranged from 32 to 68 minutes.

4.2.3.2 **Materials**

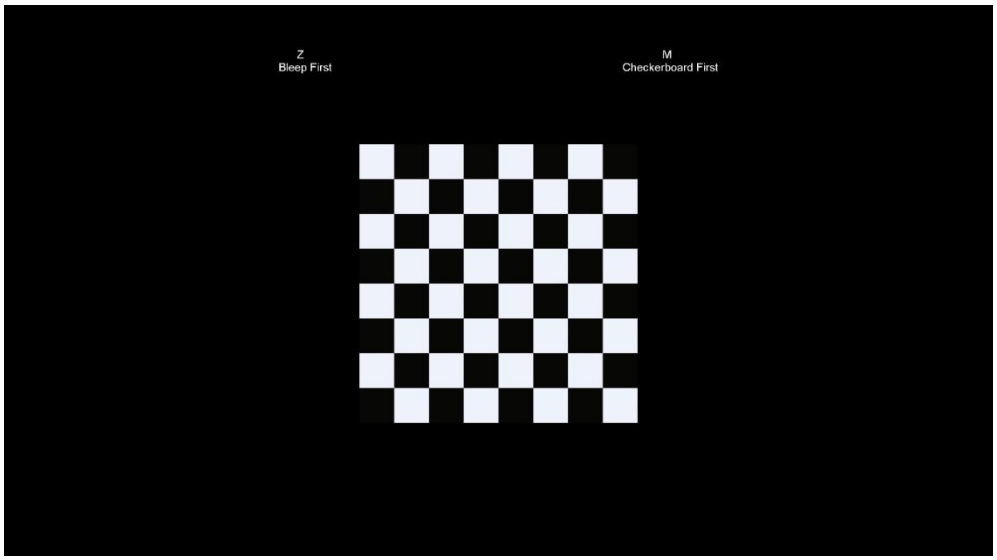
Temporal Binding Tasks

The individuals' temporal binding window was measured using two of the three tasks used in study 1 (RT and TOJ). The tasks were re-coded in millisecond Inquisit to measure millisecond-accurate online responses. All tasks used the same flash-beep stimuli as Study 1. However, the equipment which presented the stimuli could not be recorded as individuals provided their own monitors and headphones. Volume was requested to be set at 40% for each participant.

The tasks were identical to Study 1, except for a small visual reminder on the screen highlighting which button related to which judgment in the TOJ task (Figure 7). The visual reminder was included as the researcher could not be present to remind individuals if they forgot which key related to which judgment.

Figure 7

Screenshot of a single trial from an online temporal order judgment (TOJ) task



VR Hardware and Simulation

There was no limitation on which VR equipment participants could use to widen the range of participants who could participate in the study. Instead, participants were asked to report the hardware they used. Whilst the majority of participants reported using an Oculus Rift headset (6) similar to study 1, the range of headsets used also included HTC Vive (3), Gear VR (3), PlayStation VR (2), Oculus Rift S (1), Oculus Quest (1), Keplar (1), and Google Cardboard (1). To support the differences in specifications of various headsets, participants were also allowed to choose the simulation they would use during the study, which they reported. The most common simulation (3) was Beat Saber (Oculus Studios). The remaining simulations were unique: Abode, Distance, Eve: Valkyrie, Everybody's Golf VR, Half-Life: Alyx, Lone Echo, Marvel: Powers United, Ocean Rift, Puzzle Maze, Rush, Ski Game, Skyrim, Subnautica VR, Tetris, and 6x9 Solitary Confinement.

It is acknowledged that allowing such a range of headsets and simulations could reduce the pilot's rigour. Nevertheless, allowing participants to use their headsets and simulations had

two major benefits. First, allowing more freedom in recruitment permitted a wider selection of people to participate. A wider selection of participants had the desired effect of increasing the online pilot's recruitment speed. Secondly, if a relationship is identified between TBW and presence across a wider range of scenarios, the effect of TBW on presence can be considered more generalizable rather than limited to a specific scenario.

Questionnaires

Two presence questionnaires were used to assess individuals' self-reported feelings of presence within their chosen virtual environment. This study used the Slater, Usoh, and Steed Presence Questionnaire (SUS; Slater et al., 1994) as an alternative to the Igroup Presence Questionnaire (IPQ; Schubert et al., 2001). The change in the presence questionnaire was necessary because the realism scale of the IPQ did not correlate with the Spatial Presence Experience Scale (SPES) in Study 1, while the SUS does not include a measure of realism. In addition to the presence questionnaires, the Simulator Sickness Questionnaire (SSQ; Kennedy et al., 1993) was utilized to measure simulator sickness. Unlike in Study 1, these questionnaires were coded into the Inquisit software for online distribution with the tasks.

Spatial Presence Experience Scale (SPES): The first questionnaire used in the online pilot was identical to that previously used in Study 1. The SPES was found to have good construct validity through its correlations with other presence measures, supporting past research identifying the same (Hartmann et al., 2016). The SPES was retained for this online pilot due to its relationship to the underpinning theory of presence (Wirth et al., 2007).

Slater, Usoh and Steed Presence Questionnaire (SUS): The version of the SUS utilised in this pilot study was the 3-item version (Slater et al., 1994). The first question, '*In the computer-generated world I had a sense of 'being there'…*' relates to a general sense of presence. This question is also used as the general measure of presence in the IPQ (Schubert et

al., 2001), and responses range from ‘not at all’ to ‘very much.’ The second ‘*There were times during the experience when the computer-generated world became more real or present for me compared to the "real world"...*’ relates to spatial updating with answers ranging from ‘at no time’ to ‘almost all the time’. The final question, '*The computer-generated world seems to me to be more like…*’ relates to the suspension of disbelief and is answered on a scale from 'someplace I saw' to 'somewhere I visited'.

The three items are rated on a Likert-type scale of 1 to 7. The scoring strategy has changed over time since its development (between binary and continuous measures of presence). However, in this study, the questionnaire was scored continuously, which aligns with the process model's theory of presence. Therefore, the questionnaire provides a range of responses from 3 to 21.

Simulator Sickness Questionnaire (SSQ): The simulator sickness questionnaire (SSQ) was developed by Kennedy and colleagues (Kennedy et al., 1993). It comprises of 16 questions that are loaded onto three different scales. Firstly, a Nausea scale was developed to assess the severity of nausea, stomach awareness, increased salivation, and burping. Secondly, the Oculomotor Discomfort Scale identifies the severity of eyestrain, difficulty focusing, blurred vision, and headache. Finally, a Disorientation scale is used, which involves questions measuring the severity of dizziness and vertigo. Together, the questions form one general simulator sickness scale, which is often used as its components or total score (for a full guide to scoring see Section 4.2.4.3).

Previous Gaming Experience: The final questionnaire was the same as Study 1, a measure of previous gaming experience. Although the measure did not correlate with any other variables in Study 1, it was retained. See Section 3.2.2.3 for more details. As with all of the

questionnaires, the questions were coded into the experiment script rather than given as paper copies.

4.2.3.3 **Procedure**

Participants followed the advertised link on Prolific, which redirected them to the experiment hosted on Inquisit Web. They were asked to download a small piece of client software to enable the smooth running and recording of the experiment. Each participant was alternatively assigned to one of two counterbalanced groups, which determined whether the redundant target (RT) task or the temporal order judgment (TOJ) task was presented first.

Participants were introduced to the study via on-screen text and reminded of their right to withdraw. In addition, participants were required to meet certain criteria for participation in the study, including possessing a functional VR headset, which was set up, having access to a stopwatch, and setting the headphone volume on their computer to 40%. The following screen asked participants to confirm they had a VR headset and write which one in an open field text box. Finally, participants were informed about the structure of the study, that they would complete two different cognitive tasks, then go and experience their VR simulation for 10-30 minutes before finally filling in some questionnaires. If the participants were happy with the above, they pressed continue, which started the study.

Participants were introduced to the temporal binding window tasks through practice trials before continuing to complete the tasks as highlighted above. Once participants reached the end of the tasks, they were advised that if they wished to take a short break, they could do so. Furthermore, they were instructed that once they pressed continue, they would have a minimum of 10 minutes to play in the VR of their choosing up to a maximum of 30 minutes. They were instructed to use a stopwatch to record how long they spent in the VR and also informed that a timer would display on the screen to outline the minimum time expected for them to experience the VR. Once the participant clicked continue, a countdown timer was

displayed, counting down from 10 minutes. When the timer finished, the next instruction screen was displayed, ready for when the participant re-joined the experiment. On average, participants reported spending 13 minutes and 43 seconds in VR.

This final instruction screen informed participants that they would have a series of questionnaires to fill in. The first screen asked for participant information such as age and gender. The screen that followed asked an attention check question to monitor for low response and 'bot' participants: '*What were the three response keys you were required to press?*'. Participants provided their answers in an open text box, which was reviewed after they completed the study. The next screen asked individuals to report details of their VR experience, asking questions about what VR equipment they used, what VR software they experienced, and how long they spent within it. Following these questions, the participants were asked to answer the four questionnaires outlined above (SPES, SUS, SSQ, and the Previous Gaming Experience).

Afterwards, participants were thanked for their time and asked to provide any feedback in a textbox. Finally, they were displayed a debrief and redirected back to Prolific with a completion code to confirm that they had completed the study.

4.2.3.4 **Ethics**

The research followed the BPS guidelines for conducting ethical online research. Consent was obtained, and participants were thoroughly informed about every aspect of the study, including the right to withdraw. Furthermore, the study operated per Prolific's rules and conditions, which are designed to safeguard participants and researchers. Participants all owned their own VR equipment, which would have been sold with health and safety warnings from the manufacturers. During this study, participants were not asked to do anything that would go against these health and safety warnings. Therefore, because the researcher was not physically present, the participant was responsible for adhering to these health and safety measures. The

study was first approved for in-person testing by the York St John University Research Ethics Committee; however, later amendments were also approved for the adjustment to online research.

4.2.4 *Analysis*

4.2.4.1 **Participant Data Checks**

Two steps were taken to check the reliability of the participant's data. Firstly, participant responses to the attentional questions were checked to see if any responses failed to identify a reasonable level of attention paid to the study. Participants who were deemed inattentive were excluded from further analysis. Inattentive individuals included those who incorrectly identified the response keys they were supposed to press, failed to specify which VR headset they used, or provided obviously random responses (e.g., inserting unrelated words in text boxes).

After removing the data from eight participants who failed the attention checks, the remaining data was checked for spam/repeat responses. Any participant who displayed a 1-2 repeat response strategy, a series of 1 or 2 response strategies, or those individuals constantly providing incorrect judgments for extreme stimulus onset asynchronies (SOAs) were excluded from further analysis. None of the remaining 18 participants exhibited any problematic response strategy. The raw data from these participants were then used to calculate each individual's temporal binding window widths. These window widths were generated for each task, in line with the analysis conducted in Study 1 (Section 3.3).

4.2.4.2 **Calculating the TBWs**

As highlighted in the analysis of Study 1 data, although the different tasks relate to temporal binding, they address slightly different processes. However, in Study 1, this comparison was made between three different tasks, and no model considered ways to analyse all three tasks within one set of calculations. In comparison, this study included only two tasks,

the TOJ and the RT. As previously highlighted, these tasks have a calculation encompassing a method to analyse both tasks within the same model. This is the canonical adaption to the twin model (Diederich & Colonius, 2015). However, as this pilot study aimed to identify if Study 1 could be replicated on a larger scale online, the decision was made to analyse each task data using the same models as before. Using the same models allows a direct comparison between the windows generated in the pilot and first studies. The window generation for each task followed the same process as highlighted in the analysis section of Study 1 (see Sections 3.3.2 and 3.3.3). One participant's RT TBW was calculated to be above 3000ms and was completely removed from all further analyses. The descriptive data for the two temporal binding windows are displayed below (Table 6).

Table 6

Descriptive Statistics for the TBW Tasks in Study 2a

Task Type (N)	Mean (ms)	Min (ms)	Max (ms)
Redundant Target (17)	21.88	9.59	56.38
Temporal Order (17)	118.31	4.16	299.98

4.2.4.3 Descriptive Statistics for the Questionnaires

The following section details how the raw data for each questionnaire was converted to total scores and reports the descriptive statistics for each measure.

The SPES scores were calculated the same way as in the first study; both subscales (Self-location and Perceived possible actions) were calculated by adding the first and last four questions, respectively. These scores ranged from a minimum of 4 to a maximum of 20. The SPES total was then calculated by adding the scores for each subscale, ranging from 8 to 40.

Participants' previous gaming experience was also calculated similarly to the first study. The average daily gaming usage was multiplied by the average number of days per week spent

gaming to provide a continuous gaming experience score. Additionally, previous VR experience was calculated by multiplying the number of VR sessions played by the average duration of each session.

Previously, the SUS has been calculated in two separate ways based on a researcher's theoretical stance on presence (Slater et al., 1994; Slater et al., 2000), either by counting only the top two responses of each question or counting the questionnaire continuously. In line with the other presence questionnaire (SPES) used here and the theoretical stance within this thesis that presence is continuous, the SUS total was calculated on a continuous scale. Each question ranging from 1 to 7 provided a total score of 3 to 21.

The SSQ consists of four distinct but interconnected scales, each calculated using a different method. The overall score is calculated by adding all the items and multiplying the sum by 3.74 to achieve a total score ranging from 0 to 179.52. The three other subscales, nausea, oculomotor discomfort, and disorientation, are each calculated by adding seven items (some overlap onto several factors), then multiplying each value by a specific weight. For nausea, the sum of items 1,6,7,8,9,15, and 16 is multiplied by 9.54 to give a total score ranging from 0 to 200.34. The total score for oculomotor discomfort is calculated by multiplying 7.58 by the sum of items 1,2,3,4,5,9, and 11, giving a range of 0 to 159.18. Finally, the sum of items 5,8,10,11,12,13, and 14 multiplied by 13.92 gives the total disorientation score, ranging from 0 to 292.32. The descriptive statistics for all questionnaires are reported in the table below (Table 7).

Table 7

Descriptive Statistics for Questionnaires in Study 2a

Questionnaire (N)	Mean	Standard Deviation	Minimum	Maximum
Gaming experience (17)	18.68	13.97	1.00	56.00
VR Experience (16)	39.93	51.49	1.00	200.00
SPES Total (17)	27.29	6.55	14.00	39.00

SPES SL (17)	13.59	3.95	6.00	20.00
SPES PPA (17)	13.71	2.89	8.00	19.00
SUS Total (17)	12.18	3.68	4.00	20.00
SSQ Total (17)	45.10	31.97	11.22	134.64
Nausea Score (17)	32.55	29.61	0.00	114.48
Oculomotor Score (17)	36.12	24.05	7.58	98.54
Disorientation Score (17)	54.04	44.53	0.00	153.12

Note: SL = Self-Location; PPA = Perceived Possible Actions; SSQ = Simulator Sickness Questionnaire

4.2.4.4 **Construct Validity of the Presence Scales**

In order to check the construct validity of the SPES against the SUS, nonparametric correlations (Kendall's Tau) were conducted. Kendall's Tau was chosen because neither scale (SPES or SUS) met the parametric assumptions, and this correlation is considered better at dealing with small sample sizes (Bonett & Wright, 2000).

A significant relationship was identified between the SPES and the SUS (T (17) = .465, *p*=.012). Indicating that both scales were indeed measuring the same construct. Additionally, the SUS was positively correlated with the Self Location subscale of the SPES, showing a correlation of (T(17)= .434, *p*= .022). There was also a positive correlation between the SUS and the Perceived Possible Actions (PPA) subscale of the SPES, with (T (17) = .480, *p*= .012). These findings indicate that the SUS is a valuable alternative to the IPQ.

4.2.4.5 **Comparing Online Pilot Results to In-person**

The first aim of this pilot study was to see whether the measures used in Study 1 could be feasibly used online and replicate comparable results. Therefore, each measure used in the pilot was compared against comparable measures in Study 1 using Independent T-tests or its nonparametric equivalent (Mann Whitney U) when assumptions were not met.

As the Redundant Target (RT) and Temporal Order Judgment (TOJ) TBWs were calculated in the same way for both studies, they could be directly compared. Neither test was significant, indicating that the measures for RT TBW (N_1= 26, N_2= 17, U= 159, *p*=.124) and

TOJ TBW (N_1= 28, N_2= 17, U= 183, *p*=.198) did not significantly differ between in-person and online testing.

The SPES was also used identically in both studies, allowing for direct comparison. Independent T-tests revealed that the SPES total (t(43) = .970, *p*=.169) and the SPES Self Location subscale (t(43) = -.528, *p*=.300) were not significantly different between online and in-person testing. However, the SPES PPA subscale differed significantly (t(43) = 2.781, *p*=.004).

Gaming and VR experience were also calculated in the same way between studies. It was expected that individuals who owned a VR headset (a requirement of the online pilot) would have at least more VR experience, if not also more gaming experience than those in-person. No significant differences were found between the studies for Gaming experience (N_1= 28, N_2= 17, U= 174, *p*=.134). However, as hypothesised, VR experience did significantly differ (N_1= 28, N_2= 17, U= 73.5, *p<.001*), with the in-person group having an average of 18.53 minutes of VR experience and the online group having 38.27 minutes.

The SSQ and SSU were used for the first time here and, therefore, could not be compared to previous data. However, item 1 on the SUS is identical to the General presence measure on the IPQ. Therefore, comparisons were made between these two items, and no significant difference was found (t(43) = 1.368, *p*=.089).

4.2.4.6 **Assessing the Relationship between Presence, Simulator Sickness and TBW**

The second aim of this online pilot was to investigate whether simulator sickness was a confounding variable between individuals' TBW and reported feelings of presence. Kendall's Tau correlations were conducted between each of the SSQ measures, the TBW measures, and the presence measures (see Table 8 below). The TBW measures positively correlated with

some of the SSQ measures. However, simulator sickness was not found to correlate with any presence measure.

Table 8
Correlation Between Simulator Sickness, TBW and Presence Measures in Study 2a

	1.	2.	3.	4.	5.	6.	7.	8.	9.
1. SSQ Total									
2. Nausea Score	**.785****								
3. Oculomotor Score	**.604****	**.465****							
4. Disorientation Score	**.696****	**.553****	**.346***						
5. RT TBW	**.307***	.277	**.460****	.295					
6. TOJ TBW	**.322***	.277	.304	.140	.015				
7. SPES Total	.068	.047	.142	.110	.149	.075			
8. SUS Total	-.070	-.056	-.186	-.008	-.168	.015	**.465****		
9. SPES SL	.095	.097	.139	.122	.147	.039	**.926****	**.434***	
10. SPES PPA	.016	-.016	.141	.041	.219	.110	**.802****	**.480****	**.691****

Note: SSQ= Simulator Sickness Questionnaire, RT = Redundant Target, TBW = Temporal Binding Window, TOJ = Temporal Order Judgement, SL = Self-Location, PPA = Perceived Possible Actions, *. Correlation is significant at the 0.05 level (1-tailed). **. Correlation is significant at the 0.01 level (1-tailed).

Due to the small sample size associated with a pilot study, no regression analysis was conducted to explore these relationships further.

4.2.4.7 **Results Summary**

Participants' responses to the online pilot were examined to check for authenticity. Once responses to the attention check and valid responses to the TBW tasks had been verified, participants' TBWs were calculated per the procedures outlined in Study 1 (Section 3.3). Questionnaire scores were calculated from the raw data, and the SPES was confirmed to have construct validity against the SUS. A series of independent T-tests were conducted, and no significant difference was found between online and in-person testing on all measures except for the PPA subscale. Kendall Tau correlations were used to preliminarily investigate the relationship between simulator sickness, TBW and presence. These identified a relationship

between TBW tasks and the total score for the SSQ. However, no relationship was reported with presence.

4.2.5 *Discussion and Amendments*

The aims of the online pilot study were twofold. First, it investigated whether moving to online testing was feasible and replicable of in-person testing. Second, it investigated the possibility of simulator sickness confounding the relationship between individuals' TBWs and reported feelings of presence.

The results from the pilot study support the replicability of the measures used in Study 1 online. Out of all the measures used in Study 1 and the online pilot, most of the results displayed no significant difference. Only two measures were found to have significantly different responses between the online and lab studies: the PPA subscale and VR experience. One reason for the significant differences in Perceived Possible Actions (PPA) is the range of simulations experienced. Whilst allowing a range of simulations was an attempt to provide a more generalisable understanding of TBW and Presence, its potential impact on PPA should be avoided. Investigating the difference between the least interactive simulation (6x9: Solitary confinement) and the most interactive simulation (Marvel: Powers United) demonstrates the range of possible actions available within the simulations. The first places an individual in a prison cell with no interaction, and the second allows individuals to select a superhero with a range of abilities to fight computer-generated characters.

In addition to the choice of simulation, the range of head-mounted displays (HMDs) used could also have been an influential factor in the differences between PPA in the online study. Although most individuals used the same headset as Study 1, several used headsets with no physical controls, such as the Google Cardboard or the Gear VR headset. These headsets

not only have less immersive qualities but also require head movements rather than physical actions to control elements of the simulation.

The second aim of the online pilot was to determine whether there was a relationship between simulation sickness, TBWs, and presence. An unexpected finding of the pilot was that no significant relationship was identified between TBW and presence. This is likely due to the limited sample size included in the pilot, which prevented regression analyses from being conducted. A larger sample size is necessary to assess the impact of TBW on presence more accurately. Although the sample size was insufficient for a robust regression analysis, correlation analyses did reveal relationships between TBW and simulation sickness. However, no relationship was found between simulation sickness and presence. These results align with findings from previous research (e.g., Sadiq and Barnett-Cowan, 2022; Grassini et al., 2021).

As temporal binding windows increased, so did the reported feelings of simulator sickness. This finding is in line with past research, which suggests that simulator sickness is caused by individuals binding incongruent information from multiple senses due to having a larger binding window (Sadiq, 2019; Sadiq & Barnett-Cowan, 2022; Zilka & Bonneh, 2022). Its demonstration over both RT and TOJ windows indicates that the underlying mechanisms for both windows contribute to this phenomenon.

The lack of a relationship between presence and simulator sickness also supports past literature on the topic. Although Weech and colleagues' (2019) review found the most support for a negative relationship between the two measures, they outlined that these negative relationships typically occur with older head-mounted displays (HMDs) rather than more modern equivalents as used by participants within this study. Studies using more modern headsets have found no correlation between simulator sickness and presence (e.g., Ryan & Griffin, 2016; Grassini et al., 2021). However, simulator sickness can differ between headsets

and simulations (Weech et al., 2019). Controlling for the content and HMD (Saredakis et al., 2020) may provide alternative results and allow for more direct comparisons of individuals' susceptibility to simulator sickness with fewer confounding factors (Tian et al., 2022) .

Based on the influence the variety of simulation and VR headsets might have had on the measures of PPA and SS, two amendments were made to the study before releasing the full version online. Individuals were limited on the headset or simulation that they could use. A pre-screen survey was distributed to individuals via Prolific, asking them which VR headset they owned and if they owned a copy of Minecraft VR or were willing to download a free trial. This helped select individuals who owned the same headset used in Study 1 (Oculus Rift) alongside the same simulation (Minecraft VR).

The second amendment allowed participants from around the world to join, rather than limiting recruitment to the UK only. This change was made to address the reduced pool of participants resulting from the specific criteria now included in the pre-screen survey. A question was added to the demographic questionnaire asking individuals to report their nationality.

No further changes were made besides these two amendments. The tasks, stimuli, questionnaires, and procedures were the same as listed in the online pilot. In alignment with the findings from Study 1, the size of an individual’s TBW is expected to influence their sense of presence while using a simulation that potentially includes conflicting cues, such as Minecraft VR. Specifically, wider windows are anticipated to lead to a reduction in the sense of presence. Additionally, based on the relationships identified in Study 2a, TBWs are predicted to affect the level of simulator sickness experienced by an individual in the opposite manner; wider windows are likely to increase simulator sickness. This suggests a correlation

may be present among the three concepts. The following hypotheses were derived from these foundations:

H4: In a virtual environment containing conflicting information (such as in Minecraft VR), individuals with a narrower TBW will report higher levels of presence.

H5: Simulator sickness will correlate with individuals' TBW and sense of presence. It is hypothesised that the level of simulator sickness experienced by individuals will be related to their TBWs, with wider TBWs associated with greater levels of simulator sickness.

4.3 Study 2b: Online Investigation of the Relationship between Temporal Binding Windows, Simulator Sickness and Presence

4.3.1 *Participants*

Nine hundred fifty-seven participants were recruited to complete the pre-screen survey through Prolific using the following criteria: they all owned a VR headset, were fluent in English, and had not completed the pilot study. Out of the 957 participants, 932 submitted complete data. Each participant was paid £0.20 for taking part. Of the 932 pre-screened participants, 331 were invited to complete the study. Of these 331, 62 participants completed the study. However, only 58 participants (38 male, mean age = 31.14 years, SD = 9.81 years, range = 18-59 years) submitted data for all sections. Each participant was paid £6.00 for participation.

4.3.2 *Materials*

4.3.2.1 **Pre-screening Survey**

Due to the recruitment site's (Prolific) selection criteria not being specific enough to the niche sample required to participate in the study, a pre-screening survey was developed to find

eligible participants. The survey was coded in Inquisit millisecond and distributed through Prolific, using the same selection criteria as in the pilot study. The pre-screen survey consisted of 4 questions checking individuals' eligibility to partake in the study. The first two questions, related to the hardware, ask: “Do you have a working Oculus Rift headset which is ready to use?" and "If the previous answer is 'Yes', please specify which Oculus Rift headset you have available. (If the answer is 'No', please specify which headset you do own under other.)". The second two related to software: "Do you have Minecraft VR ready to play on your headset?" and "If the above answer were 'No', would you be willing to download a free trial for the experiment. (Link provided through experiment). Leave blank if the above answer was 'Yes'". Only participants who had an Oculus Rift headset and had or were willing to download Minecraft VR were invited to participate in the complete study.

4.3.2.2 Simulation and Hardware

While the pilot study demonstrated results comparable to those of Study 1 with a range of VR headsets, the complete study aimed to reduce extraneous variables and provide a more direct comparison to the lab-based study. Therefore, participants were instructed to use their own Oculus Rift headset (the same hardware as Study 1) and the same simulation (Minecraft VR).

However, two minor changes from the simulation software used in Study 1 are worth noting. Firstly, all participants in the previous study used Minecraft VR version 1.7. Disabling updates in the lab allowed for a controlled software version. However, online participants in this study will have used either the most up-to-date version at the time of download or the most recently updated version on their own devices. Whilst these versions will differ slightly, the basis of the simulation remains the same. Therefore, there is no concern surrounding the impact of extraneous variables arising from the different versions.

Alongside the different software versions, the second change between the present and the first study's simulation was the absence of a world seed. As highlighted in Study 1, a world seed can be used to ensure every participant experiences the same environment. However, world seeds only generate the same environment within specific versions (each update changes the world seed formula). Therefore, as the version participants had access to could not be controlled, a world seed could not be introduced.

4.4 Analysis

4.4.1 *Participants Data Checks*

Consistent with the online pilot, participant responses to the attention checks were examined to see if any participant failed to pay adequate attention to the study. No participants were identified as being inattentive. However, 4 participant's data were removed from further analysis for not completing the tasks in line with the instructions given (e.g., using the wrong simulation).

Once the data for these participants had been removed, the remaining data was checked for response strategies. No responses were identified that would indicate a strategy had been used. The remaining 54 participants' raw data was used to calculate the widths for each individual's temporal binding window, which was in line with the analysis conducted in the online pilot.

4.4.2 *Calculating the TBWs*

Due to Inquisit updating the coding language partway through testing, data for the RT task was collected only from 15 participants. The rest of the data was complete. The binding windows were calculated in line with the procedure outlined above. The twin model (Diederich & Colonius, 2015) was used to calculate RT windows from the 15 participants' data for the RT task. Alcalá-Quintana and García-Pérez's model (2013) calculated participants' windows from

the TOJ task. Window widths were checked for outliers. No participants displayed unrealistic window widths. Descriptive statistics for the TBWs are displayed in Table 9 below.

Table 9

Descriptive Statistics for the TBW Tasks in Study 2b

Task Type (N)	Mean (ms)	Min (ms)	Max (ms)
Redundant Target (15)	53.62	16.55	206.53
Temporal Order (54)	162.54	4.16	447.42

4.4.3 *Descriptive Statistics for the Questionnaires*

The four questionnaires used in the present study were identical to the pilot. Each total questionnaire score was calculated in the same way as outlined in Section 4.2.4.3. The descriptive statistics for each questionnaire are reported below (Table 10).

Table 10

Descriptive Statistics for the Questionnaires in Study 2b

Questionnaire (N)	Mean	Standard Deviation	Minimum	Maximum
Gaming experience (54)	21.03	20.47	2.00	100.00
VR Experience (54)	55.03	61.70	0.17	333.33
SPES Total (54)	30.44	5.17	18.00	40.00
SPES SL (54)	14.63	3.03	7.00	20.00
SPES PPA (54)	15.81	2.71	10.00	20.00
SUS Total (54)	13.91	3.33	4.00	20.00
SSQ Total (54)	86.16	26.10	59.84	175.78
Nausea Score (54)	95.05	30.84	66.78	200.34
Oculomotor Score (54)	80.57	27.24	53.06	174.34
Disorientation Score (54)	139.46	46.17	97.44	264.48

Note: SL = Self-Location; PPA = Perceived Possible Actions; SSQ = Simulator Sickness Questionnaire

4.4.4 *Handling Missing Data*

Missing data can be treated in multiple ways. A common but unadvised method is to replace missing values with the mean (Field, 2009). Two alternative methods involve Bayesian

statistics: maximum likelihood estimation (MLE) and multiple imputation (White, 2009). The latter is preferred within psychological sciences due to MLE not being most suited to cope with the presence of categorical dummies, such as Gender (Enders, 2017). However, the ability to calculate missing data from multiple imputations has specific requirements.

Firstly, the data should ideally be missing at random (MAR) rather than missing completely at random (MCAR) or missing not at random (MNAR) (Little & Rubin, 2019). MAR refers to the probability that the missing data is independent of the unobserved data but dependent on the observed data. MCAR relates to the probability that the absence of data does not depend on either observed or unobserved data. Meanwhile, with MNAR, the probability of data being missing depends on the unobserved data, which is conditional to the observed data (see Little & Rubin, 2019). Due to the RT task data being missing from a software error, it is classed as MCAR.

Alongside the missing data being MCAR, the literature contains arguments regarding the maximum percentage of missing values and sample size. Most statisticians argue for less than 40% missing data and sample sizes over 100. In this study, the missing data is 72.3%, and the sample is 54.

These factors, combined with individual differences, which are the theoretical underpinning of the thesis, present the argument against using imputation for the missing data. Instead, the further analysis removed the RT TBW variable, focusing only on the TOJ TBW and the other measures.

4.4.5 *Correlations Between Predictors, Outcomes and Control Measures*

Nonparametric calculations were conducted between the presence questionnaires, TOJ TBW, simulator sickness questionnaire, and the control variables. The control variables in the online study were Age, Gender, Gaming Experience, VR Experience and Total Time within

the VR Simulation. Two control measures displayed significant correlations with either outcome measures or predictors. Age correlated with PPA (r(54) = .291, *p*=.033) and SSQ Total (r(54) = -.255, *p*=.031). VR Experience correlated with Oculomotor Score (r(54) = -.325, *p*=.016). Gender, Gaming Experience, and Time within the simulation did not correlate with any outcome or predictors. The relationship between TBW and simulator sickness was non-significant for all SSQ subscales (see Table 11 below).

Table 11
Correlations Between Predictors, Outcomes and Controls for Study 2b (N=54)

	1	2	3	4	5	6	7	8	9	10	11	12	13
1.SPES													
2.SL	**.901****												
3.PPA	**.896****	**.640****											
4.SUS	**.703****	**.676****	**.574****										
5.TOJ TBW	-.171	-.159	-.182	**-.330****									
6.SSQ Total	-.092	-.048	-.152	.052	.101								
7.Nausea	-.061	-.060	-.083	.068	.099	**.892****							
8.Oculomotor	**-.244***	-.192	**-.287***	-.107	.087	**.913****	**.730****						
9.Disorientation	-.080	-.059	-.126	.089	.097	**.933****	**.826****	**.831****					
10.Age	.161	-.009	**.291***	.131	-.198	**-.255***	-.181	-.204	-.177				
11.Gender	-.044	-.038	-.022	-.006	.043	-.140	-.168	-.147	-.202	-.054			
12.Time	.008	.060	-.093	-.111	.034	.124	.151	.105	.090	.080	.035		
13.Game Exp	-.133	-.060	-.168	-.134	.065	-.038	-.021	-.051	-.078	**-.601****	.129	-.056	
14.VR Exp	-.053	-.151	.090	-.120	.207	-.215	-.096	**-.325****	-.183	-.035	**.333****	-.072	**.244***

Note: RT = Redundant Target, TBW = Temporal Binding Window, TOJ = Temporal Order Judgement, SL = Self-Location, PPA = Perceived Possible Actions, *. Correlation is significant at the 0.05 level (1-tailed). **. Correlation is significant at the 0.01 level (1-tailed).

4.4.6 *Regression Analysis*

Two multiple regression analyses were conducted to investigate the relationship between the TOJ temporal binding window and self-reported presence scales. Backwards entry was used due to the uncertainty of including the SSQ subscales within the analysis. Both

control variables (Age and VR Experience), previously identified to correlate with TBW, simulator sickness or presence, were included in the analysis as covariates. Whilst the regression for the SPES was non-significant ($F(2,51)= 3.144$, $p=.052$), a significant regression was identified for the SUS ($F(3,50)= 4.468$, $p=.007$). Participants predicted SUS score was equal to 14.734 - .007 (TOJ width) - .081 (Oculomotor discomfort Score) + .049 (Disorientation Score), where TOJ width is measured in milliseconds. Participants self-reported SUS score decreased by 0.007 with every millisecond increase in the TOJ width and 0.081 with an increase of 1 on the oculomotor discomfort score. However, the SUS score increased by 0.049, with an increase of 1 on the participants' disorientation score. All three predictors were significant.

4.4.7 *Study 2b Discussion*

Study 2b made certain adjustments from the pilot to implement the full online replication and expansion of Study 1. These included a pre-screening questionnaire, a larger sample, and stricter HMD and virtual reality environment (VRE) guidance. The procedure was almost identical to that conducted in Study 1, with three main adjustments: Participants' TBWs were only estimated through the RT and TOJ tasks; individuals' self-reported presence was captured via the SPES (Hartmann et al., 2016) and the SUS, rather than IPQ (Slater et al., 1994); and alongside the other questionnaires a measure of simulator sickness was also taken (SSQ: Kennedy et al., 1993). The TBWs, as estimated from the TOJ task, were a significant predictor of presence scores measured by the SUS (Slater et al., 1994). The direction of the findings was the same as in Study 1, with narrower windows predicting higher levels of presence. Therefore, the fourth hypothesis (H4: *TBWs would be a significant predictor of presence*) is largely supported, in line with the overarching thesis.

The results pertaining to the fifth hypothesis (H5: *Simulator sickness will be correlated to individuals' TBW and sense of presence*) were less conclusive. Correlations between TOJ-based TBWs and simulator sickness were positive but did not reach statistical significance. Simulator sickness, however, emerged as a significant predictor of presence scores on the SUS, with specific components of simulator sickness demonstrating mixed relationships with presence. The oculomotor subscale (e.g., eyestrain and difficulty focusing) was a negative predictor of presence, while the disorientation subscale (e.g., dizziness and vertigo) positively predicted presence. These findings reflect the multidimensional nature of simulator sickness and its complex interactions with presence, consistent with prior research (Weech et al., 2019; Kim et al., 2005; Wilson et al., 1997).

4.5 Discussion

The studies presented within this chapter aimed to replicate and expand on the findings from Study 1. Specifically, they evaluated whether an online replication could produce comparable results and further investigated the relationships between temporal binding windows (TBWs), presence, and simulator sickness.

The results from the pilot study (Study 2a) indicated that online replication was feasible, as there were no significant group-level differences between online and in-person testing in TBWs, presence, or simulator sickness. The only exceptions were related to VR experience and the PPA subscale. Overall, these findings support the conclusion that online methods produce results comparable to those obtained through lab-based testing. The study failed to reject the null hypothesis of H1, which stated that significant differences would be observed between Presence, RT TBW, and TOJ TBW scores in the online pilot study and the lab study if online testing was unsuitable. The findings also lent partial support to H2, which posited that

online study participants would have significantly higher past gaming and VR experience compared to lab-based participants.

The primary aim of these studies was to investigate the relationships between TBWs, presence, and simulator sickness. The pilot study provided partial support for the hypothesis that individual TBWs correlate negatively with presence and positively with simulator sickness (H3). Positive correlations were observed between TBWs (measured through RT and TOJ tasks) and simulator sickness, particularly with the total SSQ score and oculomotor discomfort. However, presence did not show a significant correlation with TBWs or simulator sickness, which diverges from the predicted relationships. The lack of correlation between simulator sickness and presence aligns with previous research, which indicates that the relationship between these two variables is complex (Weech et al., 2019). Additionally, limitations such as variability in hardware and environments may have contributed to the challenges in obtaining replicable results between TBWs and presence. The differences introduced by various headsets and simulations could be a factor in this null finding, consistent with earlier studies (Angelov et al., 2020; Riva et al., 2007).

Adjustments made in the full study (Study 2b) to improve methodological rigour, such as pre-screening participants, using a larger sample, and standardizing VR equipment and simulation, helped refine the investigation of these relationships. The TOJ-based TBWs were significant predictors of presence as measured by the SUS, supporting H4. Narrower TBWs predicted higher presence levels, which aligns with the results of Study 1 and fits within the broader theoretical framework of this thesis.

The relationship between simulator sickness and TBWs/presence proved more challenging to interpret (H5: *Simulator sickness will be correlated to individuals' TBW and sense of presence*). In line with past research (Sadiq & Barnett-Cowan, 2022; Zilka & Bonneh,

2022), wider TBWs correlated positively (.087-.101) with simulator sickness in the pilot study, though these correlations did not reach statistical significance in the full study. Importantly, simulator sickness emerged as a statistically significant predictor of presence scores on the SUS, demonstrating nuanced relationships between specific components of simulator sickness and feelings of presence. The oculomotor subscale negatively predicted presence, aligning with Wilson et al. (1997), while the disorientation subscale positively predicted presence, consistent with Kim et al. (2005).

4.5.1 *TOJ TBW, Presence, and SS*

Although simulator sickness did not correlate with TOJ TBWs in this study, it was found to positively predict presence scores related to disorientation and negatively predict them regarding oculomotor discomfort. These findings support past research that found evidence for both relationships between simulator sickness and presence (Ling et al., 2013; Witmer et al., 1996; Wilson et al., 1997; Kim et al., 2005). However, they conflict with the hypothesis that more modern HMDs, such as the Oculus Rift, would no longer display a negative relationship (Saredakis et al., 2020; Weech et al., 2019). This contradiction suggests that, despite advancements in HMD technology, simulator sickness continues to influence presence in virtual environments.

An alternative explanation is that sensory conflict, rather than just HMD quality, plays a significant role in these outcomes. When incongruent information is bound together from multiple sensory streams (e.g., visual vection in a VR environment and vestibular cues), individuals may experience simulator sickness regardless of the immersive quality of the headset. This supports recent reviews which suggest that content and individual differences may be more influential than HMD technology in shaping presence and simulator sickness (Saredakis et al., 2020; Tian et al., 2022). For instance, in Minecraft VR, content that creates

incongruence between visual vection and vestibular cues can lead to varying degrees of simulator sickness. In this environment, participants remain stationary but use a handheld controller to navigate through the virtual world, causing a sensory conflict. Despite the same environmental setup, participants display a wide range of simulator sickness and presence scores. This variability underscores the role of individual differences—such as temporal binding windows—in influencing how individuals react to the same virtual content. Thus, both individual factors and the content's sensory cues contribute to the differing experiences of presence and discomfort, regardless of the technology used.

Moreover, the regression analysis in Study 2b revealed that increasing TOJ TBW and oculomotor discomfort were associated with reduced presence scores. This result supports the idea that presence will likely be diminished if conflicting cues are bound into a single percept/reference frame (Lohmann & Butz, 2017; Wirth et al., 2007). This is consistent with sensory conflict theory (Reason & Brand, 1975) and the Process Model of Presence (Wirth et al., 2007), both of which suggest that such conflicts are likely to reduce presence in virtual environments.

However, one issue with concluding that individuals with wider temporal binding windows are more likely to bind conflicting cues in virtual environments is the lack of a significant relationship between TBW (as estimated from the TOJ task) and simulator sickness in the data. Although the correlation was in the expected direction (Sadiq, 2019), two factors may have contributed to this lack of statistical significance. Firstly, the covariates identified—age and VR experience—may have played a role. In the present study, the mean past VR experience was 55 minutes. This is compared to an average 20 minutes past VR experience in Study 1. Prior VR experience has been highlighted to reduce simulator sickness in previous studies (e.g. Howard & Van Zandt, 2021) and it displayed a significant negative relationship with oculomotor discomfort in the study above ($r(54) = -.325$, $p=.016$). In addition, age has

been argued to correlate negatively with simulator sickness, with symptoms reducing in adults over 35 (Sardakis et al., 2020). In the present study, age ranged from 18-59, with a mean of 31, compared to study 1, which had a range of 18-32, with a mean of 21.5. Furthermore, age did negatively correlate with total SSQ score ($r(54) = -.255$, $p=.031$). By accounting for these covariates, their effects were removed from the analysis, which may have attenuated the relationship between TBW and simulator sickness.

On the other hand, a relationship between the TOJ TBW and simulator sickness might not have been identified because of the level at which both phenomena operate. TBWs estimated from TOJ tasks involve a higher-order feedback process to judge a forced choice – in this case, determining which of the audio or visual stimuli occurred first (Colonius & Diederich, 2018). If simulator sickness is indeed linked to sensory conflict, as opposed to postural sway (Lawson et al., 2021; Reason & Brand, 1975), then no higher-order process is necessary. This is because it is a physiological phenomenon that pertains to how the vestibular network interacts with visual information (Fetsch et al., 2009). Consequently, it operates at a lower level, akin to responses displaying multisensory facilitation in the RT task (Stevenson & Wallace, 2013). It is possible that had the RT data been recorded, the TBWs derived from the task would have been more likely to be associated with simulator sickness. This is because both processes are more concerned with sensory-level multisensory integration (Ferrè et al., 2015; Miller, 1982). Although both tasks measure an individual's TBW, it is acknowledged that they estimate different aspects of the same phenomenon (Colonius & Diederich, 2018). Thus, TBWs derived from TOJ tasks can predict presence in the same direction as TBW estimates from the RT task. Nevertheless, they still do not predict presence, as reported using the SPES.

4.5.2 *Limitations and Issues with Online Testing*

The shift from lab-based testing to online testing brought several significant limitations that impacted the studies discussed in this chapter. One major challenge was the decreased control over experimental conditions that comes with remote testing. In face-to-face settings, it is easier to minimize distractions and environmental factors. However, online participants completed tasks in uncontrolled environments, which could introduce variability that might affect outcomes such as presence and simulator sickness.

Additionally, technical issues with data collection software led to missing data for critical variables in the second study. This limitation was only identified after a significant portion of the data had been collected, reducing the statistical power of key analyses and limited the conclusions that could be drawn. While the speed of data collection offered by online testing is beneficial, it also raises the risk that undetected errors could affect a larger sample before any corrective actions can be taken.

The limitations mentioned highlight the difficulties of conducting rigorous experimental research in online environments. While online testing was essential for this phase of the research, the findings emphasize the advantages of in-person testing. In-person settings allow for better control over experimental conditions and enable researchers to identify and resolve procedural issues in real time.

4.5.3 *Summary and Conclusions*

The studies presented in this chapter aimed to further understand individual differences in presence by investigating a phenomenon (simulator sickness) that has been identified to relate to both presence (see Weech et al., 2019) and TBWs (Sadiq, 2019; Sadiq & Barnett-Cowan, 2022; Zilka & Bonneh, 2022). Two flash-bleep tasks were converted to be distributed to participants online: a redundant target (RT) task (Todd, 1912) and a temporal order judgment (TOJ) task (Bald et al., 1942). Participants then completed a virtual reality experience on their

own devices before responding to questionnaires regarding their experience. The questionnaires captured the positive user experience of presence (SPES: Hartmann et al., 2016) and the negative aspect of simulator sickness (SSQ: Kennedy et al., 1993). Capturing both elements of the user experience allowed the relationships between the three concepts to be explored. These analyses are based on the theoretical proposal that individuals with a wider temporal binding window are more likely to combine conflicting information. This merging of information can lead to an increase in simulator sickness (Reason & Brand, 1975; Oman, 1990; Grassini et al., 2021) and can hinder the development of a sense of presence (Wirth et al., 2007).

Despite the limitations, the pilot study identified a significant positive relationship between TBWs and simulator sickness. In addition, both TOJ TBWs and oculomotor discomfort negatively predicted feelings of presence as reported on the SUS presence questionnaire. The data from Study 2b partially confirms that in environments where conflicting information is presented, such as when vection is visually perceived without corresponding vestibular input, individuals with a wider temporal binding may feel less present. This is likely due to their tendency to combine these conflicting cues.

However, Wirth and colleagues (2007) highlight the importance of multisensory consistency in their Process model of Presence. If the assumptions of the model hold true, it can be reasonably inferred that individuals with a wider tolerance for offset stimuli, or a broader "TBW," are more likely to experience a greater sense of presence compared to those with a narrower tolerance. This could be due to individuals with a wider TBW being better at maintaining multisensory consistency. Interestingly, the previous studies within this thesis have shown conflicting results, with individuals possessing a narrow window reporting a stronger sense of presence. However, it is important to note that these studies utilised virtual reality environments (VREs) with conflicting cues, which inherently disrupt multisensory

consistency. The findings suggest that there may be two distinct processes at play. Some individuals can effectively process conflicting cues by separating them into distinct egocentric reference frames (Ehrenfeld & Butz, 2013) and selecting the most salient reference frame (Riecke & Von der Hyde, 2003). While others may tolerate offset information related to the same event (Lohmann & Butz, 2017). If these assumptions are correct, then when a VRE provides multisensory stimuli regarding events and does not contain conflicting information, individuals with a wider window should feel more present. In addition, if an individual's ability to accurately bind congruent sensory information is related to presence, it should only occur when individuals experience a multisensory environment. When individuals are presented with a unisensory experience, there is no need to bind multisensory information. Therefore, the relationship between an individual's binding window and presence should no longer occur.

Chapter 5: Moderating Temporal Binding within Uni v Multisensory Environments (Study 3)

5.1 Introduction

When individuals are presented with sensory cues regarding a virtual environment, they develop a mental model of that environment (e.g. Slater & Steed, 2000; Slater et al., 2022; Wirth et al., 2007). This mental model is formed of a primary egocentric reference frame (PERF) built on the sensory information the environment provides (Riecke & Von der Heyde, 2003). If this mental model passes a perceptual hypothesis, which tests whether the environment is most likely to be where the individual is situated, they experience presence. These stages are the fundamental structure of the process model (Wirth et al., 2007). However, the elements apply to multiple proposals of how presence is formed (e.g. Biocca, 2003; Nunez, 2007; Schubert et al., 1999b; Slater et al., 2022; Slater & Steed, 2000). Regardless of the model of presence, one key underlying principle is repeated: when the environment provides multisensory information, it must be consistent, because a lack of multisensory consistency diminishes presence (e.g. Slater et al., 1994; Steuer, 1992).

Consistent multisensory information can be viewed in at least two ways: synchronous or congruent. Synchronous multisensory stimuli refer to cues provided by an object which relate to the same event but are perceived by different modalities (Colonius & Diederich, 2004; Diederich & Colonius, 2009; Ernst, 2007; Pouget et al., 2002; Stevenson et al., 2013; van Wassenhove et al., 2007; Wallace & Stevenson, 2014). A typical example used in this thesis is of an individual talking. The movement of a person's lips provides visual cues. At the same time, their mouth produces the audio cues, and whilst one is detected by the eyes and the other by the ears, they are perceived as a single synchronous event.

Congruent cues, however, relate to multisensory cues from two temporally aligned events, which do not cause multisensory inhibition if they are bound together (e.g. Van der Burg et al., 2008). One example is walking in the natural world, whilst a visual overlay depicts walking in a virtual world. Although these events are separate and involve different sensory cues, if the stimuli are presented within an individual's temporal binding window, an illusory conjunction of the two cues might occur (e.g. McGurk & MacDonald, 1976; Van der Burg et al., 2008). In this case, the event can be considered congruent; conversely, if the cues did not align, they would be incongruent (conflicting).

In virtual reality (VR), consistent multisensory information can either be provided by the technology or perceived by the user. It is essential for both the equipment and the design of the VRE to deliver sensory information in a congruent and synchronised manner. This helps create a unified experience that enhances feelings of presence and reduces sensory conflicts, which can lead to simulator sickness. However, current VR technology is unable to produce truly synchronous multisensory stimuli (Warburton et al., 2022). Research indicates that small offsets in stimulus presentation (approximately 20 ms) are present in various VR headsets (Tachibana, 2022).

Furthermore, certain VRE designs incorporate conflicting cues, as seen in the VRE used in Studies 1 and 2b. These conflicting cues arise from the navigation techniques required to explore the VRE, leading to incongruence. Despite the lack of synchrony and the existence of conflicting cues, individuals often report feeling located within these VREs. However, there is variability among individuals, as demonstrated by the results of Studies 1 and 2. Therefore, understanding individual differences in the perception of consistent multisensory information—particularly through TBWs—contributes to our knowledge of how feelings of presence and simulator sickness may develop.

The results from Studies 1 and 2 demonstrated that individuals' temporal binding windows are a significant predictor of self-reported presence in VR. Individuals with narrower TBWs reported higher levels of presence, which suggests they were more likely to separate multisensory events into unisensory occurrences if offsets were too great, as may be the case in VR. One possible explanation could be that there is no need to maintain multisensory consistency but to develop prominent egocentric reference frames from these unisensory representations; however, it is unlikely that this is the case. Multisensory VREs have been demonstrated to generate a larger sense of presence than unisensory VREs (e.g. Cooper et al., 2018; Dinh et al., 1999; Slater et al., 1995). Instead, a dual process explanation is favoured. An alternative hypothesis is that two processes occur when an individual is presented with a multisensory VRE: one process of separating conflicting cues and the other of maintaining asynchronous (but related) cues.

The first process relates to avoiding the negative elements of sensory conflict. Sensory conflict can occur across multiple senses (Grassini et al., 2021) but primarily occurs between visual and vestibular cues in the context of head-mounted displays (HMDs) (e.g. Bonato et al., 2008; Kennedy et al., 2010; Reason & Brand 1975). Narrow temporal binding windows could prevent the sensory representations from being combined, thus reducing sensory conflict. This aligns with the findings from the previous two studies, which involved an environment which contained a degree of sensory conflict. When these sensory conflicts are removed (e.g., in room-scale environments), the relationship between the TBW width and presence may show the opposite pattern. This is the second process of maintaining asynchronous but related multisensory cues. Wider windows would be associated with maintaining the unified perception of multisensory events if the cues are not presented synchronously. This can be demonstrated if the relationship between window width and presence increases in a multisensory VRE but not in a unisensory environment.

This hypothesis is developed from three premises derived from the literature. Firstly, as highlighted above, most theories of presence argue that the consistency of multisensory information is important for the development of presence (e.g. Slater et al., 2022; Wirth et al., 2007). This is supported by studies which demonstrate an increase in presence with an increase in the number of sensory modalities stimulated (Cooper et al., 2018; Dinh et al., 1999; Grassini et al., 2021). Secondly, whilst the technology aims to present this multisensory information consistently, it has been documented that lag and latency within the technology are still relatively high when it comes to millisecond accuracy (e.g. Warburton et al., 2022). The limitations of the technology mean that some of the ability to maintain the consistency of information must be due to an individual's perception. Finally, as previously discussed, individuals with a larger TBW are less accurate at perceiving delayed information as asynchronous (e.g. Colonius & Diederich, 2018). Instead, these individuals will likely combine information into a single event (Calvert & Thesen, 2004; Stein & Meredith, 1993). When the information should not be combined, this leads to adverse outcomes, such as environments containing conflicting cues. However, when the information should be combined, these individuals may benefit from having a less accurate perception of the delay between stimuli. The following study aims to investigate if this second process, the ability to maintain the consistency of multisensory information when no conflicting information is present, exists.

5.1.1 *Presence*

The past two studies have identified that different measures of the temporal binding window are related to different presence measures. The Spatial Presence Experience Scale (SPES) and Igroup Presence Questionnaire (IPQ) have been predicted by TBWs estimated from the RT task (Chapter 3). In comparison, the Slater, Usoh and Steed (SUS) presence questionnaire was predicted by TBWs estimated from the TOJ tasks (Chapter 4). Correlations between the SPES and SUS demonstrate that although they share construct validity, they are

not perfectly correlated. The differences between the two presence measures may explain why the TBWs related differently to these questionnaires. The 3-item SUS is a simplistic measure of general presence, focusing on 'transportation' rather than on embodied presence (Hein et al., 2018; Youngblut, 2003). In comparison, the SPES includes both transportation and embodiment of perceived possible actions (Hartmann et al., 2016). TOJ TBWs are based on the ability to accurately discern whether the information is congruent or not, which can be considered a top-down process or higher order (Eskes et al., 2007). On the other hand, RT TBWs are calculated from the multisensory gain that binding different sensory streams provide, which can be considered a bottom-up process (Miller & Schwarz, 2006). The difference between the bottom-up RT TBWs and top-down TOJ TBWs highlights the benefit of continuing to use multiple questionnaires to assess the complex phenomena of presence.

5.1.2 *Simulator Sickness*

Study 2 investigated whether temporal binding windows, presence and simulator sickness were related. The results provided some evidence to suggest that they were. TBWs estimated from the RT task, correlated with multiple elements of the simulator sickness questionnaire (SSQ). In addition, simulator sickness was a significant predictor of presence on the SUS. Simulator sickness is thought to be caused by conflicting sensory inputs (Reason & Brand, 1975). When individuals integrate conflicting cues into a singular percept, they are more likely to experience the negative effects associated with simulator sickness. Individuals with a wider window are likelier to bind inconsistent multisensory information than those with a narrow one. Study 2 identified a positive relationship between TBWs and simulator sickness, supporting the notion that individuals with a wider window are more likely to bind conflicting sensory inputs compared to those with a narrower window. However, as highlighted above, it is important to note the difference between asynchronous and incongruent (conflicting) information (Su, 2014). Conflicting information relates to sensory information from two

modalities simultaneously representing different events (e.g. McGurk & MacDonald, 1976). For example, if the vestibular system receives information of being sat stationary, whilst the visual system receives information indicating vection (movement), the two systems are likely to be in conflict. On the other hand, asynchronous information relates to sensory information from two modalities representing the same event but containing a delay (Ernst, 2007; Pouget et al., 2002). For example, a screen produces the image of someone talking, yet the audio follows 20 milliseconds later through headphones. Of the two inconsistent multisensory experiences, simulator sickness is only likely to occur in the first: conflicting cues. More so, an individual with a wider temporal binding window will likely experience these symptoms more significantly because they integrate a wider range of conflicting cues (Sadiq & Barnett-Cowan, 2022). The results of Study 2a support this hypothesis, as the relationship between the SSQ and TBWs was identified to be positive. By removing conflicting cues, simulator sickness should be reduced, regardless of an individual's TBW size. Therefore, the SSQ was retained for measuring simulator sickness in this study, with the expectation that it would not correlate with TBWs in a room-scale environment that lacks conflicting cues.

5.1.3 *Current Study: Aims and Hypothesis*

As highlighted above, TBWs are theorised to have a dual role in virtual environments by separating conflicting cues and maintaining asynchronous multisensory information. For individuals to experience a sense of presence, they need to integrate sensory inputs to create a coherent experience. This suggests that individuals with a larger TBW would be more successful in maintaining a sense of presence. On the contrary, previous studies have demonstrated that individuals with a narrower TBW report higher levels of presence. This phenomena is hypothesised to occur due to conflicting cues being present within the previous VREs. The study presented in this chapter aimed to test the relationship between TBW and presence in an environment that did not contain conflicting cues. This approach would allow

for a direct examination of the hypothesized secondary role of TBW in the formation of presence: the maintenance of asynchronous multisensory information. Based on this premise, the first hypothesis for this study was formulated:

H1: By removing conflicting cues, individuals' TBW would positively predict their sense of presence in a multisensory space. There will be no relationship between simulator sickness and either TBW or presence.

To test this hypothesis, a VRE devoid of conflicting cues was necessary. Previous studies utilized Minecraft VR, which, despite allowing for essential elements such as interaction and navigation, introduced conflicting cues through joystick-controlled locomotion. Interaction within the environment is crucial because perceived possible actions significantly enhance the feeling of being situated within a VRE (Glenburg, 1997; Schubert et al., 1999b; Wirth et al., 2007). In addition to interaction, the ability to navigate through a virtual world plays an essential role in developing mental models of that space (e.g., Gyselinck et al., 2006). Minecraft VR provided six degrees of freedom (6-DoF): Forward/Back, Left/Right, Up/Down, Roll, Pitch, and Yaw. This means that individuals can move and rotate freely across three axes (Aldoma et al., 2012). To ensure comparability with prior VREs, the environment used in this study also needed to incorporate similar elements. Thus, a room-scale VRE—'Job Simulator' (Owlchemy Labs, 2019)—was selected for the current study. In this environment, participants navigate a virtual office cubicle that mirrors the physical lab space, ensuring that movement is consistent between the virtual and physical worlds and eliminating visual-vestibular conflicts. The inclusion of interactive elements within this confined space ensures ecological validity and maintains a sense of embodied presence, providing an optimal setup to test the second role of the TBW in the experience of presence.

Besides removing conflicting cues, the present study also aimed to investigate a confirmatory hypothesis regarding the relationship between TBWs and presence. Up to this point, the relationship has only been explored through correlations and coefficients in multisensory experiences. However, if an individual's ability to bind sensory information is important for feeling present in VR, it should only have a positive relationship in multisensory experiences. This is because, in unisensory experiences, all individuals should experience the same degree of VR-related information. In these unisensory experiences, there is no multisensory facilitation, no maintenance of offset stimuli and the egocentric reference frame that is developed, will only be of one modality from the virtual world. It is acknowledged that removing all sensory information besides one modality, is a near-impossible task. Even when only visual information is displayed to an individual from a VR world, other senses are likely still present - the smell of the laboratory, the internal sense of the room’s temperature, etc. However, research has shown that the three most important senses within VR are visual, audio, and haptic feedback (Cooper et al., 2018). Therefore, turning off haptic feedback in both conditions can assess a more direct comparison between multisensory (Audio-Visual) and unisensory (Visual) VREs. The following hypothesis is generated based on this manipulation:

H2: The relationship between TBWs and presence will be moderated between multisensory and unisensory VR environments. It is hypothesised that if there is a positive relationship between TBW and Presence, it should only be found within the multisensory condition.

5.2 Methods

5.2.1 *Participants*

Sixty-eight participants (47 females, mean age 21.38, SD 7.40) were recruited through a course participation scheme at York St John University, and all were awarded course credits

for taking part. All participants had normal hearing and normal or corrected-to-normal vision and were fluent in English. Each participant was randomly assigned to one of two groups. '*Sound-Present*', which experienced a multi-sensory version of the VRE with both audio and visual information (34 participants). The other group, '*Sound-Absent*', experienced a unisensory version of the VRE with only visual information (34 participants).

An a priori power analysis was conducted using GPower (version 3.1) to determine the minimum sample size required to detect a large moderation effect ($f^2 = 0.35$) with 95% power and an alpha level of .05. The analysis was based on a multiple regression model including five predictors: a dummy-coded condition variable, a continuous predictor, their interaction term, and two covariates. The power calculation was performed using the linear multiple regression: fixed model, R^2 increase option in GPower, testing the unique contribution of the interaction term above the other predictors. The analysis indicated a required sample size of 63 participants. The final sample met this threshold, providing sufficient power to detect large moderation effects.

5.2.2 *Design*

The main measures within this study were identical to those presented in Study 2, except participants' temporal binding windows were measured offline (in person). The '*Sound-Present'* participants experienced the VRE with audio and visual information. The other group, '*Sound-Absent,*' experienced the VRE with only visual information. In both groups, haptic feedback was turned off to reduce additional sensory input.

5.2.3 *Materials*

5.2.3.1 **Temporal Binding Tasks**

Individuals' TBWs were measured using the same RT and TOJ tasks used in the previous studies. Like in the second study, the tasks were coded using Millisecond Inquisit.

However, unlike the online experiment, these tasks were run offline to prevent software updates from affecting the code within the experimental scripts.

Due to this study taking place within a lab under the supervision of a researcher, each participant completed the task in the same manner. The visual stimuli for each task consisted of a 5x5(cm) black and white checkerboard, which appeared in front of a fixation cross on a black background (the same as in study 1). This was presented approximately 60 cm in front of the participants on a Lenovo Legion Y25f-10 monitor with a 62.23cm screen size and a 144Hz refresh rate. The auditory stimuli were a 440HZ pure tone with a 33ms duration presented binaurally via Sony MDR-ZX110 headphones. The reminders for which keys to press were presented at the top of the screen. All other task elements were identical to those presented in the online study.

5.2.3.2 **Questionnaires**

While the online study did not establish a relationship between TOJ TBW and the SPES, the initial study identified a connection between the SPES and the RT TBW. The initial finding suggests that if this measure had been included in the online study, it could have emerged as a significant predictor of presence. These findings, coupled with the SPES supporting the presence theory being examined in this thesis, influenced the decision to retain this measure from the previous study.

The Slater, Usoh and Steed presence questionnaire (SUS) was found to be significantly predicted by individuals' TOJ, and two measures from the Simulator Sickness Questionnaire (SSQ): Oculomotor Discomfort and Disorientation. Moreover, it demonstrated satisfactory convergent validity with the SPES in both the online pilot and the complete study. As a result, the SUS was retained as the second measure of presence in this research.

The relationship between simulator sickness, TBW, and presence continues to be a central focus of this research. This study's design aims to eliminate the impact of simulator sickness by removing conflicting cues. The previous study identified a relationship between the SSQ and one of the presence measures (SUS). Therefore, including the SSQ in this study is important to check that a similar relationship is not observed.

The final questionnaire utilised in this study was identical to the fourth questionnaire from the online study. While the gaming experience questionnaire did not correlate with either presence measure in the last two studies, VR experience correlated with oculomotor discomfort in the online study. This correlation occurred in a sample where participants' average duration of VR use was significantly higher than in the first study. Additionally, various studies (e.g., De Leo et al., 2014; Di Luzio et al., 2021; Donohue et al., 2010) have indicated that gaming experience might influence self-reported presence. Therefore, it was important to include the measure to continue to account for samples with different levels of average VR or gaming experiences.

5.2.3.3 **VR Hardware and Simulation**

The equipment used to elicit a sense of presence in this study was the Oculus Quest (Figure 8). A standalone head-mounted display (HMD) with a resolution of 1440 x 1600 pixels per eye, a 93˚ field of view, and a 72Hz refresh rate. The HMD provided head and position tracking through an “inside-out” tracking system. The HMD uses a Qualcomm Snapdragon 835 ‘System-on-Chip’ with 4GB of RAM and an Adreno 540 GPU. In the virtual environment, participants can control movement and interactions using the Oculus Quest wireless controllers. These are tracked using a simultaneous localisation and mapping system using AI, accelerometers, and the built-in camera to track infrared diodes on the controllers.

Figure 8
Oculus Quest Headset and handheld controllers used in Study 3

The virtual reality simulation was changed for this study for two reasons. Firstly, Minecraft© VR involves a virtual character moving large distances within a virtual environment, while the individual remains stationary. This clash in proprioception could potentially heighten feelings of simulator sickness. In keeping with this thesis's focus on Audio-Visual sensory binding, a more stationary alternative was considered.

The simulation chosen for this study was 'Job Simulator: The 2050 Archives', a video game developed by Owlchemy Labs. The simulation involves participants completing simulated jobs selected from a museum run by robots. For this study, participants completed the 'Office Worker' simulation (see Figure 9). Throughout the simulation, participants are instructed to complete various tasks using their handheld controllers, whilst confined to an office cubicle that matched the dimensions of the lab room. Tasks included, plugging in the computer, looking through emails, and selecting new robots to hire. One group of participants completed the simulation with both audio and visual information regarding the virtual environment displayed. The auditory information that was presented contained background office noise as well as task-specific sounds, such as the sound of the keys being pressed on the

keyboard. The other group only received visual information, with instructions displayed via subtitles. In both groups, haptic feedback from the controls was disabled.

Figure 9
Image taken from 'Office Worker' in Job Simulator: The 2050 Archives illustrating the office cubicle VRE

5.2.4 *Procedure*

Each participant was randomly assigned to the '*Sound-Present*' or '*Sound-Absent*' group and completed the study in a 2-hour session. Participants were seated at a computer, where they filled out a consent form and signed a safety information sheet detailing the use of the VR headset in line with university procedure. Afterwards, they filled in three open text boxes relating to demographic questions, asking them to report their age, gender, and nationality. Each individual within the two groups completed the computerised TBW tasks in a counterbalanced order. For example, participant one in the '*Sound-Present*' group completed the RT task first, followed by the TOJ task. In contrast, participant 2 completed the TOJ task first, followed by the RT. The same was true for the '*Sound-Absent*' group. To minimise any influence or distraction, the researcher left the room while participants worked on the tasks.

Following the temporal binding tasks, participants engaged in a virtual reality simulation. Researchers introduced each piece of equipment: headset and controllers. Participants were instructed on how to use the controls for the virtual hand movements. At this point, individuals attached the controllers to their wrists using the safety straps and confirmed that they understood the instructions provided regarding controls by pressing the correct buttons. The researchers also provided guidance on adjusting the headset for optimal clarity. Participants in the visual-only group were instructed not to adjust the volume, a setting that the researcher verified through the Oculus app.

After being introduced to the equipment, participants donned the HMD and adjusted the straps and lenses, to ensure they had a clear view of the writing on a virtual screen. They were then asked to pick up an orange cartridge located on a floating tray next to the virtual screen and place it into the top of the display. This exercise ensured that participants were familiar with the controls before entering the actual simulation.

When ready, participants pulled a lever beside the virtual screen to start their simulation. The researcher remained present throughout the process and was aware of what was happening at all times, although they only spoke if participants asked an important question. If individuals became stuck on specific tasks, prompts were provided, such as "Have you looked under the desk?" or "Have you pressed the red button?" Participants had up to 30 minutes to explore the virtual environment. They were permitted to exit the VR equipment at any time. The duration of their time spent in the VR environment was recorded.

After the simulation, participants completed four computerised questionnaires in the following order: the SPES, the SUS, the SSQ, and the Gaming Experience Measure.

5.2.5 *Ethics*

The research followed the BPS guidelines for ethical practice, which aligned with the ethics procedure outlined in study one. However, a few additional ethical considerations were put into place, due to the timeframe of the research and the difference in simulation and hardware. Firstly, after each participant had completed the study, all equipment was sterilised using alcohol wipes. Each participant was allowed to wear a face mask. However, they were also advised that this might impact their experience within VR. No participant wore a face covering; however, the researcher did. Due to the simulation utilising a confined space, participants were not seated throughout their experience. Instead, the safety features of the technology set up a 'guardian' boundary which they cannot go beyond. This safety feature was set up approximately half a meter from each edge of the lab room, and overlapped with the boundaries of the virtual office cubicle participants were situated within. The researcher remained present throughout each simulation to ensure the 'guardian' was correctly set and individuals remained safe from harm. The York St John University School of Education, Language and Psychology Ethics Committee approved the study.

5.3 Analysis

5.3.1 *Calculating the TBW's*

Raw data from the redundant target task were checked for outliers (<30ms and >3000ms), with individual responses removed. The raw data for the TOJ task was not removed based on the reaction time to stimuli. Instead, participants' responses were reviewed, and any identified to be using a response strategy were removed from subsequent analysis (e.g., responses such as ZZZMMM or ZMZMZM). No participant was identified using such a strategy; therefore, no data was removed for the TOJ task. The window widths for each participant were then calculated using MATLAB scripts adapted from Diederich and Colonius' (2015) extended twin model. The scripts use *fmincon* to optimise the objective function:

$$\sum_{\tau}\left\{\left(\frac{mean\left[RT_{VA,\tau}\right]-E\left[\widehat{RT_{VA,\tau}}\right]}{SE\left(mean\left[RT_{VA,\tau}\right]\right)}\right)^{2}+\left(\frac{f_{VA,\tau}-\widehat{\Psi(\tau)}}{SE\left[f_{VA,\tau}\right]}\right)^{2}\right\}$$

where mean [$RT_{VA,\tau}$] and $f_{VA,\tau}$ relate to the mean reaction times and relative frequency of "visual first" judgements for each stimulus onset asynchrony (SOA) (τ) in the RT and TOJ task, respectively. $E[\widehat{RT_{VA,\tau}}]$ and $\widehat{\Psi(\tau)}$ are the predicted values for each SOA based on the twin models parameters: visual processing time (λV), auditory processing time (λA), the common processing stage (μ), the TBW for the tasks (ω), and multisensory temporal gain (Δ). Therefore, the objective function calculates the deviation between measured and estimated values, returning the fit with the smallest deviation and identifying the TBW based on the twin model calculations (see appendix 5 for full script).

5.3.2 *Descriptive Statistics*

Raw questionnaire data for the SPES, SUS, SSQ and gaming experience questionnaires were converted into total scores in the same way as in the previous study. Each measure was checked for extreme outliers (Over 2.5SD). Three participants were removed from further analysis due to being classed as extreme outliers on multiple variables. Before conducting the moderated regression analysis, descriptive statistics were calculated for the sound-present and sound-absent groups to assess the comparability of these groups on the key variables of interest: presence, simulator sickness and temporal binding windows (TBW). Table 12 presents the mean and standard deviations for these variables across the two groups.

Table 12
Descriptive Statistics for Presence, SS and TBW by Sound Condition in Study 3

		Group Means (SD)	
	Measure	*Sound-Absent (N=30)*	*Sound-Present (N=35)*
Presence	SUS	14.97 (3.32)	15.54 (3.38)
	SPES: SL	15.27 (3.26)	16.63 (2.65)

	SPES: PPA	17.53 (2.36)	16.54 (2.69)
	SPES Total	32.8 (4.92)	33.17 (4.82)
Simulator Sickness	SSQ Total	78.42 (19.46)	76.83 (18.84)
	Nausea Score	81.41 (17.31)	79.59 (15.85)
	Oculomotor Score	75.04 (22.73)	72.98 (21.91)
	Disorientation Score	127.14 (36.32)	122.10 (37.15)
Temporal Binding Window	RT	58.48 (130.39)	37.40 (28.25)
	TOJ	234.45 (194.96)	179.07 (94.73)

Note: SL = Self-Location; PPA = Perceived Possible Actions; SSQ = Simulator Sickness Questionnaire; RT = Redundant Target; TOJ = Temporal Order Judgement

5.3.3 *Between-group Homogeneity*

Mann-Whitney U tests assessed whether the two groups were comparable on all key variables. Results indicated that there was no significant difference between the groups on SUS score (U=474, p= .499), Self-Location (SL) score (U=399.5, p= .096), Perceived Possible Actions (PPA) score (U=416, p= .147), or SPES score (U=516.5, p= .911). Alongside presence scores, neither TBW measures significantly differed between groups: RT TBW (U=509, p= .833) and TOJ TBW (U=477, p=.528). The simulator sickness measures were also found to be comparable across groups with no significant differences amongst them: SSQ Total (U=519, p=.937), Nausea score (U=509, p=.826), Oculomotor score (U=508, p=0.820), and Disorientation score (U=483, p=.567). Therefore, the two groups can be assumed to have comparable baseline equivalence.

5.3.4 *Control Variables and Simulator Sickness*

The relationship between the control variables (age, gender, gaming and VR experience), TBWs, and presence was explored through nonparametric correlations, to identify if any covariate would need to be included in the moderation analysis. Neither gender nor gaming experience correlated with any primary measure. However, age and prior VR experience were identified to correlate with multiple presence measures (Table 13).

Table 13
Correlations Between Control Variables, TBW and Presence Measures in Study 3

		Control Variables			
		Age	Gender	Gaming Experience	VR Experience
Sound Absent	SPES: SL	.435**	-0.161	-0.268	-.523**
	SPES: PPA	0.026	-0.177	-0.057	-0.177
	SUS	.339*	-0.268	-0.073	-.365*
	SPES	0.287	-0.192	-0.193	-.417*
	RT TBW	-0.200	0.183	0.102	-0.077
	TOJ TBW	0.072	0.190	0.036	-0.103
Sound Present	SPES: SL	.399**	-0.098	-0.097	-0.188
	SPES: PPA	0.184	0.044	0.069	0.029
	SUS	.428**	-0.176	-0.013	0.156
	SPES	.304*	0.008	0.012	-0.084
	RT TBW	0.116	0.115	-0.106	-0.142
	TOJ TBW	-0.163	0.111	0.039	0.064

Note: RT = Redundant Target, TBW = Temporal Binding Window, TOJ = Temporal Order Judgement, SL = Self-Location, PPA = Perceived Possible Actions, *. Correlation is significant at the 0.05 level (1-tailed). **. Correlation is significant at the 0.01 level (1-tailed).

Spearman's rho correlations were conducted between the TBW measures, the Presence scores, and the simulator sickness measures to identify whether removing conflicting cues successfully removed the impact of SS. No significant correlations were identified between the simulator sickness measures and the main variables (see Table 14). The lack of significant correlation indicates that conflicting cues were successfully mitigated/reduced.

Table 14
Correlations Between Simulator Sickness, TBW and Presence Measures in Study 3

		Simulator Sickness Measures			
		SSQ Total	Nausea Score	Oculomotor Score	Disorientation Score
Sound Absent	SPES: SL	0.066	0.022	0.054	0.063
	SPES: PPA	-0.003	0.013	0.048	-0.057
	SUS	-0.008	-0.047	0.020	-0.030
	SPES	0.035	0.010	0.075	-0.013
	RT TBW	0.188	0.325	0.139	-0.040

	TOJ TBW	0.127	0.164	0.010	-0.118
Sound Present	SPES: SL	-0.199	-0.056	-0.232	-0.082
	SPES: PPA	-0.268	-0.161	-0.328	-0.132
	SUS	-0.172	-0.215	-0.115	-0.036
	SPES	-0.268	-0.100	-0.323	-0.111
	RT TBW	-0.063	0.236	-0.070	-0.192
	TOJ TBW	0.012	0.203	0.003	-0.091

Note: RT = Redundant Target, TBW = Temporal Binding Window, TOJ = Temporal Order Judgement, SL = Self-Location, PPA = Perceived Possible Actions, *. Correlation is significant at the 0.05 level (1-tailed). **. Correlation is significant at the 0.01 level (1-tailed).

5.3.5 *Moderated Multiple Regressions*

To identify whether multisensory environments moderated the relationship between an individual's TBW and their sense of presence, moderated multiple regressions were conducted for each presence measure. The assumptions of moderated multiple regression were checked. The residuals were uncorrelated and distributed normally, and there was no multicollinearity. However, the data displayed heteroscedasticity. Therefore, heteroscedasticity-consistent standard errors were used within the multiple moderated regression using the Hayes Process module (model 1). The results for each scale are discussed below.

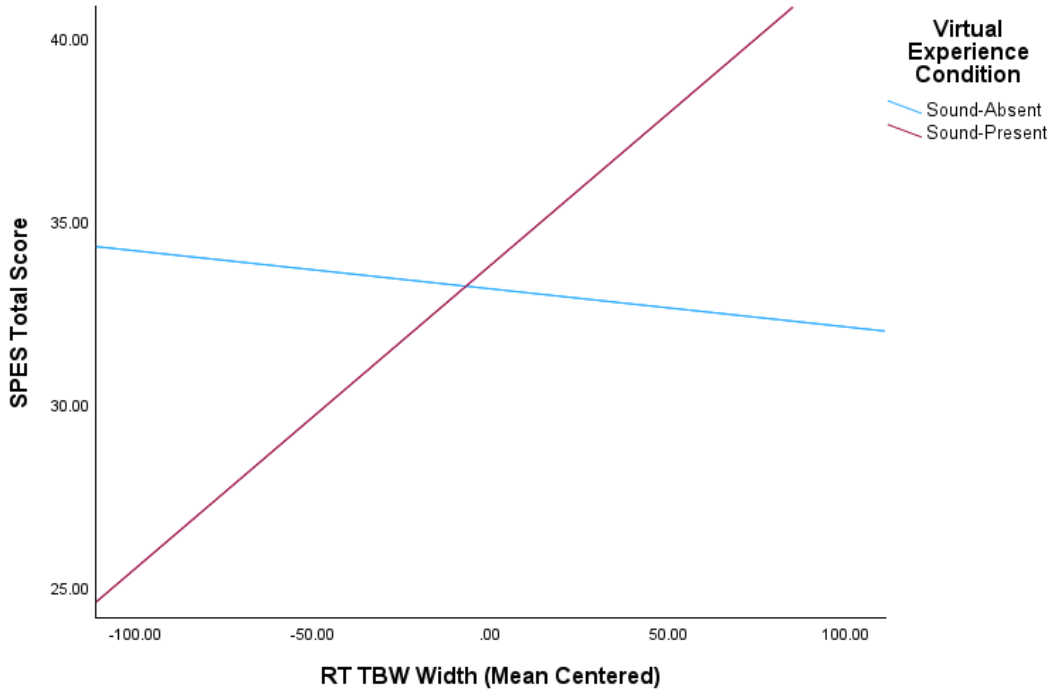
5.3.5.1 **Spatial Presence Experience Scale**

The first regression explored the relationship between individuals' TBWs as estimated from the RT task and the SPES total score, whilst controlling for age and prior VR experience. There was a significant effect of individuals' TBW on the total SPES score (b = .036, CI [.021, .051], t = 4.833, p < .001). Additionally, age was found to be a significant covariate (b = .189, CI [.124, .256], t = 5.739, p <.001), indicating that older participants tended to report higher SPES scores. Prior VR experience also significantly affected the SPES score (b = .022, CI [.008, .035], t = 3.209, p = .002), indicating that increased prior VR experience was associated with higher SPES scores.

There was not a significant main effect of sound condition on the SPES score (b = .322, CI [-.710, 1.354], t = .625, p = .535). Nonetheless, these effects were moderated by a significant interaction between sound condition and RT TBW (b = .047, CI [.032, .061], t = 6.287, p < .001). The results highlighted that as the size of individuals' TBWs increased in the sound condition, so did their SPES score (b = .083, CI [.053, .112], t = 5.6268, p < .001). In contrast, the increase in TBWs in the no sound condition had a smaller negative effect on individuals' SPES score (b = -0.010, CI [-.015, -.006], t = -4.404, p < .001). This indicates that the presence of sound moderates the relationship between presence and individuals' binding of multisensory information (Figure 10).

Figure 10

Interaction effect of the presence of sound on the relationship between individuals' temporal binding window and self-reported presence score on the spatial presence experience scale



The second regression explored the relationship between individuals' temporal binding windows (TBWs) from the TOJ task and the SPES total score, whilst controlling for age and prior VR experience. The results of the multiple regression did not display a significant effect

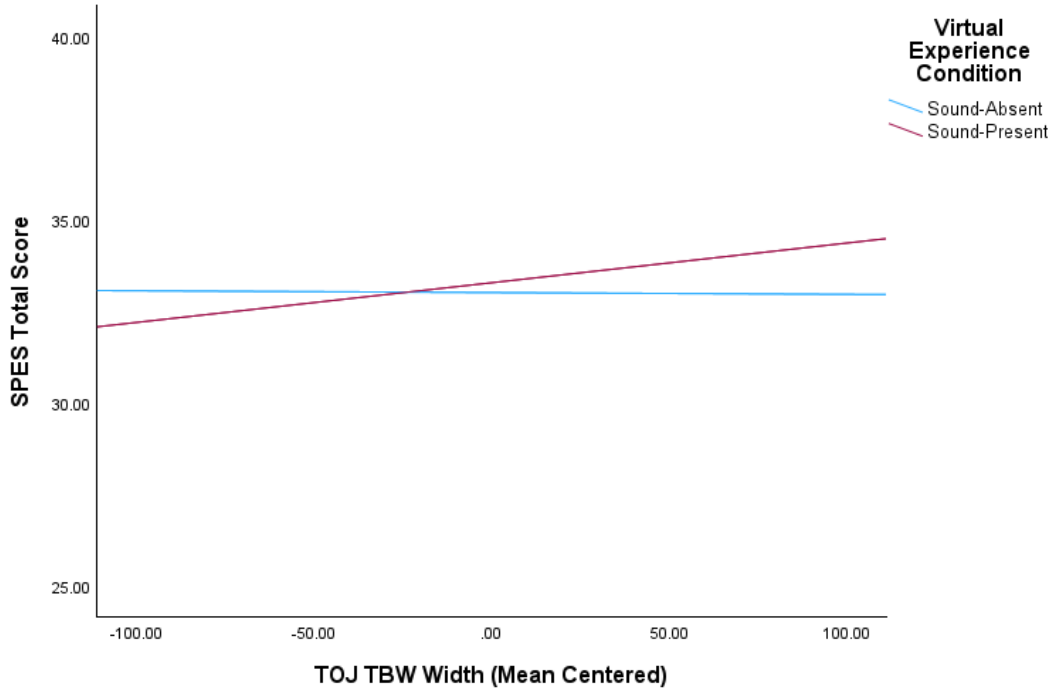
of individuals' TOJ TBW on the SPES score (b = .005, CI [-.003, .014], t = 1.242, p = .219), failing to find a relationship between window width and SPES Score. Similarly, the sound condition did not report a significant effect on the SPES score (b = .134, CI [-1.03, 1.30], t = .230, p = .819), indicating that the addition of sound did not affect individuals' sense of presence.

However, both covariates were significant predictors of the SPES score. Age had a significant positive effect on the SPES score (b = .192, CI [.121, .264], t = 5.391, p < .001), indicating that older participants tended to report higher levels of presence. Prior VR experience was also a significant predictor (b = .015, CI [.000, .030], t = 2.065, p = .043), suggesting that participants with more prior VR experience reported slightly higher presence scores.

The moderation effect between the sound condition and TOJ TBW was not significant (b = .006, CI [-.003, .014], t = 1.334, p = .184), meaning that neither TOJ TBW nor the sound condition significantly affected presence as measured by the SPES (Figure 11).

Figure 11

Interaction effect of the presence of sound on the relationship between individuals' temporal binding window and self-reported presence score on the spatial presence experience scale



5.3.5.2 Slater, Usoh and Steed Presence Questionnaire

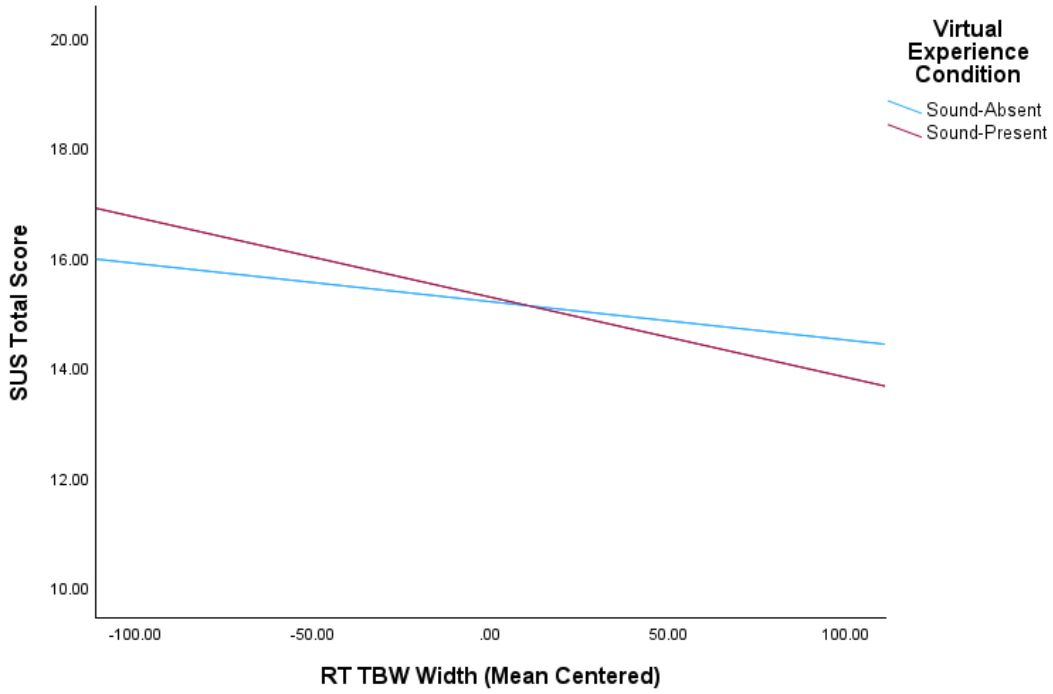
The following regression was conducted between individuals' redundant target (RT) temporal binding windows (TBW) and the SUS score, while controlling for age and prior VR experience. The results of the multiple regression did not display a significant effect of individuals' RT TBW on the SUS score (b = -0.011, CI [-.026, .004], t = -1.458, p = .150), failing to find a relationship between window width and presence as measured by the SUS. Similarly, the sound condition did not report a significant effect on the SUS score (b = .041, CI [-.715, .796], t = .108, p = .914), indicating that the addition of sound did not affect individuals' sense of presence.

However, age was a significant predictor of the SUS score (b = .160, CI [.115, .205], t = 7.108, p < .001), suggesting that older participants tended to report higher SUS scores. In contrast, prior VR experience was not a significant covariate in this analysis (b = .009, CI [-.005, .022], t = 1.280, p = .206).

The moderation effect between the sound condition and RT TBW was not significant (b = -.004, CI [-.019, .011], t = -.517, p = .607), meaning that neither RT TBW nor the sound condition significantly affected presence as measured by the SUS (Figure 12).

Figure 12

Interaction effect of the presence of sound on the relationship between individuals' temporal binding window and self-reported presence score on the Slater, Usoh, and Steed presence questionnaire



The fourth regression was conducted between individuals' temporal order judgment (TOJ) temporal binding windows (TBWs) and the SUS score, while controlling for age and prior VR experience. The results of the multiple regression did not display a significant effect of individuals' TOJ TBW on the SUS score (b = -.005, CI [-.009, .000], t = -1.897, p = .063), suggesting no clear relationship between window width and presence as measured by the SUS. Similarly, the sound condition did not have a significant effect on the SUS score (b =

.018, CI [-.738, .775], t = .049, p = .961), indicating that the addition of sound did not affect individuals' sense of presence.

The moderation effect between sound condition and TOJ TBW was also not significant (b = -.002, CI [-.006, .003], t = -.597, p = .553), suggesting no significant interaction between TOJ TBW and sound condition in predicting SUS scores.

However, both covariates were significant predictors of the SUS score. Age had a significant positive effect on SUS scores (b = .162, CI [.115, .208], t = 6.990, p < .001), with older participants reporting higher levels of presence. Prior VR experience was also a significant predictor (b = .011, CI [.000, .022], t = 2.082, p = .042), indicating that participants with more VR experience reported higher SUS scores.

5.3.5.3 SPES Subscales: Self-Location and Perceived Possible Actions

Two additional MMRs were conducted on the measure's subscales to further investigate the significant moderated relationships between participants' ability to integrate multisensory information and the SPES score.

There was a significant main effect found between RT TBW and Self Location score (b = .013, CI [.005, .021], t = 3.205, p = .002), indicating that as RT TBW increased, so did the Self Location score. Additionally, a significant main effect of sound condition on the Self Location score was observed (b = .674, CI [.033, 1.316], t = 2.104, p = .040), suggesting that the presence of sound influenced participants' Self Location scores.

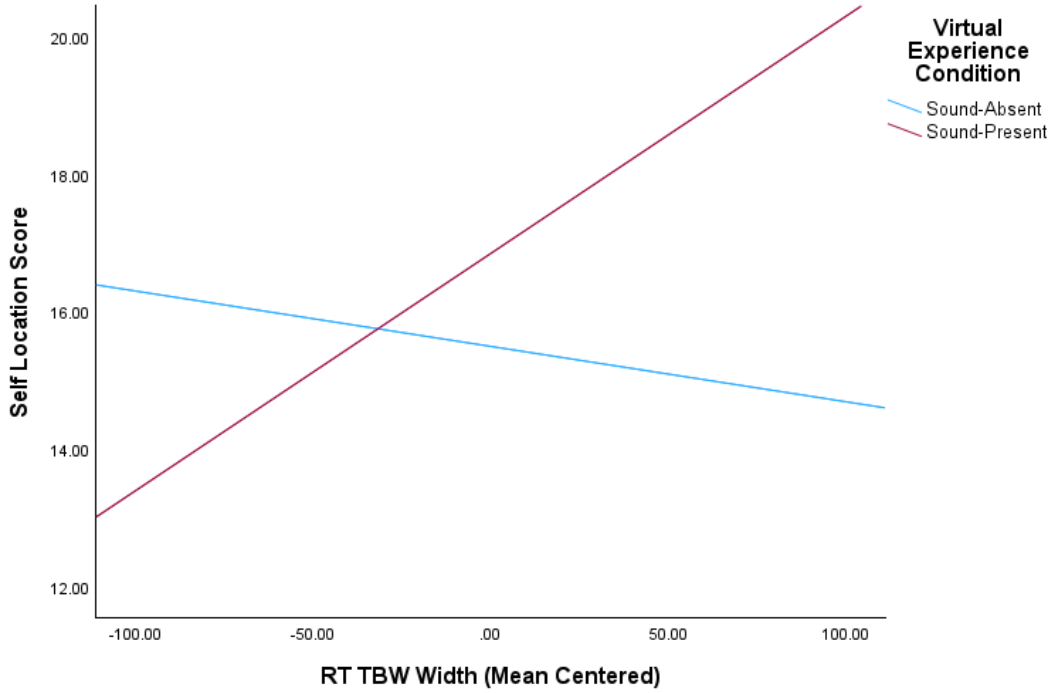
The analysis also identified a significant interaction between sound condition and RT TBW on Self Location (b = .021, CI [.013, .029], t = 5.223, p < .001). This indicates that sound moderated the relationship between RT TBW and Self Location (Figure 13).

Further analysis showed that participants who experienced the sound absent VRE displayed a negative relationship between RT TBW and Self Location score (b = -.008, CI [-.011, -.006], t = -6.546, p < .001). In contrast, participants in the sound present VRE condition displayed a positive relationship between TBW and Self Location score (b = .035, CI [.018, .051], t = 4.256, p < .001).

Age was a significant covariate, positively influencing Self Location scores (b = .123, CI [.081, .165], t = 5.849, p < .001), while prior VR experience also significantly contributed to the model (b = .013, CI [.002, .023], t = 2.474, p = .016).

Figure 13

Interaction effect of the presence of sound on the relationship between individuals' temporal binding window and self-reported score on the self-location subscale.



In addition, there was a significant main effect found between RT TBW and PPA score (b = .023, CI [.014, .032], t = 5.155, p < .001), indicating that as RT TBW increased, so did the

PPA score. However, there was a nonsignificant main effect of sound condition on the PPA score (b = -.352, CI [-.911, .206], t = -1.261, p = .212), suggesting that sound condition did not influence this outcome.

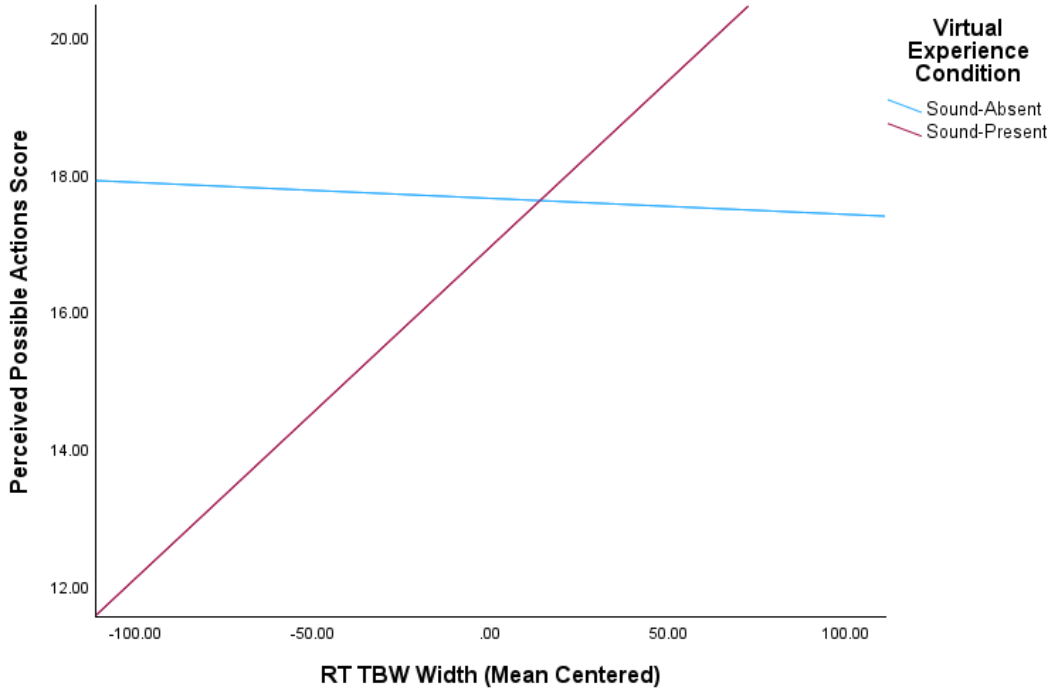
The analysis did identify a significant interaction between sound condition and RT TBW on PPA (b = .025, CI [.016, .034], t = 5.716, p < .001), indicating that sound moderated the relationship between RT TBW and PPA score (Figure 14).

Further analysis showed that participants who experienced a unisensory VRE displayed a small negative relationship between RT TBW and PPA score (b = -.002, CI [-.005, .000], t = -1.719, p = .091). However, this effect was not statistically significant. In contrast, participants in the multisensory VRE condition displayed a positive relationship between RT TBW and PPA score (b = .048, CI [.031, .066], t = 5.499, p < .001).

Age was also a significant covariate, positively influencing PPA scores (b = .067, CI [.015, .119], t = 2.563, p = .013). Prior VR experience also significantly contributed to the model (b = .009, CI [.002, .017], t = 2.410, p = .019).

Figure 14

Interaction effect of the presence of sound on the relationship between individuals' temporal binding window and self-reported score on the perceived possible action subscale



In summary, data from both temporal binding tasks were analysed using MATLAB scripts from Diedrich and Colonius' extended TWIN model (2015) to estimate the size of individuals' temporal binding windows. Nonparametric correlations confirmed that both age and prior VR experience could be potential covariates, so they were included as control variables in the subsequent moderation analysis.

Several moderated multiple regressions were conducted to determine whether the relationship between an individual's temporal binding windows and their self-reported sense of presence, was influenced by the type of virtual environment—either audio-visual or visual-only. Age and VR experience significantly contributed to each regression model, except for the SUS presence questionnaire. There were no significant main or interaction effects between temporal binding windows and the SUS presence questionnaire.

However, the sound condition did moderate the relationship between the SPES and participants' redundant target temporal binding windows. As participants' temporal binding windows increased, their reported sense of presence in a multisensory environment also

increased. Conversely, in the visual-only condition, a smaller yet significant negative relationship was observed. Further analysis revealed a similar pattern across both subscales of the SPES: Self-location (SL) and Perceived Possible Actions (PPA).

These findings support the overarching hypothesis that the way individuals process multisensory information, plays a crucial role in developing a sense of presence.

5.4 Discussion

The study presented in this chapter had two main aims. The first aim investigated whether minimising conflicting cues within the VRE would lead to a positive relationship between individuals' TBW and self-reported presence. It is argued that there is a need to process multisensory information consistently for mental models to form and be maintained (Wirth et al., 2007). The second aim was to see if the relationship between TBWs and presence would be moderated by the presence (or absence) of multisensory cues. If TBWs are important in the formation of presence, the relationship should be moderated by the number of modalities stimulated.

The results supported the first hypothesis (H1: *Individuals TBW would positively predict their sense of presence in a multisensory space. There will be no relationship between simulator sickness and TBW or Presence.*). Individuals in the multisensory experience displayed a positive relationship between TBW width measured on the RT task and presence. Wider windows (those better at maintaining the perception of multisensory consistency) predicted higher levels of presence in a non-conflicting multisensory environment. This finding was true for both subscales of the SPES, self-location and perceived possible actions. The same was also found with the overall score for the SPES. However, the TOJ task displayed no significant relationship with any presence measure, nor was the SUS predicted by the RT task. This further highlights the difference between the questionnaires and TBW tasks (Eskes, 2007;

Hartmann et al., 2016, p. 8; Miller & Schwarz, 2006). Furthermore, the second part of the first hypothesis was also supported by simulator sickness not correlating with either TBW, or presence in this VRE.

The second hypothesis was also largely supported (H2: *The relationship between TBWs and presence will be moderate between multisensory and unisensory VR environments.*). The moderated regressions identified a significant interaction between TBW widths estimated from the RT task and all SPES measures. The multisensory condition displayed a positive relationship between TBW width and SPES measures. However, an unexpected finding was that the unisensory condition displayed a negative (albeit small) relationship between individuals' TBW width and SPES measures. Neither the TOJ task TBWs nor the SUS displayed a significant moderation.

The findings indicate that individuals with a wider TBW tend to experience a stronger sense of presence in a multisensory VRE. This enhanced presence may arise from their ability to maintain consistent multisensory inputs, which supports the cognitive development of mental models. In contrast, individuals with a wider TBW experienced a slightly lower sense of presence in unisensory environments. This reduction in presence could be attributed to technological issues or influences from the real world.

5.4.1 *TBW and Presence in Multisensory Room-scale VREs*

The findings from this study emphasize the important role that individual differences in TBWs play in shaping the sense of presence within a multisensory room-scale VRE. Specifically, the study highlights that individuals who can effectively integrate asynchronous sensory inputs experience higher levels of presence, even when faced with minor delays that are common in current VR technology. This section explores how technological limitations

and cognitive processes may interact to affect the perception of multisensory consistency and presence.

While technological advances in VR headsets, such as reducing display lag and offsets, have significantly improved user experiences over the past 20 years, discrepancies between modality update times persist (e.g., Dixit & Sarangi, 2023). Even when controlling for a small offset (4ms) in VR headsets, scene instability still impacts spatial presence (Kim et al., 2020). These ongoing challenges highlight the limitations of current technology in achieving perfect congruence across sensory modalities, with delays under 16.7ms being considered good (Bailey et al., 2004; Jung et al., 2000). The results of this study highlight that, despite such technological limitations, individuals with wider TBWs are better equipped to integrate asynchronous sensory inputs, supporting a stronger sense of presence.

The room-scale simulation used in this study provided a confined space to reduce conflicting cues (Ng et al., 2020). However, in room-scale VR, where users can physically move, processing demands on the VR system are higher than compared to passive experiences (Warburton et al., 2022). This increased demand can cause latency, especially in multisensory environments, where processing multiple sensory inputs adds to the complexity (Stauffert et al., 2020). Although these technological constraints are not the main focus of this study, they provide important context for understanding the relationship between TBWs and presence in multisensory virtual environments.

The limitations of the technology, particularly issues related to latency and lag, mean that multisensory information is not always presented consistently (Stauffert et al., 2020; Tachibana & Matsumiya, 2022). This inconsistency has been identified as a problem in previous theories (Steuer, 1992; Wirth et al., 2007). Consistent presentation of multisensory cues is crucial for forming and maintaining mental models (Slater et al., 2022; Wirth et al.,

2007). For an individual to feel present in the virtual world, the mental model must be prominent and well-defined. When sensory information is inconsistent, competing reference frames can develop (Ehrenfeld & Butz, 2013), potentially undermining the construction of a coherent mental model. For example, processing audio-visual stimuli as two separate percept's can lead to distinct reference frames, which may weaken the sense of presence if one of these frames aligns more closely with 'real-world' cues (Riecke & Von der Heyde, 2003; Wirth et al., 2007).

Visual and auditory cues are the modalities individuals predominantly rely on to orient themselves within an environment (Larsson et al., 2007; Stein & Meredith, 1997). Ensuring that these cues form a consistent multisensory perception, even when slightly out of sync, would create a strong egocentric reference frame. If this reference frame pertains to virtual reality, mental models could be formed and sustained despite delays. The study found that participants with wider temporal binding windows (TBWs) reported a greater sense of presence in the multisensory VRE. This supports the notion that their ability to integrate slightly asynchronous audio-visual inputs enhances the formation of a dominant and stable egocentric reference frame, even amid minor inconsistencies.

In a multisensory environment without conflicting cues, individual differences in the ability to process offset multisensory information predict one's self-reported sense of presence. Those with a wider temporal binding window—meaning they are less likely to notice delays—report feeling more present in a room-scale multisensory virtual environment, than those with a narrower window. However, this does not clarify at which stage TBWs influence presence. It may be that these windows are significant only for forming situational spatial models—mental models related to one's location (Wirth et al., 2007). Alternatively, they could play a vital role in maintaining these models over extended periods, especially when sensory information fluctuates, or they may serve both functions. Nonetheless, the findings from this

study and previous studies within this thesis emphasise the importance of TBWs in shaping the sense of presence, highlighting individual differences at the sensory perception level.

5.4.2 *Moderating the relationship between TBWs and Presence*

The second hypothesis of this study proposed that the presence or absence of multisensory stimuli would influence the relationship between TBW width and presence. As previously discussed, in multisensory virtual reality environments (VREs), it is crucial to maintain a unified perception of the sensory information being presented (Steuer, 1992; Wirth et al., 2007). This unity helps reduce the number of egocentric reference frames that develop, allowing the primary reference frame created by the VRE to take precedence (Riecke & Von der Heyde, 2003). However, this is not the case in unisensory VREs, where visual stimuli alone can establish a prominent egocentric reference frame and mental model (e.g., Foley et al., 2015). This idea is supported by the finding that presence ratings in the unisensory VRE did not significantly differ from those in the multisensory VRE. In unisensory conditions, individuals are not required to maintain multisensory information, leading to the prediction that TBW would have no relationship with presence in this context. While the relationship was moderated, this speculation was not supported by the data collected here. However, the relationship identified between the TBW and SPES in the unisensory condition was small (b=-.011, $p<.001$).

One potential explanation for the small negative relationship between the TBW and SPES, is that individuals with a wider temporal binding window may have been sensitive to minor occurrences of scene instability. Whilst there was no relationship between TBW or Presence and simulator sickness in this study, scene instability may still occur on a smaller scale because a mismatch of visual-vestibular cues contributes to scene instability (Kim et al., 2020). Typically, this occurs between the time it takes for an individual to turn their head and

the 'scene' to update visually (Kim et al., 2021). As a result, individuals with a wider TBW might be more affected by scene instability, compared to those with a narrower window.

This situation resembles how conflicting vestibular cues have previously been shown to impact presence (Study 2). Such a relationship likely applies to both unisensory and multisensory experiences. However, the minor negative impact of incorrectly binding unstable scene cues would likely be overshadowed by the more significant role of maintaining an audio-visual egocentric reference frame in multisensory conditions (Cooper et al., 2018). The incongruent visual-vestibular mismatch would primarily play a role in unisensory conditions. This highlights the dual function that the TBW may have in the development of presence and simulator sickness.

5.4.3 *Towards a Dual Role of the TBW in Presence and Simulator Sickness*

Combining multisensory information into a single percept has long been considered essential for individuals to orient themselves within their surroundings (Gramann et al., 2010; Klatzky, 1998; Stein & Meredith, 1993). In the natural world, sensory information is predominantly presented in a congruent manner, except for events that occur at great distances (Sugita & Suzuki, 2003). However, sensory information about a virtual world in a head-mounted display (HMD), differs significantly from that in the natural world (Tachibana & Matsumiya, 2022).

In the natural environment, multiple sensory cues emerge from a singular point; for example, the audio of someone speaking is synchronised with the visual movements of their lips. In contrast, in virtual reality, multiple sensory cues associated with a single event are delivered by separate technologies, each capable of producing only one type of sensory cue (Slater, 2009). For instance, an OLED screen may depict a person’s lips moving while

headphones produce their speech. Despite these differences in how multisensory cues are generated, individuals can still orient themselves and feel present in both environments.

As explored in this chapter, when individuals are presented with multisensory stimuli from a VRE which do not contain conflicting cues, those who are better at tolerating offsets between the stimuli, go on to feel more present. This relationship does not present itself in unisensory environments. Instead, the relationship between TBWs and Presence in unisensory environments is in the same direction as when conflicting cues are present. Wider windows led to lower levels of presence.

These findings suggest a dual role of the TBW in the formation of presence. First, individuals with wider windows can integrate asynchronous multisensory stimuli effectively, enabling the formation of salient egocentric reference frames. This reference frame likely supports the construction of situational spatial models of the virtual environment. An individual's ability to tolerate further asynchronous information may help to maintain this mental model, facilitating a sustained sense of presence even when sensory inputs are slightly offset. The ability of the individual mitigates the requirement for a VRE to contain consistent multisensory information.

Second, the same tolerance for asynchronous information that supports presence in non-conflicting multisensory environments may become a disadvantage in the presence of conflicting cues. In these cases, the individual’s wider TBW may attempt to integrate unrelated and conflicting cues. This is likely to lead to difficulties in maintaining a stable egocentric reference frame. Consequently, the individual may develop a fragmented mental model of the virtual environment, resulting in a diminished sense of presence and an increased likelihood of experiencing simulator sickness.

The results presented in this chapter confirm the first element of this dual role. The second element is speculative based on the negative relationships reported in Studies 1 and 2. Given these findings, further research is needed to clarify the dynamics between TBWs, presence, and simulator sickness, particularly in environments where conflicting cues are present. The following section will discuss the limitations of the current study, which will help shape the direction of the subsequent research aimed at clarifying these dynamics.

5.4.4 *Limitations of the Current Study*

The primary aim of this study was to isolate the moderated impact of multisensory versus unisensory cues on the relationship between TBWs and the sense of presence. To achieve this, an identical environment was chosen for both groups, with one group only receiving visual information about the environment. Haptic feedback was turned off for both groups, and the HMD did not have the capability to stimulate olfactory or gustatory senses. This setup allowed for a comparison between an audiovisual group and a visual-only group.

However, while this approach targeted the most relevant sensory inputs, it is important to acknowledge the limitations inherent in the study's sensory manipulation. As previously mentioned, humans have more than five exteroceptive senses, with the vestibular system being considered the sixth (Reichenbach et al., 2016). Additionally, there are at least seven recognised interoceptive senses, including proprioception and nociception (Ceunen et al., 2016). The inability to control for all these senses represents a limitation, as their influence on presence could not be fully accounted for. Despite this limitation, participants reported varying levels of presence between the audio-visual and visual-only conditions, which appeared to depend on the size of their TBW. Although it was not possible to fully account for the influence of other senses, the comparison between visual and auditory-visual cues concentrated on two primary senses that are crucial for orienting oneself in an environment (Dinh et al., 1999;

Cooper et al., 2018). This comparison is particularly relevant because the vestibular sense was kept consistent across conditions, which helped prevent conflicting cues from disrupting the experimental manipulation (Ng et al., 2020).

A second limitation was the lack of a relationship between the SUS and either TBW measure. Although TOJ TBW predicted SUS scores in the previous study, this relationship was not identified here. Nor was there a relationship between the SUS and RT TBWs. One potential explanation for this discrepancy is that the SUS may not be sensitive enough to detect differences in user experience with more modern VR, such as the post-2019 headsets used in this study (Hartmann et al., 2016). These newer systems feature significant improvements in screen quality, field of view, and design, as well as enhanced positional tracking that provides interactions more closely aligned with real-world experiences. In contrast, the VR system used in the previous study featured a wired headset and a less realistic environment, which may have made the SUS more responsive to variations in user experience.

Given the enhanced capabilities of modern headsets, the SUS may be assessing a more general sense of presence that fails to capture more nuanced differences, such as those related to temporal binding. In fact, the means and standard deviation of the SUS were smaller than those of the SPES, suggesting more similar responses across participants. The lack of sensitivity of the SUS is consistent with findings from Laarni et al. (2015), who reported similar challenges in using the SUS. Additionally, other studies have demonstrated that while the SUS can differentiate between real and virtual conditions, it often fails to distinguish between different VRE conditions (Baños et al., 2008; Mania & Chalmers, 2001; Nunez & Blake, 2003). These findings suggest that the SUS may not be sufficiently sensitive to detect subtle differences that could be more apparent in other measures, such as those captured by the SPES (Hartmann et al., 2016).

A final limitation to note is the use of a between-subject design to investigate individual differences. This design was chosen to prevent carry-over effects, where exposure to one condition could influence responses in a subsequent condition, and to reduce the potential for demand characteristics (Greenwald, 1976). By using different participants in each group, the study aimed to eliminate any influence of prior conditions on subsequent responses, ensuring the validity of the findings. However, while a between-subjects design avoids these issues, it may reduce the power to detect subtle relationships between individuals' TBWs and their sense of presence. Although this study was adequately powered (Cohen, 1992) and displayed clear support for the hypothesis, the ability to explore individual differences more comprehensively between conditions is limited in a between-subjects design.

In contrast, a within-subjects design would allow each participant to experience all conditions, thus reducing the variability between conditions caused by individual differences. By measuring how each individual responds to different conditions, a within-subjects approach can more precisely assess the impact of TBWs on presence and simulator sickness within the same person, minimising the noise introduced by participant-level variability in a between-subjects design.

Given these considerations, the following study was designed to explore the effect of two conditions on each individual while avoiding order effects. Order effects can occur when participants are exposed to multiple conditions in a way that influences their responses to later conditions, such as through fatigue, learning, or adaptation. The design of the next study aimed to avoid these factors to ensure that the impact of the conditions could be examined independently within participants.

5.4.5 *Summary and Conclusion.*

Due to the findings from Chapter 4, the study presented here aimed to investigate the relationship between TBWs and Presence when conflicting cues are minimised. Conflicting cues are hypothesised to lead to simulator sickness and undermine the presence experience (Grassini et al., 2021). In simulations that do not have conflicting cues, it was hypothesised that wider temporal binding windows may help maintain multisensory consistency. This is compared to the results from Studies 1 and 2, which demonstrated that narrow windows may lead to higher reports of presence.

The study conducted within this chapter utilised the same RT and TOJ tasks as Study 2 to estimate the size of individuals' TBW. Unlike Study 2, however, the extended-twin model was used to estimate the window widths (Diederich & Colonius, 2015). Participants were then randomly allocated into one of two groups. Both groups experienced the same VRE experience, but one group only had visual information about the virtual environment presented to them. Afterwards, individuals completed the same questionnaires administered in Study 2: the SPES, the SUS, and the SSQ (Hartmann et al., 2016; Kennedy et al., 1993; Slater et al., 1994).

The results indicated that wider windows led to higher levels of reported presence in multisensory environments, which did not contain conflicting cues. In the unisensory environment, which did not contain conflicting cues, window widths had a small negative impact on presence. The latter is hypothesised to occur due to potential minor issues of scene instability. Nonetheless, the number of modalities stimulated moderates the relationship between TBWs and Presence, supporting both the research questions explored within this chapter.

The results of this study, along with the exploratory studies reported in chapters 3 and 4, indicate that TBWs may serve a dual role in the formation of presence and simulator

sickness. On one side, they demonstrate a difference in allowing individuals to maintain multisensory consistency beyond the capabilities of the technology, providing one explanation for why individuals experience different levels of presence whilst in the same VRE using the same HMD. On the other hand, there is support for the idea that these wider windows can also be detrimental to the formation of presence in environments with conflicting cues. Nonetheless, this hypothesis is formed from the exploratory studies and needs further exploration. Therefore, study 4 investigates this potential dual role of the TBW in presence, using a within-subject design.

Chapter 6: Individual Differences in Temporal Binding Windows Moderate the Impact of Conflicting Vestibular Cues on Presence and Simulator Sickness (Study 4).

6.1 Introduction

The overarching objective of this thesis has been to investigate how individual differences in multisensory integration impact the sense of presence in virtual reality. A fundamental assumption is that variations in temporal binding windows (TBWs), the timeframes within which individuals integrate multisensory information, affect how people experience presence in VR. This assumption aligns with established theories of presence, which emphasise the need for consistent multisensory cues for optimal user experience (Slater et al., 2022; Steuer, 1992; Wirth et al., 2007; Witmer & Singer, 1998). Individuals with varying TBW widths, particularly those with wider windows, may process multisensory cues differently, potentially leading to diverse levels of perceived presence.

The exploratory studies presented in Chapters 3 and 4 provided initial insights into this relationship. Findings indicated that individuals with wider TBWs tended to report a reduced sense of presence. These results were somewhat unexpected, as theoretical models of presence suggest that individuals with wider TBWs—typically less sensitive to minor inconsistencies in sensory information—might experience a greater sense of presence due to their higher tolerance for asynchronous cues (Stevenson et al., 2013; Wirth et al., 2007). However, contrary to these models, the exploratory studies revealed that those with wider TBWs experienced a diminished presence, suggesting that another factor might moderate this relationship.

Further analysis in Study 2 introduced simulator sickness as a covariate, revealing a positive correlation between the size of an individual's TBW and simulator sickness in Study 2a. Furthermore, simulator sickness was also a significant predictor of presence scores in Study

2b. The results pertaining to simulator sickness suggest that in VR environments where sensory cues conflict, individuals with wider TBWs may experience heightened simulator sickness and decreased presence. These studies used virtual environments which contained conflicting sensory inputs, such as visuals that suggested movement, while vestibular cues indicated stillness (Chang et al., 2020). It was proposed that individuals with wider TBWs might incorrectly bind these conflicting cues, leading to increased simulator sickness and a corresponding reduction in presence.

These findings suggest that the relationship between TBWs and presence is not straightforward. When exposed to environments with multisensory conflicts, individuals with wider TBWs may experience reduced presence as they bind conflicting cues, which undermines the mental representation of the environment. However, when VR environments avoid conflicting cues and offer a more consistent multisensory experience, the opposite effect has been observed: individuals with wider TBWs report higher levels of presence. Study 3 confirmed this hypothesis by using a "room-scale" VR setup that provided a matched visual-vestibular experience with minimal navigation requirements. Here, a moderation analysis showed that, in non-conflicting multisensory environments, those with wider TBWs reported a greater sense of presence. This dual role of TBWs—their potential to enhance and diminish presence—highlights the importance of multisensory consistency in VR experiences.

The first three studies provide evidence that TBWs significantly influence the experience of presence in VR. However, this relationship appears to depend on the sensory consistency within the virtual environment. While Study 3 confirmed the positive role of TBWs in enhancing presence under non-conflicting conditions, the negative impact of TBWs—suggested by the first two exploratory studies—warrants further investigation. Specifically, these findings raise the need for confirmatory research to verify whether individuals with wider TBWs experience reduced presence in VR environments with conflicting sensory cues.

This final empirical chapter addresses this gap by focusing on environments where sensory cues conflict, providing a confirmatory examination of the relationship between TBWs and individuals' sense of presence under these conditions. The current study investigates feelings of presence in two variations of the same VR environment: one containing conflicting multisensory cues and one with consistent, matched cues. This design allows us to assess how TBWs impact presence in each condition, clarifying the circumstances under which TBWs act as a negative or positive factor in the VR experience. Additionally, it will confirm that previous findings were not influenced by unique characteristics of the environments, aside from the level of sensory conflict attributed to them.

To enhance ecological validity and fully immerse users, the virtual environment used for this study is a non-Euclidean world where users can navigate continuously within a constrained physical space (Skrodzki, 2020). This approach ensures consistent and controlled multisensory experiences whilst minimising spatial limitations, allowing for naturalistic navigation (Grigoriadou et al., 2020; Suma et al., 2012). Avoiding conflicts in sensory cues through seamless, non-Euclidean VR looped environments presents an ideal framework for this confirmatory study.

6.1.1 *Non-Euclidean VR and Impossible Spaces*

Virtual reality provides an immersive platform capable of simulating varied spatial environments, making it possible to explore spaces that could not exist in the real world (Skrodzki, 2020). In Room-scale VREs, the virtual and physical spaces are mapped linearly, meaning that user movements in the real world correspond directly to movements in the virtual environment (Boldt et al., 2018). Whilst this one-to-one mapping is intuitive, it has limitations. This is observed when virtual spaces are designed to be large or intricate and then physical constraints, such as room size, restrict user movement (Nilsson et al., 2018; Suma et al., 2012).

One approach to overcoming these limitations is using non-Euclidean VR spaces, specifically through the concept of “impossible spaces,” which allow for the illusion of larger, more complex spaces within a physically constrained environment (Suma et al., 2012).

Non-Euclidean VR environments break the usual rules of spatial consistency, meaning that the virtual world’s dimensions, angles, and connections do not necessarily align with real-world spatial properties (Skrodzki, 2020). Impossible spaces, as an application of non-Euclidean VR, enable users to navigate seemingly infinite or complex environments within limited physical spaces (Suma et al., 2012). These environments feature self-overlapping architectures where spaces appear to occupy the same physical location, allowing for naturalistic navigation even as users encounter geometrically impossible layouts. This technique manipulates spatial cues so that users can turn, walk, and explore continuously without realising they are moving in loops or other carefully calculated paths within a restricted area (Suma et al., 2012). By creating these perceptual illusions, impossible spaces avoid the sensory conflicts that can arise when users’ visual perception of movement differs from their physical experience.

Impossible spaces achieve this effect by subtly updating the environment out of the user’s view (Suma et al., 2012). For example, as users walk through doorways, the environment shifts to create the impression of connected yet expansive rooms. This mechanism capitalises on spatial judgments, the process by which users interpret spatial relationships and distances based on environmental cues (Robb & Barwulor, 2019). In impossible spaces, consistent visual information is provided, even when spatial layouts defy physical logic, encouraging users to rely on their spatial judgments to make sense of the environment (Robb & Barwulor, 2021). Because these cues align with the user’s perception of natural movement, spatial judgments remain largely consistent with expectations, allowing users to form a coherent mental model of the space (Peck et al., 2011).

This distinctive quality of impossible spaces makes them an ideal alternative to the previous VREs for investigating the relationship between TBWs and presence. Impossible spaces eliminate conflicting sensory cues typically associated with VR locomotion systems by allowing natural locomotion (Lochner & Gain, 2021). For example, when visual cues indicate that a person is moving and the vestibular cues suggest the same, the experience of locomotion feels very similar to that of the real world. This congruent sensory environment may facilitate a higher sense of presence, particularly for individuals with wider TBWs who may otherwise experience a reduced sense of presence in VR environments with sensory conflicts. This reduction in conflicting cues aligns with research suggesting that natural movement contributes significantly to the sense of presence in VR (Cherni et al., 2020; Usoh et al., 1999).

6.1.2 *Impossible Spaces, Presence and Related Concepts*

In recent years, researchers have explored how non-Euclidean VR environments, particularly impossible spaces, influence concepts related to presence, such as spatial perception and navigation (e.g., Cherni et al., 2020; Fisher et al., 2017; Suma et al., 2012). By using spatial distortions that remain below users' detection thresholds, the point at which inconsistencies become noticeable, impossible spaces enable continuous, naturalistic movement within physically constrained areas (Suma et al., 2012). When these thresholds are not breached, users are often unaware of the non-Euclidean nature of the space, with some studies reporting moderate to high levels of presence in impossible spaces (Cherni et al., 2020; Fisher et al., 2017; Robb & Barwulor, 2019). While these studies do not explore presence in-depth, their findings align with broader research indicating that real walking, rather than artificial locomotion methods like teleportation, enhances presence by reinforcing sensorimotor contingencies (Ruddle & Lessels, 2009; Usoh et al., 1999). Additionally, research

on overlapping spaces indicates that users often believe they are navigating a larger area than physically exists, suggesting that impossible spaces can create an illusion of vastness that feels perceptually authentic (Vasylevska & Kaufmann, 2015). This may lead to better mental models of the virtual reality environment compared to other locomotion techniques, such as walking-in-place (Peck et al., 2011).

The concept of mental models is particularly relevant in understanding how users may adapt to non-Euclidean VR. As previously discussed, mental models are cognitive representations that users develop to make sense of complex environments (Glenberg, 1997). In traditional VRE, mental models align with real-world spatial logic; however, in non-Euclidean VR, users unknowingly form models based on the artificial, expansive design of impossible spaces (Warren et al., 2017; Muryy & Glennerster, 2021). The seamless integration of consistent spatial cues helps users construct mental models that accommodate the unique layout without disrupting immersion (Gibson, 2024). While direct evidence linking this mental model adaptation to presence in impossible spaces is limited, studies on spatial updating in VR indicate that users can maintain orientation and immersion despite these small spatial distortions (e.g. Nguyen-Vo et al., 2018; Peck et al., 2011; Robb & Barwulor, 2021).

Beyond spatial perception, impossible spaces may also influence action judgments (Payne & Robb, 2024) or the assessment of one's ability to perform actions within the environment. Theories of presence often emphasise perceived action affordances, where an environment's ability to support movement and interaction contributes to the feeling of being located there (e.g. Wirth et al., 2007). When users feel they can navigate freely and interact with the environment without constraints, their confidence in these judgments increases, fostering a heightened sense of presence (Schubert et al., 1999b). By allowing natural locomotion, impossible spaces create an experience where users feel physically oriented in the space, leading to an embodied interaction that enhances the believability of the environment

(Grigoriadou et al., 2020). This is particularly important for theories that conceptualise perceived possible actions as a key element in the formation/maintenance of presence (e.g. Schubert et al., 1999b; Wirth et al., 2007). Because natural locomotion aligns with the perceived possible actions expected in the real world, it might strengthen the sense of presence by reinforcing embodied interaction. Further research is needed to determine whether impossible spaces provide comparable levels of presence to real-world navigation in VR. However, findings from Cherni et al. (2020), Fisher et al. (2017), and Robb and Barwulor (2019), which report moderate to high presence levels in impossible spaces, support the broader notion that natural locomotion in these environments aids presence formation.

In summary, while direct research on the relationship between impossible spaces and presence is limited, studies reporting moderate presence levels (Cherni et al., 2020; Fisher et al., 2017; Robb & Barwulor, 2019) support the idea that seamless locomotion contributes to feeling present within a non-Euclidean VRE. By staying below detection thresholds for spatial inconsistencies (Vasylevska & Kaufmann, 2015), maintaining alignment with user-generated spatial judgments and mental models (Robb & Barwulor, 2019), and supporting naturalistic action judgments (Payne & Robb, 2024), these environments can foster an experience of immersive, expansive space (Gibson, 2024; Muryy & Glennerster, 2021). This combination of perceptual believability and navigational freedom positions impossible spaces as a valuable tool for assessing presence and simulator sickness in conflicting and non-conflicting VREs, making them particularly suitable for the following study exploring the impact of individuals temporal binding windows on simulator sickness and spatial presence.

6.1.3 *Beyond Self-Location and Perceived Possible Actions*

In addition to altering the environment, a different measure of spatial presence was employed to better understand the elements that may emerge when individuals explore a non-Euclidean VRE. While the short and refined Spatial Presence Experience Scale (SPES) has provided valuable insight into presence throughout this thesis, particularly in relation to the process model of presence (Hartmann et al., 2016; Wirth et al., 2007), it does not fully encompass the entire model. Specifically, the SPES does not address spatial situational models, higher cognitive involvement, or attention allocation, all of which play a role in the formation of presence and may be influenced by an individual's TBW width. Furthermore, each of these factors may be impacted by the non-Euclidean VRE. The Measures, Effects, Controls, Spatial Presence Questionnaire (MEC-SPQ) (Vorderer et al., 2004a) was used to investigate these additional factors potentially influenced by a non-Euclidean environment and align the findings with the process model of presence.

The MEC-SPQ was developed from 103 items across all subscales. The validity and reliability of the items were tested in a multimedia study, investigating the questionnaire's applicability to multiple mediums (Text, Hypertext, Film, and VR; Vorderer et al., 2004b). The results highlighted high internal consistency for 64 items split between the subscales (α=0.78-0.93; Vorderer et al., 2004b). Based on the results of the test study, a revised final version of the MEC-SPQ was compiled. This final version can be administered in three formats: a 64-item questionnaire, a 48-item version, and a 36-item format. The full version contains eight items per construct. The two smaller versions contain six items and four items per construct, respectively (Vorderer et al., 2004a).

The MEC-SPQ, a precursor to the SPES, includes two key scales: Self-Location (SL) and Perceived Possible Actions (PPA), that directly informed the latter questionnaire. The SPES incorporates four out of the eight questions from these MEC-SPQ scales, maintaining a partial but direct overlap. It is useful to maintain these as sub-measures because they provide

valuable insight into the relationship between individuals' TBW and presence and align with previous research into non-Euclidean VREs (Robb & Barwulor, 2019; Payne & Robb, 2024). As previously discussed, individuals navigating non-Euclidean spaces have reported spatial and action judgments comparable to those made in Euclidean environments (Robb & Barwulor, 2019; Payne & Robb, 2024), but the MEC-SPQ offers a more nuanced approach to ensure these elements can be reliably compared. This is because the MEC-SPQ also measures additional factors representing the eight constructs of the process model: Attention Allocation, Spatial Situational Model, Higher Cognitive Involvement, Suspension of Disbelief, Domain-Specific Interests, and Visual Spatial Imagery (Vorderer et al., 2004a). These additional factors may provide further context for understanding spatial presence in non-Euclidean environments, alongside exploring the moderation effects of conflicting versus non-conflicting cues on the relationship between TBWs, presence and simulator sickness.

What distinguishes the MEC-SPQ from most other presence measures, aside from its three different sizes, is its direct development from the process model theory. Theory-led development allows the questionnaire's results to be used to make direct inferences about the underlying process model itself (Wirth et al., 2007).

6.1.4 *Rationale and Aims of the Current Study*

Building on previous findings that underscore the role of TBWs in shaping presence in VR, this final study investigates TBWs within the non-Euclidean VR environment Tea for God (Void Room, 2023), which leverages "impossible spaces". Prior studies in this thesis revealed that wider TBWs can either enhance or diminish presence depending on the sensory consistency of the VR environment. When multisensory cues conflict, as in cases of visual-vestibular mismatch, wider TBWs correlate with both heightened simulator sickness and reduced presence, suggesting that sensory conflict disrupts the immersive experience for these

individuals. However, when cues align naturally, individuals with wider windows experience an enhanced sense of presence.

By using Tea for God to eliminate typical VR spatial constraints, this study creates a controlled comparison between conflicting and non-conflicting sensory conditions. This approach provides a unique opportunity to directly test whether non-Euclidean navigation reduces simulator sickness and enhances presence for individuals with wider TBWs. Additionally, this study employs the MEC-SPQ to capture a more detailed understanding of presence across cognitive and perceptual dimensions, offering insights beyond previous metrics.

This investigation serves as a test of the proposed relationship between TBWs, presence, and multisensory consistency in VR, advancing our understanding of individual differences in immersive environments and paving the way for optimised VR experiences. The following hypotheses of this study are:

H1: Individuals with wider temporal binding windows (TBWs) will report a lower sense of presence in Tea for God under conflicting sensory conditions than individuals with narrower TBWs.

H2: In non-conflicting sensory conditions within Tea for God, individuals with wider TBWs will report a higher sense of presence compared to conflicting conditions, demonstrating an enhanced VR experience when sensory cues align.

H3: Non-Euclidean navigation in Tea for God will reduce simulator sickness, particularly for individuals with wider TBWs, compared to conflicting sensory conditions.

H4: The MEC-SPQ will reveal significant differences across dimensions of presence, with individuals with wider TBWs showing distinct patterns in attention allocation, spatial

situational modelling, and cognitive involvement in Tea for God, depending on sensory consistency.

6.2 Methods

6.2.1 *Participants*

100 participants (75 Females) aged 18 – 43 (M = 21.51, SD = 4.60) were recruited for this study. Participants were all students studying at York St John University and received course credit for their participation. All participants had normal hearing and normal or corrected-to-normal vision and were fluent in English. 2 participants did not complete all study sections due to withdrawing part way through the experiment, leaving 98 participants who completed the full study.

An a priori power analysis was conducted using G*Power (3.1) to determine the minimum sample size required to detect a large effect ($f^2 = 0.15$) with 95% power and an alpha level of .05. The analysis was based on a linear multiple regression model with one continuous predictor, which is conceptually equivalent to the MANCOVA design employed in the present study. The analysis indicated a required sample size of 89 participants. The final sample consisted of 98 participants, exceeding this threshold and providing adequate power to detect large effects.

6.2.2 *Design*

Unlike the previous studies reported in this thesis, this study used a within-subject design. Each individual experienced a congruent (Natural Locomotion) VR experience and a conflicting (Stationary) VR experience. The VR experience was the same for both conditions with the exception of conflicting vestibular cues. Individuals completed the temporal binding tasks between the VR experiences. Whilst it is acknowledged that within-subject designs can

introduce carry-over effects, it was deemed important to test the hypothesis. As such, counterbalancing of VR experience and task order was implemented to reduce order effects.

6.2.3 *Materials*

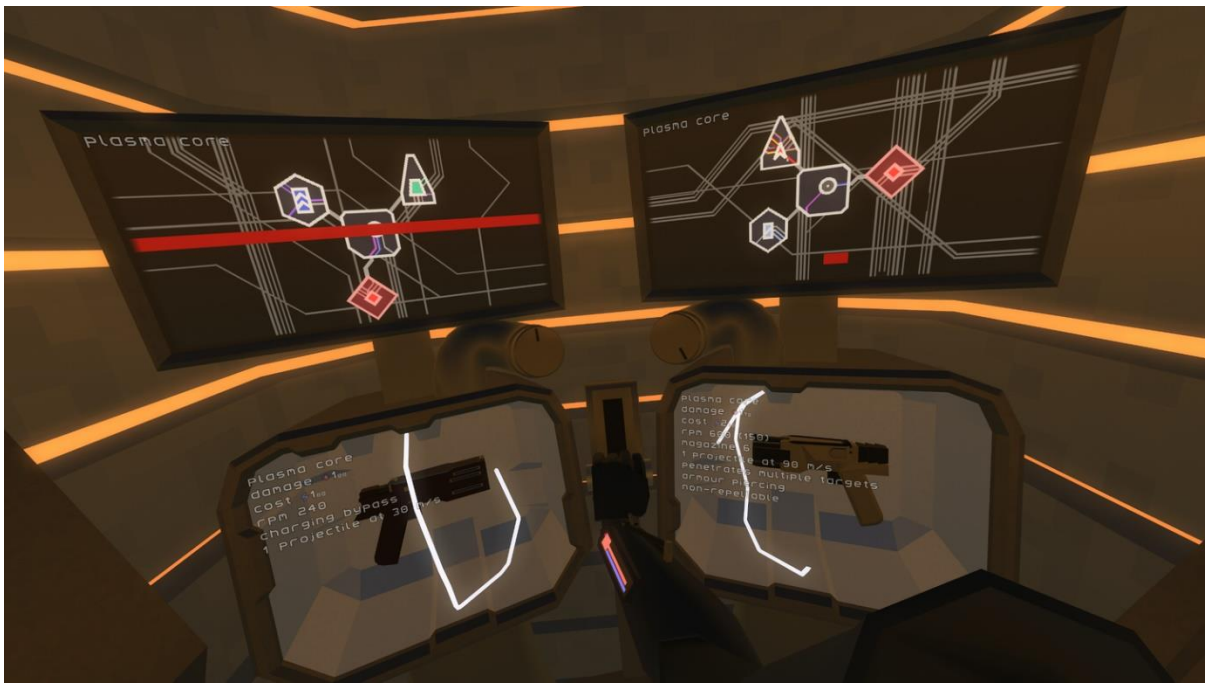
6.2.3.1 **VR Hardware and Simulation**

The equipment used to elicit a sense of presence in this study was the Oculus Quest (in line with Study 3). This is a standalone head-mounted display (HMD) with a resolution of 1440 x 1600 pixels per eye, a 93˚ field of view, and a 72Hz refresh rate. The HMD provided head and position tracking through an "inside-out" tracking system and uses a Qualcomm Snapdragon 835 'System-on-Chip' with 4GB of RAM and an Adreno 540 GPU. Participants could control movement and interactions within the virtual environment using the Oculus Quest wireless controllers. These are tracked using a simultaneous localisation and mapping system using AI, accelerometers, and the built-in camera to track infrared diodes on the controllers.

The simulation chosen for this study was 'Tea for God', a video game developed by Void Room. The simulation involves participants exploring a non-Euclidian virtual world. The environment is non-Euclidian because of the intuitive way the game adjusts to the available space. For this study, participants had 15 minutes to explore as much as possible. Throughout the simulation, participants are given opportunities to engage in combat and pick up various 'loot' items (see Figure 15). In one manipulation, participants completed the simulation by physically moving around the laboratory to navigate the virtual environment. The physical laboratory space was a 4m by 2.5m space outlined by the oculus's quest guardian boundary, which was free from obstacles. In the other condition, participants experienced the same environment but navigated it using the handheld controllers while stationary.

Figure 15

A screenshot of an interactive lever to access "loot" within the video game Tea for God, developed by Void Room



6.2.3.2 **Questionnaires**

The first two questionnaires used in this study were identical to the previous two studies. The Gaming Experience Measure, which was adapted from Lemmens et al. (2011a, 2011b) and the Simulator Sickness Questionnaire (SSQ) developed by Kennedy et al. (1993). However, unlike the previous two studies, the SSQ was used to calculate a change in simulator sickness. Individuals' levels of simulator sickness were recorded before and after each VRE condition. The SSQ subscales and totals were calculated in the same way as previously. The pre-VR score was subtracted from the post-VR score for the total and each subscale to provide a 'change in simulator sickness' measure.

Unlike the previous studies, only one measure of presence was used in this study. The questionnaire used to assess presence was the Measurements, Effects, Conditions, Spatial Presence Questionnaire developed by Vorderer and colleagues (2004a). The MEC-SPQ

directly underpins the process model of presence and was the precursor to the SPES used in previous studies (Studies 1, 2, and 3). The change to the MEC-SPQ in this study is related to the aim of understanding what elements of the process model individual differences in TBWs might affect. Therefore, the MEC-SQP was chosen because it was developed from 103 items representing the eight constructs of the process model: Attention Allocation, Spatial Situational Model, Self-location, Possible Actions, Higher Cognitive Involvement, Suspension of Disbelief, Domain-Specific Interests, and Visual Spatial Imagery (Vorderer et al., 2004a). The validity and reliability of the items were evaluated in a multimedia study examining the questionnaire's applicability across text, hypertext, film, and virtual reality (Vorderer et al., 2004b). The study found high internal consistency for 64 items across subscales. Based on these findings, a revised version of the MEC-SPQ was created and is available in three sizes: the full version with eight items per construct and smaller versions with six and four items per construct (Vorderer et al., 2004a). Despite the difference in size, the items are administered the same way for each version. Items related to the *domain-specific interest* are administered in a single block. *Visual-spatial imagery* and *higher cognitive involvement* items are presented in a randomised block. The final items relating to the other concepts are also presented in a separate randomised block. This study utilised the 4-item version of the questionnaire, which aligns with the size of the SPES subscales. The questionnaire was converted into an online format, and responses were recorded via Qualtrics. Each item is rated on a scale from 1 to 5 depending on how much people agreed ("I do not agree at all" to "I fully agree") with a statement from each of the subscales. Two items within the *suspension of disbelief* subscale were reversed scored; "I concentrated on whether there were any inconsistencies in the Virtual Reality" and "I took a critical viewpoint of the Virtual Reality presentation". Each subscale item is added to form a subscale score ranging from 4 to 20. The total score combines all eight subscales, forming a total score between 32 and 160.

6.2.4 *Procedure*

Each participant was assigned to a counterbalanced group and completed the study in a 2-hour session. The counterbalanced grouping determined whether they first experienced the stationary VRE or natural locomotion VRE. Participants first filled in a consent form and a demographic questionnaire, which asked questions about their age and gender. After this, individuals filled in a baseline measure of the SSQ to assess the change pre and post-VR. Individuals were then introduced to the VR equipment, with a standardised set of instructions for using the controllers and adjusting the headset to an optimum position. The participants then had up to 15 minutes to explore the VRE. In the stationary condition, participants were sat on a chair in the middle of the room. Their screen was shared via the Oculus app so the researcher could guide them through starting the experience. In this condition, individuals use the handheld controllers to navigate and look around the VRE. Participants had up to 15 minutes to explore as far as possible but could exit the VR equipment anytime. The duration the individuals spent in the VRE was recorded using a stopwatch. This procedure was the same for the natural locomotion condition, except that the controllers did not control the navigation or view. Instead, individuals walked around the room to navigate the VRE. The environment was limited to the room's space but updated constantly using non-Euclidean geometry, meaning that both sets of participants experienced the same distance.

After completing the first VR experience, the participants completed the three questionnaires on a computer: MEC-SPQ, SSQ, and the Gaming Experience Measure. They were then offered a short break and water before completing the TBW tasks, which collected data regarding their window widths and provided an extended break between the two VRE experiences.

Once individuals had completed the TBW tasks, they were asked to complete the SSQ for a third time. The multiple administrations of the SSQ served two purposes. First, high levels of responses surrounding simulator sickness were screened to check that individuals could ethically re-enter the VRE. Second, this provided a second pre-VR measure to calculate change before and after experiencing the next VRE (which was the counterbalance of whichever condition they experienced first).

After individuals completed the second VRE, they filled in two questionnaires on the computer: the MEC-SPQ and the SSQ. Participants were then provided with a debrief and allowed to ask any questions about the study before being thanked and dismissed.

6.2.5 *Ethics*

The research followed the BPS guidelines for ethical practice, in line with the previous research reported in this thesis. However, additional safeguarding was required due to the specific manipulation of attempting to cause visual-vestibular conflict in one condition. On the consent form and subsequent study information, attention was drawn to the fact that individuals are likely to experience simulator sickness. A separate consent form was provided regarding increased risk based on personal characteristics, including, but not limited to, suffering from a cold, flu or hang-over and having a high susceptibility to travel sickness. Participants were reminded multiple times through the study information and before entering the VRE that they could exit at any time should they need to. The researcher showed diligence in making sure every individual was safe.

Beyond introducing visual-vestibular conflict in one condition, the other manipulation allowed individuals to navigate through a non-Euclidean space. To protect individuals from harm in this condition, a guardian boundary was set up approximately half a meter away from the edge of the room. The researcher remained present at all times to ensure the guardian

boundary was still working correctly. The York St John University School of Education, Language and Psychology Ethics Committee approved this study.

6.3 Analysis

6.3.1 *Calculating the Window Widths*

Raw data from the redundant signal task were checked for outliers (<30ms and >3000ms), with individual responses removed. The raw data for the TOJ task was not removed based on the reaction time to stimuli. Instead, participants' responses were reviewed, and any identified to be using a response strategy were removed from subsequent analysis (e.g., responses such as ZZZMMM or ZMZMZM). No participant was identified using such a strategy; therefore, no data was removed for the TOJ task. The window widths for each participant were then calculated using MATLAB scripts adapted from Diederich and Colonius' (2015) extended twin model. The scripts use fmincon to optimise the objective function:

$$\sum_{\tau}\left\{\left(\frac{mean\left[RT_{VA,\tau}\right]-E\left[\widehat{RT_{VA,\tau}}\right]}{SE\left(mean\left[RT_{VA,\tau}\right]\right)}\right)^{2}+\left(\frac{f_{VA,\tau}-\widehat{\Psi(\tau)}}{SE\left[f_{VA,\tau}\right]}\right)^{2}\right\}$$

where mean [$RT_{VA,\tau}$] and $f_{VA,\tau}$ relate to the mean reaction times and relative frequency of "visual first" judgements for each stimulus onset asynchrony (SOA) (τ) in the RT and TOJ task, respectively. $E[\widehat{RT_{VA,\tau}}]$ and $\widehat{\Psi(\tau)}$ are the predicted values for each SOA based on the twin models parameters: visual processing time (λV), auditory processing time (λA), the common processing stage (μ), the TBW for the tasks (ω), and multisensory temporal gain (Δ). Therefore, the objective function calculates the deviation between measured and estimated values, returning the fit with the smallest deviation and identifying the TBW based on the twin model calculations (see appendix 5 for full script). The descriptive statistics are reported in the table below (Table 15)

Table 15
Descriptive Statistics for the TBW tasks in Study 4

Task Type (N)	Mean (ms)	Min (ms)	Max (ms)
Redundant Target (98)	84.01 (76.10)	4.75	341.81
Temporal Order (98)	181.99 (83.13)	5.44	386.20

6.3.2 *Descriptive Statistics for the Main Measures*

Raw questionnaire data for the MEC-SPQ and the SSQ were converted into total scores. Due to the repeated measures element of this study, SSQ scores were measured before and after each VR experience. The change in SSQ was calculated by subtracting the pre-VR score from the post-VR score. Each measure was checked for extreme outliers (Over 2.5 SD). No participants were removed from further analysis. Before conducting the multivariate analyses, descriptive statistics were calculated for the main measures: presence, simulator sickness and temporal binding windows (TBW). Table 16 presents these variables' mean and standard deviations across the natural locomotion and stationary conditions.

Table 16
Descriptive Statistics for Presence and SS by Navigation Condition in Study 4

		Condition Means (SD)	
	Measure	*Natural Locomotion (N=98)*	*Stationary (N=98)*
Presence (MEC-SPQ)	Total Score	126.55 (13.53)	109.41 (18.72)
	Attention Allocation	18.21 (1.86)	15.43 (3.38)
	Spatial Situational Model	16.89 (2.74)	13.78 (3.68)
	Self-Location	17.89 (2.44)	13.39 (3.89)
	Perceived Possible Actions	17.11 (2.12)	13.82 (3.53)
	Higher Cognitive Involvement	15.49 (2.69)	12.91 (3.41)
	Suspension of Disbelief	14.79 (3.08)	14.77 (2.84)
	Domain Specific Interest	11.95 (4.01)	11.34 (3.8)
	Visual-Spatial Imagery	14.22 (3.68)	13.99 (3.72)
Simulator Sickness (Change in)	SSQ Total	-2.21 (15.08)	23.62 (22.42)
	Nausea Score	-0.88 (21.79)	38.35 (33.86)
	Oculomotor Score	-7.19 (16.12)	13.07 (20.99)
	Disorientation Score	-2.84 (31.54)	39.35 (42.68)

Note: MEC-SPQ = Measures Effects Control Spatial Presence Questionnaire; SSQ = Simulator Sickness Questionnaire

6.3.3 *One-way Within Subject Multivariate Analyses*

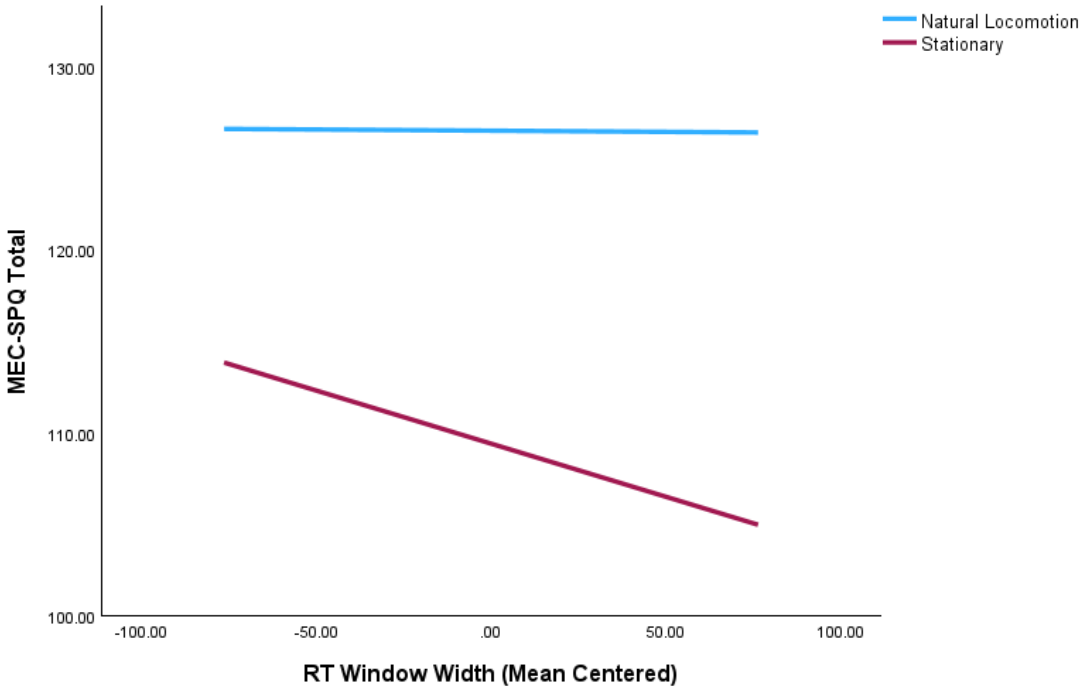
Testing the relationship between the temporal binding windows, presence, and simulator sickness in a within-subject design involves multiple steps. Firstly, two separate multivariate analyses were conducted for both temporal binding window measures. The analyses are conducted separately to reduce the high multicollinearity that arises from estimating the window widths from the same canonical model. The multivariate analyses reduce the type 1 error rate while determining the differences between presence and simulator sickness conditions, all the while examining the effect of the TBWs

The first one-way repeated measure multivariate analysis examined the effects of VR condition (natural locomotion vs stationary) on the MEC-SPQ and change in SSQ total scores, with the RT TBW as a continuous moderator. The descriptive statistics for the questionnaire totals in each condition are reported above in Table 16. On average, MEC-SPQ totals were higher in the natural locomotion condition (M=126.55, SD=13.53) than in the stationary condition (M=109.41, SD = 18.72). The opposite is the case for change in SSQ score, where stationary was higher (M=23.62, SD=22.42) than natural locomotion (M=-2.21, SD = 15.08).

Assumptions checks indicated that the data met the assumptions for repeated measures multivariate analysis. The analysis revealed a significant main effect of VR condition on MEC-SPQ total, $F(1,96) = 1325.363$, $p<.001$, Wilks $\Lambda =.068$, $\eta p2=.932$, and SSQ total, $F(1,96) = 4.948$, $p=.028$, Wilks $\Lambda =.951$, $\eta p2=.049$. The interaction effect of RT TBW by VR Condition was significant for the MEC-SPQ total, $F(1,96) = 6.872$, $p=.010$, Wilks $\Lambda =.933$, $\eta p2=.067$, but not for SSQ total, $F(1,96) = .195$, $p=.660$, Wilks $\Lambda =.998$, $\eta p2=.002$. There was also a significant three-way interaction of VR condition by presence by SSQ, $F(1,96) = 45.620$,

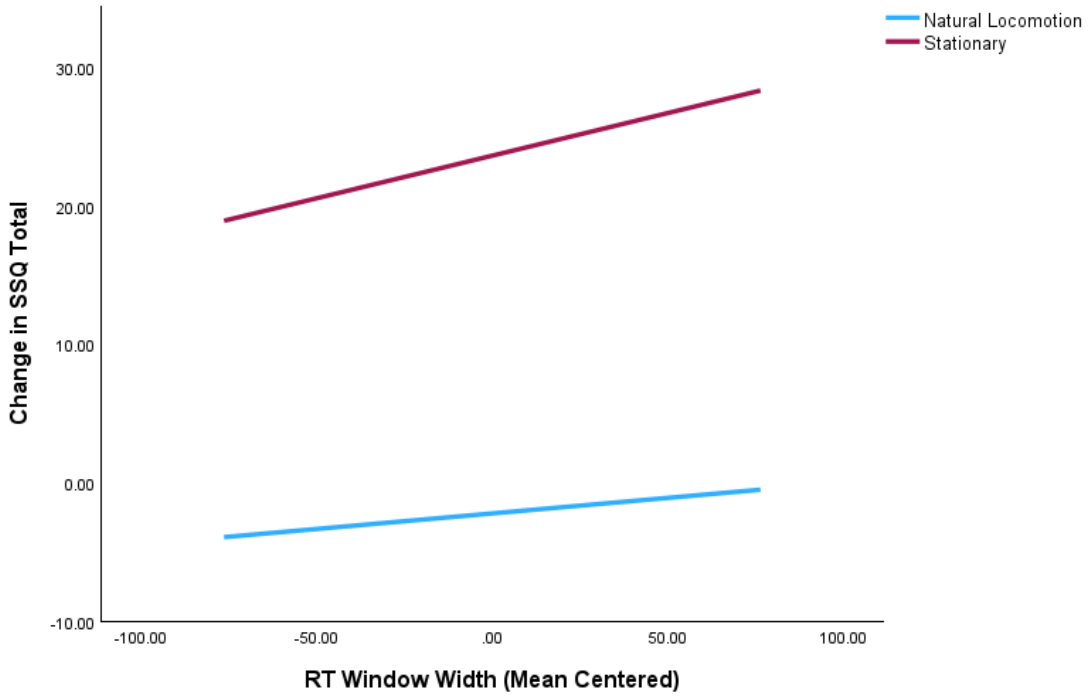
p<.001, Wilks Λ =.678, ηp2=.322, alongside a significant four-way interaction by RT TBW, F(1,96) = 4.482, p=.037, Wilks Λ =.955, ηp2=.045.

Figure 16
Interaction effect of the VR condition on the relationship between individuals' redundant target temporal binding window and self-reported score on the MEC-SPQ



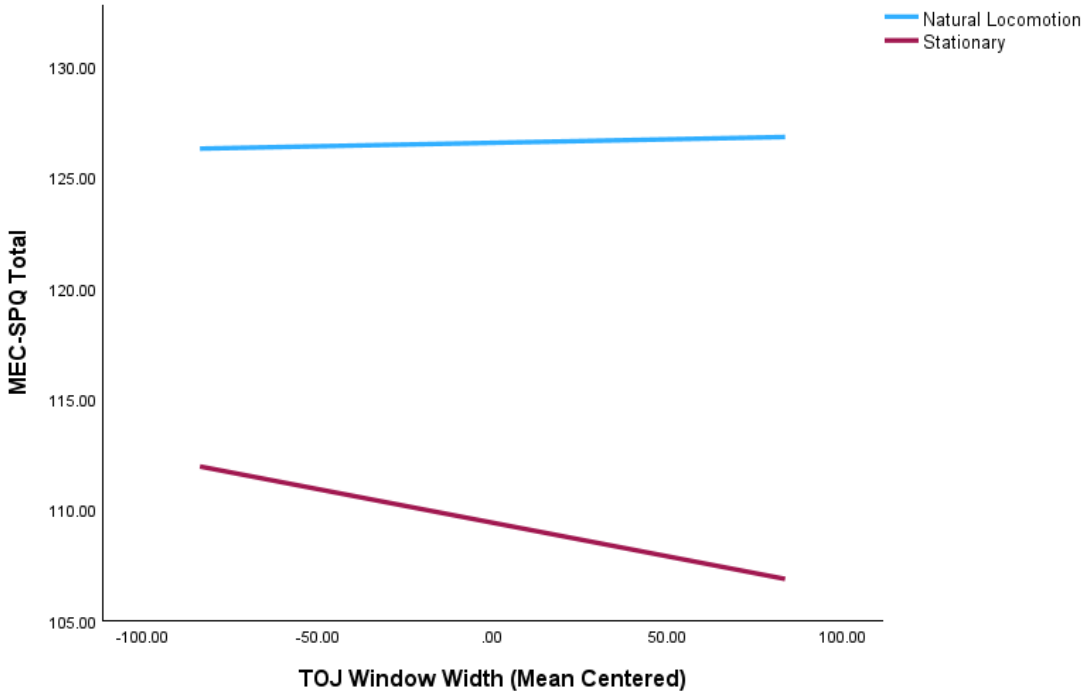
Simple slopes analysis identified a significant negative effect of the RT TBW on the MEC-SPQ total in the stationary condition (b= -.058, CI [-.107, -.010], t=-2.393, p =.019) but not in the natural locomotion condition (b=-.001, CI [-.037,.035], t=-.071, p=.944). Larger TBW windows led to lower reported presence in the stationary condition but not the room scale (Figure 16). The opposite relationship was identified for change in SSQ total (Figure 17). There was a significant positive effect of the RT TBW on Change in SSQ in the stationary condition (b=.062, CI [.003, .120], t=2.102, p =.038), but not in the natural locomotion condition (b=.022, CI [-.017, .062], t=1.121, p=.265).

Figure 17
Interaction effect of the VR condition on the relationship between individuals' redundant target temporal binding window and change in score on the SSQ



The second one-way repeated measure multivariate analysis examined the effects of VR condition (natural locomotion vs stationary) on the MEC-SPQ and change in SSQ total scores, with the Temporal Order Judgement TBW as a continuous moderator.

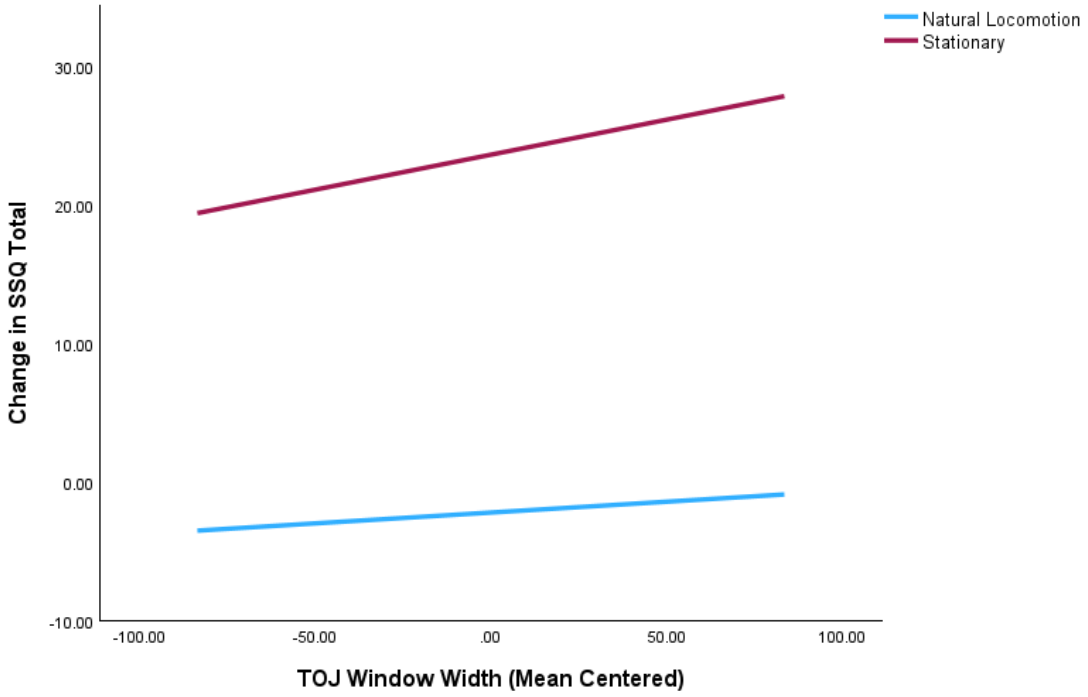
Figure 18
Interaction effect of the VR condition on the relationship between individuals' temporal order judgement temporal binding window and self-reported score on the MEC-SPQ



Assumptions checks indicated that the data for the TOJ TBW widths also met the assumptions for repeated measures multivariate analysis. The analysis revealed a significant main effect of VR condition on MEC-SPQ total, $F(1,96) = 823.982$, $p<.001$, Wilks $\Lambda = .104$, $\eta_p2 = .896$, but not for the SSQ total, $F(1,96) = 2.212$, $p = .140$, Wilks $\Lambda = .977$, $\eta_p2 = .023$. The interaction effect of TOJ TBW by VR Condition was not significant for either the MEC-SPQ total (Figure 18), $F(1,96) = 3.366$, $p = .070$, Wilks $\Lambda = .966$, $\eta_p2 = .034$, nor the SSQ total (Figure 19), $F(1,96) = .001$, $p = .970$, Wilks $\Lambda = 1.00$, $\eta_p2 = .000$. There was a significant three-way interaction of VR condition by presence by SSQ, $F(1,96) = 27.446$, $p<.001$, Wilks $\Lambda = .778$, $\eta_p2 = .222$. However, the four-way interaction by TOJ TBW was insignificant, $F(1,96) = 2.674$, $p = .105$, Wilks $\Lambda = .973$, $\eta_p2 = .027$.

Figure 19

Interaction effect of the VR condition on the relationship between individuals' temporal order judgement temporal binding window and change in score on the SSQ



6.3.4 *Post-Hoc Analyses*

To better understand what elements of the process model TBWs might be affecting, multiple repeated measure moderation analyses were conducted between TBWs and the MEC-SPQ's subscales for each condition.

The analysis revealed significant moderation effects of VR condition between RT TBW and Higher Cognitive Involvement and Self-Location. For Higher Cognitive Involvement, the results indicated a significant interaction, $F(1,96)= 12.262$, $p<.001$, Wilks $\Lambda=.887$, $\eta p2=.113$. Simple slopes analysis identified a significant negative effect of the RT TBW on Higher Cognitive Involvement in the stationary condition ($b = -0.012$, CI [-0.020, -0.003], $t = -2.642$, $p = .010$) but not in the natural locomotion condition ($b = 0.003$, CI [-0.004, 0.010], $t = 0.856$, $p = .394$).

For Self-Location, a significant interaction was also found, $F(1,96)= 9.732$, $p=0.002$, Wilks $\Lambda=.908$, $\eta p2=.092$. The simple slopes analysis revealed a significant negative effect of

the RT TBW on Self-Location in the stationary condition (b = -0.011, CI [-0.021, -0.001], t = -2.138, p = .035). In contrast, the effect in the natural locomotion condition was insignificant (b = 0.004, CI [-0.002, 0.010], t = 1.237, p = .219).

The remaining measures did not exhibit significant moderation effects after Holm-Bonferroni correction: for Suspension of Disbelief, the interaction was significant (F(1,96)= 6.762, p=0.011). However, neither the stationary (b = 0.003, p = .456) nor natural locomotion conditions (b = -0.007, p = .098) were significant. The same pattern held for Domain Specific Interests (interaction: F(1,96)= 5.974, p=0.016), Spatial Situational Model (interaction: F(1,96)= 5.920, p= 0.017), Attention Allocation (interaction: F(1,96)= 4.459, p= 0.037), Perceived Possible Actions (interaction: F(1,96)= 3.385, p=0.069), and Visual Spatial Imagery (interaction: F(1,96)= 3.323, p=0.071). None of these measures' associated stationary or natural locomotion conditions reached significance (see Appendix 6).

The analysis revealed no significant moderation effects of VR condition between RT TBW and simulator sickness as measured by the SSQ subscales. For SSQ Disorientation, the interaction was insignificant (F(1,96)= 1.949, p=0.166, Wilks Λ= 0.980, ηp2= 0.020). While the interaction for SSQ Disorientation was not significant overall, simple slopes analysis indicated a significant positive effect of the RT TBW on disorientation in the stationary condition (b = 0.147, CI [0.038, 0.257], t = 2.674, p = .009). In contrast, the interactions for SSQ Nausea (F(1,96)= 0.832, p= 0.364, Wilks Λ=0.991, ηp2= 0.009) and SSQ Oculomotor (F(1,96)= 0.372, p=0.543, Wilks Λ= 0.996, ηp2= 0.004) were not significant, with no substantial effects identified in their corresponding stationary or natural locomotion conditions. These results suggest that the RT TBW significantly influences disorientation, specifically in stationary conditions, while no significant effects were observed for the other subscales of simulator sickness.

6.4 Discussion

The study in this chapter examined the relationship between temporal binding windows (TBWs), presence, and simulator sickness in a non-Euclidean virtual reality environment (VRE) titled *Tea for God*. Participants engaged in two versions of the same non-Euclidean experience. In the first version, the VR experience relied on natural locomotion, providing a congruent visual-vestibular environment. The second condition required participants to navigate the environment using a controller while seated, introducing a visual-vestibular mismatch. It was hypothesised that individuals with wider TBWs would experience more negative effects (reduced presence and increased simulator sickness) in the stationary condition and more positive effects (increased presence) in the natural locomotion condition. This hypothesis was informed by prior studies in this thesis, which demonstrated that sensory conflicts in VR environments tend to reduce presence for individuals with wider TBWs, while congruent sensory conditions enhance it.

The first hypothesis (H1: Individuals with wider TBWs will report lower levels of presence in *Tea for God* under conflicting sensory conditions compared to individuals with narrower TBWs) was supported. Specifically, there was a negative relationship between presence scores, measured by the MEC-SPQ, and individual TBW size in the stationary condition. The relationship identified here suggests that individuals with wider audio-visual TBWs are more likely to integrate conflicting cues, such as visual-vestibular mismatch, which consequently diminishes their sense of presence within the virtual environment.

The second hypothesis (H2: In non-conflicting sensory conditions within *Tea for God*, individuals with wider TBWs will report higher levels of presence compared to conflicting conditions, thus experiencing an enhanced VR experience when sensory cues are congruent) received partial support. While participants in the non-conflicting condition reported higher

levels of presence than those in the conflicting condition, there was no significant impact of TBW size on presence in the natural locomotion condition. Possible explanations for this finding are explored in further detail below.

The third hypothesis (H3: Non-Euclidean navigation in *Tea for God* will reduce simulator sickness, particularly for individuals with wider TBWs, compared to conflicting sensory conditions) was supported. Participants reported lower levels of simulator sickness in the natural locomotion condition. A significant relationship was also observed between TBW size and simulator sickness levels in the stationary condition. This suggests that individuals with wider TBWs are more prone to binding dissonant visual-vestibular cues, leading to increased simulator sickness.

The final hypothesis (H4: The MEC-SPQ will reveal significant differences across dimensions of presence, with individuals with wider TBWs showing distinct patterns in attention allocation, spatial situational models, and cognitive involvement depending on sensory consistency) received partial support. While attention allocation and spatial situational model sub-measures demonstrated significant moderation at a conventional significance level ($p < 0.05$), they did not remain significant after a Holm-Bonferroni correction for multiple comparisons. However, the relationship between TBW size and cognitive involvement showed significant moderation by the VRE condition, with individuals with wider TBWs in the stationary condition reporting lower cognitive involvement.

The findings from a four-way multivariate analysis and its subsequent post-hoc tests highlight a nuanced relationship between TBW size and VR user experience. This relationship affects both negative aspects (simulator sickness) and positive aspects (presence) of VR user experience. Notably, redundant-target TBWs demonstrated a stronger and more direct connection to user experience than temporal order judgment TBWs, a finding which has been

common throughout the studies reported in this thesis. Overall, this suggests that lower levels of temporal perception influence distinct aspects of user experience in VR.

6.4.1 *Multisensory Facilitation is More Important than Temporal Judgement in the Formation of Presence*

A significant insight from the research presented in this thesis is that multisensory facilitation plays a far more impactful role in immersive VR experiences than strict temporal order judgment. Each study differentiated between tasks that measured the integration of multisensory cues into a facilitated response (redundant target) and the precision of temporal judgement, essentially, how well a user can determine the order of stimuli (temporal order judgement) (Colonius & Diederich, 2018). These two tasks fundamentally differ in the level of cognitive involvement (Eskes et al., 2007; Miller & Schwarz, 2006). Although they are both argued to be related to the size of an individual's binding window, the temporal order judgment tasks require feedback from a higher level to make a judgment (Eskes et al., 2007). In comparison, the redundant target (RT) task focuses on speeded responses as a product of multisensory facilitation (Miller & Schwarz, 2006). The findings from this study and others within this thesis (except Study 2) have suggested that TBWs, as measured by the RT task, are more predictive of presence in VR than those measured by temporal order judgement (TOJ).

The idea that multisensory facilitation plays a more important role than the perception of temporal order aligns with another concept necessary for the development of presence: Allocation of Attention. Attention allocation is often argued to be the first step toward developing a sense of presence in a virtual environment (Biocca, 2003; Nunez, 2007; Slater, 2002; Wirth et al., 2007). In environments with multisensory stimuli, automatic attention allocation is thought to be improved by these stimuli reaching a response criterion level faster than single modality cues (Koelewijn et al., 2010; Talsma & Woldorff, 2005). The environment

is attended to quicker because the cues are integrated faster. In addition, Miller (1982) proposed that individuals' subjective criterion level differs based on task dependency. In tasks where multisensory facilitation is more important than the accuracy of temporal order judgment, the criterion level is typically lower (Colonius & Dietrich, 2018; Schwarz, 1989). A combination of lower subjective criterion levels in RT tasks and the multisensory facilitation thought to occur in attention allocation (Tang et al., 2017) may explain why only the TBW measured from RT tasks predict presence (e.g. Biocca, 2003; Nunez, 2007; Wirth et al., 2007).

While the relationships between presence and TBWs estimated from different tasks can be explained theoretically, decisions were also made at the analytical level, which might impact the results. The TWIN model of temporal binding has been used throughout the thesis because it provides a good fit for reaction time data (e.g. Diederich & Colonius, 2007; Glazebrook et al., 2016) and it can be expanded to account for data from other paradigms such as judgment tasks (e.g. Megevand et al., 2013). These advantages align with Colonius and Diederich's aim to provide a framework for the integration of multisensory cues rather than an exact calculation (2004, 2007). However, the TWIN model's application of a canonical model (Gibbon & Rutschmann, 1969) means that the data from the TOJ task is fitted to Gaussian functions, which coincides with the probability of integration outlined in the TWIN model's second processing stage (Diederich & Colonius, 2015). The fitting of individuals' TOJ data to Gaussian functions has been criticised due to the likelihood of visual first judgments plateauing across the middle range of SOAs (Alcalá-Quintana & García-Pérez, 2013). An issue which is more commonly found in individual data rather than group analysis. The application of the canonical model may have unintendedly reintroduced this issue in its attempt to reduce cross-model variance. As a result, the model may have introduced systematic biases in individual estimates of the TOJ TBWs, potentially affecting the interpretation of their relationship with presence by either

underestimating or overestimating variability in the midsection of trials, where visual-first judgments tend to plateau.

Whilst the difference between RT TBW and TOJ TBWs may be due to the limitations of applying a canonical model to the measurement of the two separate tasks, it is important to note that even in the first study where the independent-channels model by Alcalá-Quintana and García-Pérez (2013) was utilised, TOJ had no relationship with presence. Furthermore, while the extended twin model may only provide a descriptive nature of the window widths estimated by each task (Diederich & Colonius, 2015), it has been demonstrated to align with the estimations calculated from other models (e.g. Glazebrooke et al., 2016; Megevand et al., 2013). This suggests that even if other models were used to analyse the data from the TBW tasks, the direction and relationship between them and presence would remain consistent. Thus, indicating that windows estimated from the redundant target (RT) paradigm are more important in developing presence than those estimated from the temporal order judgment paradigm. Specifically, individuals who display multisensory facilitation (i.e. bind stimuli) over a large range of offset asynchronies are more impacted by sensory conflict in VRE.

6.4.2 *Individuals' Ability to Handle Sensory Conflict Influences Presence and Simulator Sickness*

This study's findings provide further evidence for the critical role that an individual's ability to handle sensory conflict can play in shaping user experience within virtual reality. Prior research into simulator sickness and presence has expressed the need for multisensory cues to be consistent (e.g. Donohue et al., 2013; Steuer, 1992; Wirth et al., 2007). In cases where multisensory cues related to a VRE are not consistent, theories have proposed that simulator sickness will increase (e.g. Reason & Brand, 1975) and presence will decrease (e.g. Wirth et al., 2007; Witmer & Singer, 1998). The research presented within this chapter supports

this conceptualisation whilst expanding it further. It is argued that multisensory conflict causes a negative user experience in VR, but it is amplified in individuals who are more likely to integrate incongruent sensory signals.

Individuals with a wider TBW are more likely to integrate incongruent sensory signals (Lohmann & Butz, 2017). Due to their tendency to bind inconsistent cues, they are likelier to experience simulator sickness and a reduced sense of presence in conflicting information conditions. This hypothesis was tested here through a "stationary" navigation condition. As the size of individual TBWs increased, so did their reports of simulator sickness. In addition, their self-reported presence levels decreased. On the other hand, the congruent sensory information produced by the natural locomotion condition demonstrated significantly higher presence scores in these individuals with wider TBWs. However, there was no significant relationship between presence and TBW in this condition.

The lack of a significant relationship between TBWs and presence in the natural locomotion condition could be explained by two key factors, both of which build on the findings from Study 3: the lack of asynchronous cues or a minor presence of conflicting cues. In the case of the lack of asynchronous cues, the relationship was predicted to be positive between TBW width and presence in the natural locomotion condition, as those with a wider window are better at maintaining a unified perception of a multisensory event, even when the cues are asynchronous. However, if the cues provided by the environment are synchronous, then the size of an individual's window should not influence presence. All the stimuli will be bound correctly regardless of window width (Colonius & Diederich, 2018; Stein & Meredith, 1993).

In contrast, if the information was asynchronous but contained minor conflicting cues, the studies presented here suggest that individuals with a wider window would experience more

simulator sickness and a reduced sense of presence. A minor conflict could be related to the visual scene update time (Kim et al., 2020). Even in the natural locomotion condition, there may be a slight delay in updating the visual information in the VRE as an individual moves their head (Kim et al., 2021). This slight delay is more likely integrated in individuals with a wider window, causing a sensory conflict. Whilst both proposals could be valid, the results suggest that the latter explanation might be more accurate. A small, non-significant relationship between simulator sickness and TBW size was identified in the natural locomotion condition. This finding suggests that other sensory conflicts (such as visual update time) may have been present and marginally influenced individuals with a wider TBW in the natural locomotion condition. However, these individuals still reported a significantly higher level of presence and lower level of simulator sickness in the natural locomotion condition.

By examining TBWs as an influencing factor on presence and simulator sickness, this study highlights the need for a more nuanced understanding: individuals with wider TBWs are inherently more likely to integrate conflicting sensory cues, a process that, paradoxically, undermines their overall VR experience. Further analysis was conducted to investigate what elements of presence might be specifically influenced.

6.4.3 *TBWs Influence Self-Location and Higher Cognitive Involvement*

The process model of presence highlights multiple stages that might occur for individuals to feel spatially present within a medium (Wirth et al., 2007). In virtual reality, these stages include the abovementioned attention allocation and seven other key categories. Domain-specific interests directly influence attention allocation because if an individual is interested in virtual reality, it should take fewer resources to maintain their attention (Hidi, 1995). Once an individual attends to the environment, they can form a Spatial Situational Model (SSM) of that environment. The SSM is formed from the spatial cues presented by the

environment (Vorderer et al., 2003) and directly relates to the concept of egocentric reference frames (Riecke & Von der Heyde, 2003). While the spatial cues influence the SSM from 'bottom-up', the next element, Visual Spatial Imagery, influences the SSM from 'top-down'. This element is tied to an individual's spatial ability, allowing them to "fill in the gaps" of the model from memory (Dean & Morris, 2003). The SSM then undergoes a perceptual hypothesis test, which outlines whether the egocentric reference frame can be considered the primary reference frame (Wirth et al., 2007). Two indirect processes support this: Suspension of Disbelief and Cognitive Involvement. The first relates to an individual's ability to ignore distractions from the virtual environment (Bystrom et al., 1999). At the same time, cognitive involvement relates to a decrease in the attentional process to the real world due to higher involvement with the virtual content (Wirth et al., 2007). Providing that these indirect processes support the hypothesis testing process, spatial presence forms. Characterised by feeling located in the VRE, Self-Location, and being able to act within it: Perceived Possible Actions.

The MEC-SPQ provided a way for this study to assess the impact of individuals' TBWs on each of the above components across the two VRE conditions. Due to the sample size not being sufficiently large to conduct a structural equational model containing the eight subscales (Westland, 2010), multiple comparisons with the appropriate correction (Holm-Bonferroni) were conducted instead. Suspension of disbelief, Domain Specific Interests, Spatial Situational Model, and Attention Allocation did not remain significant after the correction for multiple comparisons was made. Despite each subscale demonstrating significant moderation at a conventional significance level ($p < 0.05$). Perceived Possible Actions and Visual Spatial Imagery were also not significantly moderated. However, both Self-location and Higher Cognitive Involvement displayed a significant relationship with individuals' TBWs in the two VRE conditions. Both reported a negative relationship in the condition containing the conflicting cues.

These findings have important implications for the role of TBW in the formation of presence. When it comes to spatial presence, TBWs primarily impact users' sense of self-location in VR rather than their perception of possible actions within the virtual environment. In much of the presence literature, self-location, or the sense of "being" in the virtual space, is a fundamental component of presence (e.g. Slater et al., 2022; Wirth et al., 2007). The findings of this study suggest that individuals with wider TBWs experience greater disruptions to their sense of self-location, particularly in environments with conflicting sensory inputs, where their tendency to bind incongruent cues distorts their immersive experience.

When considering that the temporal integration of multisensory stimuli helps us orient to the environment surrounding us (Gramann et al., 2010; Klatzky, 1998; Stein & Meredith, 1993), it is potentially of little surprise that the element of spatial presence, which is most negatively affected by an individual's tendency to bind conflicting cues, is self-location. This understanding is further strengthened by the disorientation subscale of the SSQ, which demonstrated a significant positive relationship with TBWs in stationary condition. The conflicting cues presented in stationary VREs reduce individuals' ability to orient themselves within their environment if combined within an Egocentric Reference Frame. The integration of such cues leads to a sensory conflict, affecting both their levels of disorientation and self-location within the VRE.

6.4.4 *The Dual Role of TBW in VR User Experience*

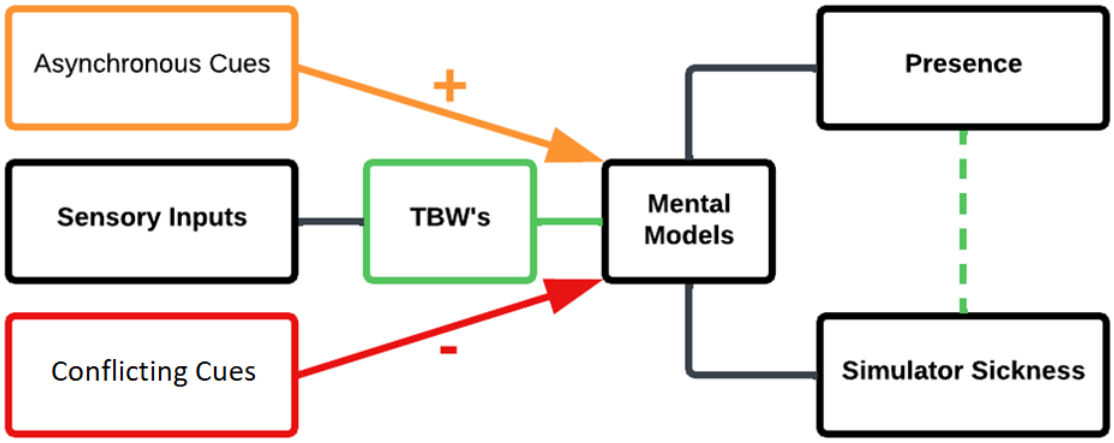
Head-mounted displays (HMDs) provide cues surrounding a virtual environment in a way that is not identical to reality (Tachibana & Matsumiya, 2022). Besides pixels replacing the natural stimuli of the world and headphones producing audio, the biggest difference is in the delay between stimuli. At best, the technology produces multisensory stimuli with a few milliseconds of delay (Bailey et al., 2004; Jung et al., 2000), referred to asynchronous in this

thesis. At worst, the technology produces sensory cues that conflict with other senses. This sensory conflict has been referred to as conflicting cues, such as the conflict between perceived vection and actual movement. At the same time, theories of presence have highlighted the need for technology to provide consistent multisensory cues (e.g. Slater et al., 2022; Wirth et al., 2007). Few have focused on individuals' role in maintaining this multisensory consistency (Kim & Lee, 2022; Marucci et al., 2021). At the point of writing this thesis, none have turned their attention to a potential mechanism that might account for the variance individuals display in maintaining these multisensory inconsistencies. The studies presented within this thesis put forward temporal binding windows, which explain part of this variance.

The role TBWs have within presence has been demonstrated to be dual-natured. Study 3 found that individuals with a wider window feel more present when multisensory information is presented without conflicting vestibular sensory cues. This was confirmed through a moderation analysis comparing uni and multisensory conditions. The study presented within this chapter served to confirm the second role that the TBW may play. Studies 1 and 2 indicated that when individuals with a wider window are presented with conflicting cues, they feel less present. The present study's findings confirmed this with a multivariate analysis comparing individuals' TBW width and their levels of presence in a congruent or conflicting VRE. Figure 16 demonstrates the dual nature that TBWs play in multisensory VREs and its impact on the feeling of 'being there' within them.

Figure 20

Proposed Model for the Dual Role of Temporal Binding Windows in the Formation of Presence and Simulator Sickness



The proposed model above highlights the positive and negative role individuals' TBW width has on their experience within VR. When information regarding a virtual environment is offset but not conflicting, those with a wider temporal binding window are more likely to maintain a unified percept of the events. This integration supports the constructed egocentric reference frame, referred to here as a mental model (Riecke & Von der Heyde, 2003). It also supports the hypothesis that this is the PERF, which leads to a stronger sense of presence in these environments (Riecke & Von der Heyde, 2003; Wirth et al., 2007). On the other hand, when information from the virtual environment is produced within an individual's window width but relates to two conflicting/incongruent events, these same individuals bind the conflicting cues. This binding of conflicting cues leads to a sensory conflict, undermining the medium-as-PERF hypothesis and increasing feelings of simulator sickness. These underlying principles of this model and its implications are developed further in the following chapter.

6.4.5 *Limitations and Future Research Directions*

Whilst this study was designed to mitigate the potential issue of using a between-subject design to investigate individual differences, the within-subject manipulation might have also impacted the results. Although counterbalancing was utilised, carry-over effects can be present in any repeated measure design (Greenwaldlo, 1976). The carry-over effects in this study may have been twofold. Individuals who experienced simulator sickness in the first condition may

not have fully recovered from all symptoms before entering the second condition despite the break to complete the TBW tasks (Kennedy et al., 2000). The data suggests that this may be true to some extent, with the average for all of the changes in simulator sickness scores in the natural locomotion condition being negative. The average participant left this VRE condition feeling less sick than they went in. Considering that individuals with a wider window reported larger increases in simulator sickness in both conditions, this carry-over effect might explain why individuals in the natural locomotive condition reported similar levels of presence regardless of TBW size.

Another dependent variable that could have been impacted by repeated testing is individuals' reported level of presence. As highlighted within the process model of presence, attention allocation and involvement are key elements toward feeling present in a VRE (Wirth et al., 2007). However, with individuals taking part in the same VRE twice, both factors might have diminished through the reduction of novelty (e.g. Karacan & Hayhoe, 2008). Controlled attention allocation might have, therefore, been reduced through familiarity with the content. At the same time, involvement may be reduced through individuals already interacting with some aspects of the VRE, with no need to repeat this in the second instance. Despite these potential impacts, presence was still significantly higher in the natural locomotion condition than in the stationary condition, as expected.

A second limitation of the present study lies in the sample. Despite the sample size being adequately powered to detect a four-way interaction between TBW, presence and simulator sickness between VRE conditions. Four post-hoc tests did not display a significant finding once adjusted for multiple comparisons. A solution to using multiple comparisons is to contain all variables within a single analysis, in this case, structural equational modelling (SEM). However, to conduct an SEM on the eight variables of the MEC-SPQ, the four subscales of the SSQ and individuals' TBWs, a minimum sample of 170 would be needed

(Westland, 2010). Due to the time limitations of this final study and the primary aims being sufficiently investigated through multivariate analysis, this was not achievable within the current research.

A final limitation concerns the direct adoption of a non-Euclidean VRE to facilitate natural locomotion (Suma et al., 2012). Given that non-Euclidean VR remains a relatively recent development, this approach relied on consistent but limited research suggesting that levels of presence in impossible spaces are comparable to those in standard Euclidean environments (e.g., Cherni et al., 2020; Lochner & Gain, 2021; Suma et al., 2012; Vasylevska & Kaufmann, 2015). The study's results support this, with levels of reported presence displaying a similar variance to those of the previous studies within the thesis. However, further insight into whether using a non-Euclidean VRE potentially impacted the results of some sub-measures (e.g., spatial situation models) remains unclear.

Future research could aim to solve these limitations. Multiple tests could be conducted on different days to reduce carry-over effects and potentially improve the strength of the moderation between conditions. The research could aim to include a much larger sample in order to investigate the structural relationship between each of the key variables. In addition, this future research may investigate the difference between non-Euclidean and Euclidean VRE with natural locomotion. Providing the laboratory space was big enough to allow for both, which was a limiting factor here.

A final suggestion for future research, further elaborated upon in the general discussion, is to propose solutions for the individual differences identified. While existing models of presence indicate that technology should strive to provide consistent multisensory cues (e.g., Wirth et al., 2007; Witmer & Singer, 1998), the research presented here highlights the challenges of replicating real-world experiences with technology. Therefore, a different

approach may be beneficial. Future studies could explore whether an individual's temporal binding window can be temporarily adjusted to enhance the virtual reality (VR) user experience. This could involve widening the window in asynchronous but non-conflicting situations or narrowing it in conditions where cues are contradictory. . This is discussed further in section 7.5.

6.4.6 *Summary and Concluding Remarks*

The study presented within this chapter aimed to provide confirmatory evidence that environments with conflicting cues negatively impacted individuals' sense of presence in line with the size of their temporal binding window. Utilising a repeated measures design, individuals took part in two non-Euclidean VR experiences, one with natural locomotion and one with motion controlled via handheld controllers. Individual TBW widths were estimated from a redundant target (RT) task and a temporal order judgment (TOJ) task, which they completed between VR experiences. After completing the VRE, individuals recorded their levels of presence on the MEC-SPQ and change in simulator sickness was measured via the SSQ. A significant four-way interaction was identified between the size of individuals' RT TBW, presence, and simulator sickness score between VRE conditions. Post hoc analyses identified a significant negative relationship between individuals' window size and self-location score within the stationary condition, driving the relationship alongside a significant positive relationship between window width and disorientation score in the stationary condition. The results of this study, alongside the other evidence from this thesis (study 3), suggest a dual role of individuals' TBW in the formation of presence. One allows individuals to bind asynchronous but relevant multisensory stimuli, and another separates incongruent stimuli into separate reference frames to avoid sensory conflict.

Chapter 7: General Discussion and Summary of Conclusions

7.1 Overview of Findings

This section summarises the key findings from Chapters 3-6, clarifying the relevance of each study to the overall development of the thesis. This serves as a foundation for discussion of the theoretical implications of TBWs in presence and simulator sickness. Table 17 provides a concise summary of the research questions, the data-derived answers, and additional implications/future research directions addressed in each chapter.

Table 17

Summary of Research Questions, Answers and Future Directions from the Present Thesis

Chapter Title	Main Research Questions	Answers Provided	Implications/ Future Directions
Chapter 3: Exploratory Analysis of Correlation between TBW and Presence	1. Does the size of an individual's TBW predict self-reported presence within a VR HMD? 2. Is the SPES a valid measure of individuals' Spatial Presence? 3. Are there any individual differences that confound a relationship between temporal binding and presence?	1. Yes: As window width decreased, individuals reported presence increased. 2. Yes: The SPES displayed good convergent validity with the IPQ (except realism). 3. No: There were no variables which correlated with both TBW and presence in this study.	• Future work should consider that individual differences in the ability to process multisensory information affect the development of presence in stationary games. • The SPES is a quick and easily deployed questionnaire which assesses spatial presence in a convergent manner with other long-standing presence questionnaires.
Chapter 4: Exploring the Relationship Between TBW, Presence and Simulator Sickness	1. Can the relationship between presence and TBWs be feasibly replicated online? 2. Does Simulator Sickness, TBWs and Presence correlate with each other? 3. Do individuals with a narrower TBW feel more present in environments with conflicting information?	1. Yes, but not without issues arising from online data collection. 2. Mixed results regarding correlations; SS was related to TBW in the pilot, but not in the full study. Presence was not correlated with SS in the pilot, but was in the full study. 3. Mixed results regarding TBW's: Due to a fault in the recording of data, windows widths were only generated for the TOJ task. This was only a predictor for the SUS rather than the SPES.	• Online data collection contains many issues, which prevents research questions from being fully answered. In person testing is considered the best approach for exploring the concepts of TBW and Presence. • Modern headsets may still pose the risk of simulator sickness. Further research should explore ways to reduce this and its impact on presence.
Chapter 5: Moderating Temporal Binding within Uni v Multisensory Environments	1. Does removing conflicting cues change the relationship between TBW, Presence and Simulator sickness? 2. Does the relationship between TBW and Presence only exist in Multisensory VRE's?	1. Yes, in a multisensory VRE with no conflicting cues, individuals who have a wider TBW felt more present. Simulator sickness did not correlate with either TBW or Presence. 2. Mixed results; Moderation analysis demonstrated a significant moderation between TBW and Presence based on the absence or presence of multisensory stimuli. However narrow windows reported more presence in the unisensory condition.	• Future research could replicate these findings using a range of non-conflicting multisensory VREs. Including other sensory modalities. • Individuals TBWs are related to presence, even in unisensory VREs.
Chapter 6: Individual Differences in Temporal Binding Windows moderate the Impact of Conflicting Vestibular Cues on Presence and SS.	1. Do individuals with a wider TBW feel less present in a vestibular conflicting environment? 2. Are wider TBW likely to experience more SS in cue-conflicting environments?	1. Yes, as TBW width increases presence decreases in the conflicting environment. There was no effect of window size on a non-conflicting environment. 2. Yes, as TBW width increases, SS increases in a conflicting environment. A significant four-way interaction confirmed the relationship between presence, SS, and TBW.	• Individual differences in processing conflicting sensory information are an important consideration to be made in theoretical models of Presence and SS • Future work could investigate the malleability of TBWs to see if they can be recalibrated for particular VREs.

7.1.1 *Chapter 3:*

Chapter 3 outlined the first exploratory study, which investigated whether individuals' temporal binding windows (TBW) were related to presence. Based on responses to flash-bleep tasks, individuals' TBWs were estimated in three separate ways across two experimental paradigms. Simultaneity Judgment (SJ) and Temporal Order Judgement (TOJ) tasks relate to an individual's ability to accurately distinguish between the two stimuli (flash/bleep) (Stevenson & Wallace, 2013). TBW width estimates for these tasks are calculated based on the participants' responses that indicate the stimuli were perceived as one percept (Parise & Ernst, 2016). On the other hand, TBWs estimated from Redundant Target (RT) tasks utilise the process of multisensory facilitation (a speeded response to perceptually combined stimuli) to determine the TBWs width (Colonius & Diederich, 2018). The study's main aim in Chapter 3 was to identify if a relationship existed between the size of an individual's TBW and their feelings of presence. Although the direction of this relationship was not predicted, it was hypothesised that an individual's ability to process multisensory stimuli would explain a part of the variance between individuals' levels of presence within VR (Slater et al., 2022; Wirth et al., 2007). The study's results provided evidence that individuals with a narrow TBW felt more present in a VR environment, which involved navigating through an open-world simulation using a handheld controller while remaining seated.

7.1.2 *Chapter 4:*

Chapter 4 presented two studies (Study 2a and Study 2b) that expand upon the results in Chapter 3. The studies were conducted online to replicate the findings of Study 1 in a higher-powered sample. Study 2a was a pilot conducted to see if data collected from online participants was comparable to that of the in-person study (Study 1). The majority of the data from the online group did not significantly differ from the in-person group, with the exception of

Perceived Possible actions (PPA) and past VR experience. The latter is explained by the requirement for online participants to have access to/own a VR headset. The difference in PPA could also be linked to the range of simulations available to the participants in the pilot. To reduce extraneous variables that might have influenced a difference in PPA, Study 2b used the exact same virtual reality environment (VRE) as Study 1 – rather than giving individuals a choice over the VRE.

A change in the coding language between the online pilot (Study 2a) and full data collection (Study 2b) was particularly detrimental to the analysis of Study 2. Data for most of the participants' responses to the RT task were not recorded, meaning TBWs could not be estimated. However, the TOJ data was recorded and negatively correlated with the Slater, Usoh and Steed (SUS) presence questionnaire (Slater et al., 1994). This was comparable to the results of Study 1, which found a negative relationship between the SPES and RT TBW width. Unlike Study 1, Studies 2a and 2b also included a measure of simulator sickness. Simulator sickness was included because it is known to have a relationship with presence (Weech et al., 2019) and TBWs (Sadiq, 2019). However, in this study, the relationship between simulator sickness and individual TBWs (as estimated by the TOJ task) did not reach a level of statistical significance. Although, simulator sickness was found to be a statistically significant predictor of presence (as measured by the SUS).

7.1.3 *Chapter 5:*

In Chapter 5, an attempt was made to isolate the TBWs' impact on feelings of presence by removing the open-world element of the VREs previously used. It was hypothesised that by providing a “room-scale” VRE, conflicting cues between the real world and the virtual world would be minimised (Ng et al., 2020). Once conflicting cues are removed, the difference between unisensory and multisensory environments allowed the researcher to directly test the

direction of the relationship between TBWs and presence. It was hypothesised that individuals with a wider temporal binding window could better maintain information in a multisensory VRE that did not have conflicting information. It was argued that head-mounted display (HMD) technology still produces issues with latency (lag/offset) when providing multisensory information (Kim et al., 2020; Bailey et al., 2004). Therefore, individuals with a wider window are less likely to notice the offset and maintain a coherent mental model of the VRE.

In Study 3, individuals with a larger temporal binding window were found to feel more present in multisensory VREs. This supports the idea that, in the absence of conflicting cues, these individuals can better maintain a mental model provided by VR compared to those with a narrower window width. Notably, measures of simulator sickness showed no significant correlations with presence or TBW variables, confirming that conflicting cues and, thus, the effects of simulator sickness were successfully minimised. Additionally, it's worth noting that individuals with a narrower window felt more present in a unisensory condition. This suggests that there are two processes occurring when it comes to multisensory integration and presence: a development process (related to correctly combining multimodal stimuli) and a maintenance process (related to the discerning of separate non-related stimuli).

7.1.4 *Chapter 6:*

Chapter 6 built on the foundations developed in the previous chapters to explore the two multisensory processes involved in the development of a mental model and, ultimately, presence. Whereas Chapter 5 isolated the mechanism that allows individuals to cope with offsets in the presentation of multisensory stimuli, Study 4, presented in Chapter 6, intended to investigate the mechanism which coped with conflicting information. Temporal binding windows are considered important in both cases, with one relating to the ability to combine offset (but related) stimuli into a single percept and then a mental model. The other case

considers the ability to accurately discern two separate conflicting events and not combine them.

Study 4, therefore, used a non-Euclidean VRE to provide either vestibular-visual consistent cues or vestibular-visual conflicting cues regarding vection. It was hypothesised that individuals with a wider window would likely combine the conflicting cues into a single percept, undermining their sense of presence and increasing levels of simulator sickness. The results supported this hypothesis and led to the development of a set of principles, provided in more detail below.

A significant insight from Study 4 relates to the continued contrast observed between TBWs estimated from Temporal Order Judgment (TOJ) tasks and those estimated by the Redundant Target (RT) tasks. While both assessments provide an estimate of the temporal binding window (TBW), they are fundamentally different in nature. TOJ tasks predominantly rely on higher-level cognitive judgments for determining the sequence of stimuli (Eskes et al., 2007; Miller & Schwarz, 2006). In contrast, RT tasks emphasise automatic, low-level multisensory facilitation, which leads to a more efficient integration of sensory inputs (Colonius & Diederich, 2018). Given that presence is influenced by automatic attentional processes and sensory integration (Wirth et al., 2007), TBWs derived from RT tasks appear to be more predictive of presence. This observation aligns with the findings of Studies 1, 2, and 4, which indicate that multisensory facilitation, rather than precise temporal ordering, plays a crucial role in shaping user experience within virtual environments.

Having summarised the key findings and potential contributions of each study, the full range of evidence presented from the studies will now be explored more uniformly to summarise the role that temporal binding windows may have in the two key elements of user experience in VR: presence and simulator sickness.

7.2 Rationale for Establishing New Principles

The discussion of presence in this thesis has been framed using the process model of presence proposed by Wirth et al. (2007). While this model is neither the most comprehensive nor the most up-to-date framework for understanding presence (Slater et al., 2022), it was selected for its ability to capture the shared elements across various theories (e.g. Biocca, 2003; Nunez, 2007; Schubert et al., 1999b; Slater & Steed, 2000). These shared elements outline the stages of presence formation: attending to sensory cues in the virtual environment (e.g. Nunez, 2007; Schuemie et al., 2001; Slater & Steed, 2000); constructing a mental model of that environment (e.g. Biocca, 2003; Slater & Usoh, 1993); and perceptually testing the accuracy and relevance of that model (e.g. Slater, 2002; Wirth et al., 2007). The sense of presence relies on these interconnected processes (Slater et al., 2022).

Alongside the interconnected process outlined above, the consistency of multisensory information provided by the technology has long been recognised as critical for inducing presence (Steuer, 1992; Wirth et al., 2007). Early work by Witmer and Singer (1998) emphasized the importance of sensory consistency, though they primarily referred to aligning sensory cues between the virtual and real worlds, placing the responsibility on technology and design. More recently Slater and colleagues (2022) have expressed how interconnected the place illusion and sensorimotor contingencies are, with an emphasis on possible individual differences impacting feelings of 'being there' within a virtual world.

Although the process model provides a useful foundation, it has limitations in addressing these individual differences in the ability to process sensory information. While Wirth et al. (2007, p 504) acknowledged that "*internal processes based on user characteristics also affect the construction and quality of the SSM*", they do not propose any particular mechanism which might vary between individuals. While it was not the intent of the process

model to provide a completely comprehensive theory of every process, which does lead to presence, it does leave room for further expansion. Conversely, Nunez's (2007) Capacity Limited Cognitive Constructionist (CLCC) model of presence highlights a mechanism that processes sensory information into a mental model, referred to as 'Memory Modules' in the CLCC, known as the stimulus attenuator. However, this model does not differentiate between asynchronous and incongruent stimuli, nor does it explain how individuals vary in their responses to these two distinct forms of inconsistent sensory cues. Instead, the model argues that "*incongruent stimuli are less likely to be allocated working memory capacity*" (Nunez, 2007, p135). A suggestion which would not explain the findings presented here within this thesis.

Instead, a set of principles are proposed to build upon, rather than replace, existing theories of presence. These principles could be adapted and tested within any theoretical framework for understanding the formation of presence. Aligning with Slater et al.'s (2022) concept of place illusion, spatial presence is viewed as a spectrum ranging from 0 (no presence) to 1 (complete presence), depending on sensorimotor contingencies. For instance, if all sensory inputs in a virtual environment were identical to those of the real world, the expected place illusion would reach a value of 1. However, Slater and colleagues emphasise that individual differences must be considered when measuring spatial presence, as these differences contribute to variations beyond what sensorimotor contingencies alone can predict (2022). The proposed principles introduce individual differences in temporal binding windows (TBWs) as a key factor in explaining this variation.

7.3 The Finalised Set of Principals

In the natural world, cross-modal sensory information originating from a single event often arrives with slight delays, requiring a time window of integration to maintain a coherent

perception of events (Poppel, 1988; Sugita & Suzuki, 2003). This integration enables multisensory events to be experienced as unified (Stein & Meredith, 1993). Individuals differ in their perceptual thresholds for these delays—referred to as TBWs—which can influence their experience of the same virtual environment (Colonius & Deiderich, 2018; Stevenson & Wallace, 2013, Kim et al. 2021). Those with wider TBWs may tolerate greater asynchrony between sensory inputs, while those with narrower TBWs may require stricter alignment for successful integration (e.g. De Gelder & Bertelson, 2003).

The proposed set of principals incorporates a spectrum of sensory cue consistency in virtual environments, categorised as consistent (aligned with the event), asynchronous (temporally offset), or conflicting (incongruent). When cues are consistent, individuals engage with the virtual world as they would with the real one, as sensorimotor contingencies align with their natural expectations. While current VR technology has not yet achieved perfect sensory consistency, and may never do so, it often provides stimuli close enough to real-world timing for most users to feel present and experience minimal simulator sickness. However, as the offset between stimuli increases, individual differences in sensory processing become more evident. The three sensory principles are:

The Consistent Cue Principle: In virtual environments, when sensory information from different modalities related to an event is presented simultaneously, it creates a clear mental model of the environment. This occurs regardless of the width of an individual's TBW. As a result, individuals tend to report a strong sense of presence and low levels of simulator sickness.

The Asynchronous Cue Principle: In virtual environments that present sensory information from different modalities with offsets from one another, an individual's TBW influences the formation of their mental model. Those with a wider TBW are able to maintain a perception of synchrony between sensory inputs and develop a clear mental model of the environment. These

individuals report a strong sense of presence and experience low levels of simulator sickness. Conversely, individuals with a narrower TBW tend to separate sensory information into two distinct mental models. While they still can feel a sense of presence in one of these models, this feeling is diminished. However, their levels of simulator sickness remain low and are not affected by the dual representations of the environment.

The Conflicting Cue Principle: In virtual environments where sensory information from different modalities is incongruent, an individual's TBW affects their ability to create distinct mental models free from sensory conflict. Individuals with a wider TBW tend to incorporate conflicting cues into a compromised mental model. As a result, they often report a significantly diminished sense of presence and an increased feeling of simulator sickness. Conversely, individuals with a narrower TBW are able to separate this sensory information into two distinct mental models. By selecting one of these models, they can still experience a sense of presence similar to that of environments with asynchronous cues. However, the existence of a conflicting alternative mental model may lead to a slight increase in feelings of simulator sickness.

7.4 Understanding Findings through the Sensory Principles

The following section utilises the set of principles to describe the findings of each study within this thesis and includes examples of other recent studies to explain elements of their results.

7.4.1 *The Principles Applied to Studies 1-4*

While the study conducted in Chapter 3 was exploratory, the relationship between temporal binding and presence appeared to be somewhat counterintuitive. Theories of presence highlight the need for multisensory consistency, yet individuals more likely to notice inconsistencies went on to feel more present. The principles highlighted above, along with the findings from Studies 2a, 2b and 4, suggest that this relationship may be linked to the VRE in

Study 1 containing incongruent/conflicting cues. The ‘conflicting cue principle’ highlights the potential perceptual process that occurred within the individuals in this study. The Minecraft environment used in study 1 involved individuals moving throughout an open world whilst sitting stationary. The principle suggests that in this situation, individuals with a narrow window are more likely to separate the two representations. A visual representation of the world being navigated and a vestibular/proprioceptive representation of being seated in a lab room. Individuals still have the potential to feel present in the visual representation (as evidenced in Study 1 from an average SPES score of 29.54). Still, there exists a higher chance of breaking presence due to another representation being available (a mental model of the lab room/chair being sat on). On the other hand, individuals with a wider window, i.e. those more likely to bind offset/irrelevant sensory cues, are thought to attribute elements from both modalities into a single representation. This causes an undermined mental model, leading to a reduction in presence.

Study 2 closely replicated the findings of Study 1, whilst also examining simulator sickness. Using the same VRE, which contained incongruent cues, the results are supported by the potential processes occurring within the ‘conflicting cue principle’. Individuals with a wider TBW struggled to accurately discern separate representations, resulting in a diminished sense of presence. This suggests that wider TBWs facilitated the integration of conflicting sensory information, weakening the mental model of the environment. However, given the exploratory nature of the study and issues surrounding data collection, a larger and more confirmatory study was necessary to investigate these relationships further. Study 4 addressed this need.

The findings of Study 4 strongly align with the proposed principles. This study compared a condition with deliberately conflicting cues to a matched cue condition. In the conflicting condition, individuals with narrower TBWs reported higher levels of presence. This can be explained via perceptual processes occurring within the ‘conflicting cue principle’. As

these individuals are potentially better at distinguishing conflicting sensory cues, they may reduce the sensory conflict within the mental model of the environment by separating the cues into distinct representations. In contrast, those with wider TBWs, who are more likely to integrate these conflicting cues, experienced reduced presence and higher levels of simulator sickness.

Conversely, the natural locomotion (matched cue) condition can be explained through either the ‘consistent cue principle’ or a combination of 'asynchronous cue' and 'conflicting cue' principles. If the consistent cue principle is correct, it suggests that the sensory cues provided by the technology and virtual environment aligned with the individuals' sensorimotor contingencies to maintain a coherent mental model. However, as highlighted previously, current VR technology remains limited by latency and lag (Stauffert et al., 2020; Tachibana & Matsumiya, 2022), preventing full consistency.

In such cases, it is possible that incongruent cues, such as visual update time, could interact with asynchronous cues, like delays in the audio-visual synchrony of an event. For individuals with wider TBWs, tolerance for delays may mitigate the effects of asynchronous cues but result in reduced presence due to sensory conflicts in the mental model. In contrast, those with narrower windows may experience the opposite effect. The proposed principles suggest that these opposing influences balance out, leading to the observed similarity in presence scores between groups in the matched cue condition. Untangling these effects would be challenging if not for the results of Study 3.

The asynchronous cue principle explains the results of Study 3, which intentionally removed conflicting cues by simulating a simple room-scale office cubicle. Unlike the environment in Study 4, this setting involved no locomotion or impossible spatial

configurations, keeping visual update time to a minimum and allowing for the exploration of naturally occurring asynchrony compared to a unisensory condition.

In the multisensory condition, individuals with larger temporal binding windows exhibited significantly higher presence scores, which aligns with the asynchronous cue principle. As noted, the technology currently available for simulating virtual realities cannot fully replicate the complexity of reality. Offsets between stimuli are likely to occur, and individuals who can manage these offsets tend to feel more present. However, asynchronous information is unlikely to lead to simulator sickness, regardless of window width, because the information is not in conflict. The asynchronous cue principle captures this, and the results showed no relationship between simulator sickness and TBW or presence.

Additionally, the small negative relationship between presence and TBW width in the unisensory condition may be partially explained by the conflicting cue principle; although the virtual reality environment (VRE) provided only visual information, the external environment still contributed sensory information. While this conflict is not as severe as a conflict between visual and vestibular information, which involves a greater intensity of stimuli, it might have caused individuals with the widest windows to combine redundant signals from the real world, slightly diminishing their sense of presence.

7.4.2 *The Principles in the Context of Recent Literature*

Beyond the scope of the studies which led to the formation of the principles, it is important to consider how well these principles can account for findings that have been published outside of their conceptualisation.

Gim et al. (2024) recently developed a VR in-car simulator that navigated a straight path or a route with minimal to significant swerves to investigate the effect of redirected

steering. The participants were placed in a car in the real world which drove a straight route, matching the velocity of the VR simulation they were displayed.

In the straight-line condition, the perceived possible actions were severely limited because the participant did not need to move the steering wheel. This lack of interaction negatively affected the overall presence score for the straight-line condition, as the PPA subscale received low ratings.

Removing this effect from the comparison and focusing on the Self-location element of the SPES scores, Gim and Colleagues demonstrate that as the virtual swerves deviated more from the real straight path, the severity of Simulator Sickness Questionnaire (SSQ) scores increased, and the sense of presence decreased (2024). This finding suggests that the matched sensory cue condition (Straight line) aligned with the consistent cue principle, indicating sensory consistency. In contrast, the condition with the greatest sensory conflict (The largest swerves in the virtual route) appears to align with the conflicting cue principle, characterised by the lowest self-location awareness and the highest SSQ severity. While no individual differences were recorded, the standard deviation indicates a degree of variability that may be related to the individual differences in TBWs in the conflicting cue principle.

The principles can also be used to explain the variability of the results in a study investigating design effects on presence and simulator sickness in VR exhibitions. Kim et al. (2024) utilised various navigation techniques to measure spatial presence using the Igroup Presence Questionnaire (IPQ) and simulator sickness with the SSQ. Their study involved three different navigation methods: Joystick, Teleport, and Redirected Walking (RDW) across two types of environments: static (where NPCs did not move) and dynamic (where NPCs did move). The results indicated that RDW achieved the highest scores in both IPQ and SSQ across the two conditions. At the same time, Teleporting recorded the lowest SSQ in both conditions,

although the IPQ scores varied. The Joystick condition recorded scores in between the two other conditions with regard to both measures, noting a significant amount of variability in each.

In their discussion, Kim and Colleagues (2024) highlight that participants using the joystick verbalised experiencing the highest levels of simulator sickness. The conflicting cue principle conceptualised here clearly accounts for the simulator sickness (SSQ) and presence (IPQ) in the joystick condition. In contrast, a combination of the consistent cue and asynchronous cue principles appears to explain the Teleport condition and the observed variations.

Interestingly, the RDW condition resulted in increased scores for both SSQ and IPQ. This variability can be attributed to individual differences interacting with the combined asynchronous cue and conflicting cue principle. When stimuli are simply offset (as in normal walking), participants predominantly abide by the asynchronous cue principle. However, once the physical boundaries are reached, the introduction of incongruent stimuli for a brief period leads to individual differences experienced in line with the conflicting cue principle. This is due to the virtual reality subtly rotating the room to allow for further exploration. This visual discrepancy is argued to be imperceivable (Razzaque et al., 2002); however, in applying the set of principles to Kim and colleagues (2024) study, it is suggested that the visual discrepancy was perceived as incongruent, if only for short periods of time. This is particularly relevant to simulator sickness, which will likely build over time.

The two studies reported as examples above are far from an exhaustive list of all the studies in which the principles can be used to help explain variability. However, they are some of the most recently published and relevant studies with their focus on attempting to investigate locomotion techniques to reduce simulator sickness and increase presence.

7.5 Current Limitations of the Proposed Principles and Future Directions

Considering that the proposed principles are based on the studies presented in this thesis, it is important to note some potential limitations regarding there current applications. Although the principles are suggested to be adaptable to various theories of presence, the connection between temporal binding windows (TBWs) and other factors that contribute to variability in presence has not been thoroughly investigated. One such factor is the allocation of attention to the environment. Throughout this thesis, it has been highlighted that TBWs have a complex relationship with attention (e.g. Talsma & Woldorff, 2005).

While the final study employed an attention allocation subscale of the Measurement, Effects, Conditions, Spatial Presence Questionnaire (MEC-SPQ) (Vorderer et al., 2004) to explore and capture aspects of this relationship, it did not assess an individual's ability to focus on stimuli. This ability can vary significantly between individuals (Talsma et al., 2010) and likely plays a unique role in the relationship between TBWs and presence in virtual reality (VR). Further research could explore this relationship, examining how individual differences in TBWs and attention interact to influence the experience of presence.

A second limitation of the proposed principles arises from the existing literature on TBWs. It is acknowledged that the conceptualisation of TBW is broad and varied within the literature (e.g. Colonius & Diederich, 2018; Miller, 2016; Parise & Ernst, 2016), and this thesis only focuses on a specific aspect. Although an attempt was made to apply a canonical model that addresses two different paradigms in the literature (Redundant Target and Estimation/Judgment), there remains a significant amount of literature that does not fit this framework. This includes works that deny the existence of temporal integration (e.g., Stoffregen et al., 2017), diverse conceptualisations of a temporal window (see Wilson, 2023

for a recent review), and studies on the neuroscientific structures related to temporal integration (e.g., Wolff et al., 2022).

Furthermore, this thesis conceptualises the TBW primarily through the measurement of audio-visual integration. Although prior research has demonstrated correlations within individuals' TBWs across multisensory pairs, such as audio-visual, visual-tactile, and audio-tactile (Noel et al., 2015), it remains unclear whether visual-vestibular TBWs align similarly with audio-visual TBWs in individuals. Thus, the relationship between visual-vestibular and audio-visual integration represents a noteworthy gap within current TBW conceptualisations, highlighting an important opportunity for future investigation.

However, the primary aim of this thesis is not to explore the complexities of conceptualising TBWs but rather to provide insight into how individual differences in perception affect feelings of presence. By utilising the TWIN model (Colonius & Diederich, 2018) to estimate an individual's temporal binding window, the principles presented can effectively explain variations in a person's sense of presence when exposed to different sensory experiences.

While the principles are defined within the current scope, it is hoped that further investigation, along with the TWIN model's simplicity as a framework rather than a method for precise temporal binding estimates, will encourage future research to expand the applicability of these principles to various conceptualisations of temporal binding.

A final limitation and suggestion for future research arises from the overlap between the different principles. It has been conceptualised that experiences within VR can be attributed to environments that provide either a single category of sensory cues (i.e., all matched or all incongruent) or a combination of cues (e.g. conflicting and asynchronous). However, this

assumption does not address whether individuals become more or less susceptible to the various sensory inputs as time progresses.

Research into temporal integration has demonstrated that, while there are individual differences in the size of TBWs, these windows are malleable. Some studies have reported changes in the size of these windows based on perceptual training (e.g., Stevenson et al., 2013), while others have observed rapid recalibration depending on task demands (e.g., Van der Berg et al., 2018).

Although the outcomes of the cue combinations remain consistent with what is described in the principles, individual differences in temporal recalibration may introduce a greater degree of variability. Some individuals may have a stronger tendency to adjust their temporal binding window when exposed to virtual stimuli. Therefore, among the suggestions listed, this area represents a particularly important direction for future research. Understanding whether individuals can adjust their temporal binding window within virtual reality based on the type of sensory cues presented in the environment could help reduce simulator sickness and enhance the sense of presence for all users.

One promising direction for future research would be to investigate the ethical considerations and practical implications associated with either widening or narrowing individuals' temporal binding windows (TBWs). Previous studies aimed at training the temporal binding window have predominantly concentrated on the narrowing individuals' TBWs through experiments that provide feedback on their judgments of stimulus offset (Powers et al., 2009, 2016). Research has shown that individuals can maintain a sustainably narrow TBW over prolonged durations (Stevenson et al., 2013) and can also adjust their TBWs swiftly on a trial-by-trial basis (Van der Berg et al., 2018). In the context of virtual reality, where visual-vestibular conflicts arise, the narrowing of the TBW could prove particularly

advantageous in alleviating simulator sickness and enhancing users' immersive experiences. Suggested study designs to further investigate this phenomenon could encompass multi-day experimental frameworks, which would integrate baseline TBW measurements, pre- and post-training VR sessions featuring conflicting cues, as well as structured training sessions aimed specifically at narrowing the TBW.

The malleability of the temporal binding window in the alternative direction—namely, its widening—remains largely unexplored, likely due to valid ethical and practical considerations. Narrower windows are associated with enhanced accuracy in temporal judgments within real-world scenarios, whereas wider windows tend to be characteristic of neurodivergent or clinical populations (e.g. Foss-Feig et al., 2010). Nonetheless, if it were feasible to temporarily widen an individual's TBW specifically during VR experiences, this could mitigate the effects of latency and lag, consequently improving the depth of presence experienced by users. Investigating the possibilities for short-term, controlled widening of the TBW may yield significant advancements in VR design and user experience. Together, the examination of temporal recalibration in both directions offers substantial opportunities for the practical implementation of the principles outlined in this thesis within practical, real-world VR applications.

7.6 Conclusion

The research presented in this thesis goes beyond the main objective outlined in Chapter 2, which aimed to determine if there is a relationship between temporal binding widths and the feeling of presence in virtual reality (VR). Through a series of empirical studies summarised at the beginning of this chapter (see Section 7.1, Table 17), a set of principles has been proposed to explain the relationship between temporal binding windows, sense of presence, and

simulator sickness, all of which depend on the sensory information provided by the VR technology.

These principles categorise sensory information into consistent, asynchronous, and conflicting cues, providing a framework for understanding variability in VR user's experience of presence and simulator sickness. Environments which provide consistent cues enable all users, regardless of TBW width, to form coherent mental models, leading to a strong sense of presence and minimal simulator sickness. Environments which contain asynchronous cues highlight individual differences in sensory processing. Individuals with wider TBWs tolerate greater delays and maintain a sense of presence, while narrower TBWs perceive distinct representations, slightly diminishing presence without increasing simulator sickness. When conflicting cues are present within the VRE, they demonstrate the limitations of sensory integration, where individuals with a wider TBW combine incongruent information, leading to reduced presence and higher sickness. Individuals with narrower TBWs separate the cues, persevering a sense of presence at the cost of a minor increase in sickness.

The principles can be utilised to explain varying user experiences within different virtual realities. They account for phenomena such as the heightened sense of presence in individuals with narrower TBWs when experiencing scenarios which contain conflicting cues (e.g. Studies 1, 2, & 4), and the mitigating role of wide TBWs in asynchronous environments (e.g. Study 3). Additionally, they account for the overall variability in user responses to VR experiences that involve a wide range of sensory inconsistencies introduced through various locomotive methods (e.g. Gim et al., 2024; Kim et al., 2024). This framework establishes a basis for future research aimed at minimising simulator sickness and improving presence by utilising insights into individual differences in cross-modal temporal binding.

Future theoretical accounts of VR user experience will need to take into account individual differences in multisensory processing. It is argued that this is best done by understanding the three main types of multisensory cues which a VR environment may provide and understanding the principles of how the size of an individual's TBW might interact with them. This can help predict and explain the variability in the user experience of presence and simulator sickness. The research presented within this thesis demonstrates that consistent, asynchronous, and conflicting cues elicit distinct interactions with TBWs, shaping user experiences in separate ways. By incorporating these principles into theoretical accounts of VR user experience, developers and researchers can create environments that are not only more inclusive but also tailored to accommodate differences in individual's cross-modal temporal binding. While these findings represent an initial step, further empirical work is needed to refine these principles and explore their application in more complex and diverse virtual environments.

References

Abraham, W. T., & Russell, D. W. (2008). Statistical power analysis in psychological research. Social and Personality Psychology Compass, 2(1), 283-301.

Alais, D., & Burr, D. (2004). Ventriloquist Effect Results from Near-Optimal Bimodal Integration. Current Biology, 14(3), 257–262. https://doi.org/10.1016/S0960-9822(04)00043-0

Al-Ashwal, W., Asadi, H., Mohamed, S., Alsanwy, S., Kooijman, L., Nahavandi, D., ... & Nahavandi, S. (2021, October). Cybersickness measurement and evaluation during flying a helicopter in different weather conditions in virtual reality. In 2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (pp. 2152-2157). IEEE. doi: 10.1109/SMC52423.2021.9659215.

Alcalá-Quintana, R., & García-Pérez, M. A. (2013). Fitting model-based psychometric functions to simultaneity and temporal-order judgment data: MATLAB and R routines. Behavior Research Methods. https://doi.org/10.3758/s13428-013-0325-2

Aldoma, A., Marton, Z. C., Tombari, F., Wohlkinger, W., Potthast, C., Zeisl, B., ... & Vincze, M. (2012). Tutorial: Point cloud library: Three-dimensional object recognition and 6 dof pose estimation. IEEE Robotics & Automation Magazine, 19(3), 80-91.

Angelov, V., Petkov, E., Shipkovenski, G., & Kalushkov, T. (2020, June). Modern virtual reality headsets. In 2020 International congress on human-computer interaction, optimization and robotic applications (HORA) (pp. 1-5). IEEE.

Anwyl-Irvine, A., Dalmaijer, E. S., Hodges, N., & Evershed, J. K. (2021). Realistic precision and accuracy of online experiment platforms, web browsers, and devices. Behavior research methods, 53(4), 1407-1425.

Arcioni, B., Palmisano, S., Apthorp, D., & Kim, J. (2019). Postural stability predicts the likelihood of cybersickness in active HMD-based virtual reality. Displays, 58, 3-11.

Aymerich-Franch, L. (2010). Presence and Emotions in Playing a Group Game in a Virtual Environment: The Influence of Body Participation. Cyberpsychology, Behavior, and Social Networking, 13(6), 649–654. https://doi.org/10.1089/cyber.2009.0412

Bailenson, J. N., Blascovich, J., Beall, A. C., & Loomis, J. M. (2003). Interpersonal Distance in Immersive Virtual Environments. PERSONALITY AND SOCIAL PSYCHOLOGY BULLETIN, 29(10), 1–15. https://doi.org/10.1177/0146167203253270

Bailey, R. E., Arthur III, J. J., & Williams, S. P. (2004, August). Latency requirements for head-worn display S/EVS applications. In *Enhanced and Synthetic Vision 2004* (Vol. 5424, pp. 98-109). SPIE.

Bald, L., Berrien, F. K., Price, J. B., & Sprague, R. O. (1942). Errors in perceiving the temporal order of auditory and visual stimuli. Journal of Applied Psychology, 26(3), 382–388. https://doi.org/10.1037/h0059216

Balk, S. A., Bertola, D. B., & Inman, V. W. (2013, June). Simulator sickness questionnaire: twenty years later. In Driving Assessment Conference (Vol. 7, No. 2013). University of Iowa.

Bannigan, G. M., de Sousa, A. A., Scheller, M., Finnegan, D. J., & Proulx, M. J. (2024). Potential factors contributing to observed sex differences in virtual-reality-induced sickness. Experimental brain research, 242(2), 463-475.

Baños, R. M., Botella, C., Rubió, I., Quero, S., García-Palacios, A., & Alcañiz, M. (2008). Presence and emotions in virtual environments: The influence of stereoscopy. CyberPsychology & Behavior, 11(1), 1-8.

Baren, J. van, & IJsselsteijn, W. (2004). Measuring Presence: A Guide to Current Measurement Approaches. Measurement, 0, 1–86. http://www.mendeley.com/research/measuring-presence-guide-current-measurement-approaches/%0Ahttps://ispr.info/about-presence-2/tools-to-measure-presence/omnipres-guide/

Barreda-Ángeles, M., Aleix-Guillaume, S., & Pereda-Baños, A. (2020). Virtual reality storytelling as a double-edged sword: Immersive presentation of nonfiction 360°-video is associated with impaired cognitive information processing. Communication Monographs, 0(0), 1–20. https://doi.org/10.1080/03637751.2020.1803496

Basharat, A., Adams, M. S., Staines, W. R., & Barnett-Cowan, M. (2018). Simultaneity and temporal order judgments are coded differently and change with age: an event-related potential study. Frontiers in integrative neuroscience, 12, 15.

Battaglia, P. W., Jacobs, R. A., & Aslin, R. N. (2003). Bayesian integration of visual and auditory signals for spatial localization. Journal of the Optical Society of America. A, Optics, Image Science, and Vision, 20(7), 1391–1397. http://www.ncbi.nlm.nih.gov/pubmed/12868643

Beauchamp, M. S. (2005). See me, hear me, touch me: Multisensory integration in lateral occipital-temporal cortex. Current Opinion in Neurobiology, 15(2), 145–153. https://doi.org/10.1016/j.conb.2005.03.011

Biocca, F. (1997). Is this body really "me"? Self presence, body schema, self-consciousness, and identity. *The Cyborg's Dilemma: Progressive Embodiment in Virtual Environments, JCMC, 3*(2), 295-302.

Biocca, F. (2003). Can we resolve the book, the physical reality, and the dream state problems? From the two-pole to a three-pole model of shifts in presence. EU Future and Emerging Technologies, Presence …. http://www.mindlab.org/images/d/DOC705.pdf TS - BibTeX%5Cnhttp://www.mindlab.org/images/d/DOC705.pdf

Biocca, F., & Levy, M. R. (1995). Communication Applications of Virtual Reality. Communication in the Age of Virtual Reality.

Biocca, F., Kim, T., & Levy, M. R. (1995). The vision of Virtual Reality. In Communication in the Age of Virtual Reality (Routledge Communication Series).

Boldt, M., Bonfert, M., Lehne, I., Cahnbley, M., Korschinq, K., Bikas, L., ... & Smeddinck, J. (2018, March). You shall not pass: Non-intrusive feedback for virtual walls in VR environments with room-scale mapping. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (pp. 143-150). IEEE.

Bonato, F., Bubka, A., Palmisano, S., Phillip, D., & Moreno, G. (2008). Vection change exacerbates simulator sickness in virtual environments. Presence: Teleoperators and Virtual Environments, 17(3), 283-292.

Bonett, D. G., & Wright, T. A. (2000). Sample size requirements for estimating Pearson, Kendall and Spearman correlations. *Psychometrika, 65,* 23-28.

Burr, D., Silva, O., Cicchini, G. M., Banks, M. S., & Morrone, M. C. (2009). Temporal mechanisms of multimodal binding. Proceedings of the Royal Society B: Biological Sciences, 276(1663), 1761–1769. https://doi.org/10.1098/rspb.2008.1899

Bystrom, K. E., Barfield, W., & Hendrix, C. (1999). A conceptual model of the sense of presence in virtual environments. *Presence: Teleoperators & Virtual Environments, 8*(2), 241-244.

Calvert, G. A., & Thesen, T. (2004). Multisensory integration: Methodological approaches and emerging principles in the human brain. Journal of Physiology Paris, 98(1-3 SPEC. ISS.), 191–205. https://doi.org/10.1016/j.jphysparis.2004.03.018

Caserman, P., Garcia-Agundez, A., Gámez Zerban, A., & Göbel, S. (2021). Cybersickness in current-generation virtual reality head-mounted displays: Systematic review and outlook. *Virtual Reality, 25*(4), 1153-1170.

Ceunen, E., Vlaeyen, J. W., & Van Diest, I. (2016). On the origin of interoception. Frontiers in psychology, 7, 743.

Chang, E., Kim, H. T., & Yoo, B. (2020). Virtual reality sickness: a review of causes and measurements. International Journal of Human–Computer Interaction, 36(17), 1658-1682.

Chen, A., DeAngelis, G. C., & Angelaki, D. E. (2010). Macaque parieto-insular vestibular cortex: responses to self-motion and optic flow. Journal of Neuroscience, 30(8), 3022-3042.

Chen, Y.-C., Dong, X., Hagstrom, J., & Stoffregen, T. A. (2011). Control of a virtual ambulation influences body movement and motion sickness. Proceedings of the BIO Web of Conferences, 1, 1–3. https://doi. org/10.1051/bioconf/20110100016

Cherni, H., Métayer, N., & Souliman, N. (2020). Literature review of locomotion techniques in virtual reality. *International Journal of Virtual Reality*.

Cohen, J. (1992). Statistical power analysis. Current directions in psychological science, 1(3), 98-101.

Colavita, F. B. (1974). Human sensory dominance. Perception & Psychophysics, 16(2), 409–412. https://doi.org/10.3758/BF03203962

Coleman, T., Branch, M. A., & Grace, A. (1999). Optimization toolbox. For use with MATLAB. User's guide for MATLAB, 5.

Colonius, H. (1990). Possibly dependent probability summation of reaction time. Journal of Mathematical Psychology, 34(3), 253–275. https://doi.org/10.1016/0022-2496(90)90032-5

Colonius, H., & Diederich, A. (2004). Multisensory interaction in saccadic reaction time: A time-window-of- integration model. Journal of Cognitive Neuroscience. 16(6), 1000–1009. https://doi.org/10.1162/0898929041502733

Colonius, H., & Diederich, A. (2018). Formal models and quantitative measures of multisensory integration: A selective overview. European Journal of Neuroscience, December. https://doi.org/10.1111/ejn.13813

Cooper, N., Milella, F., Pinto, C., Cant, I., White, M., & Meyer, G. (2018). The effects of substitute multisensory feedback on task performance and the sense of presence in a virtual reality environment. PloS one, 13(2), e0191846. https://doi.org/10.1371/journal.pone.0191846

Coxon, M., Kelly, N., & Page, S. (2016). Individual differences in virtual reality: Are spatial presence and spatial ability linked? *Virtual Reality, 20*, 203-212.

Cummings, J. J., & Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology, 19*(2), 272-309.

Curry, C., Peterson, N., Li, R., & Stoffregen, T. A. (2020). Postural precursors of motion sickness in head-mounted displays: drivers and passengers, women and men. Ergonomics, 63(12), 1502-1511.

De Dieuleveult, A. L., Siemonsma, P. C., Van Erp, J. B., & Brouwer, A. M. (2017). Effects of aging in multisensory integration: a systematic review. Frontiers in aging neuroscience, 80.

De Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends in Cognitive Sciences, 7*(10), 460-467.

De Leo, G., Diggs, L. A., Radici, E., & Mastaglio, T. W. (2014). Measuring sense of presence and user characteristics to predict effective training in an online simulated virtual environment. Simulation in Healthcare, 9(1), 1-6. DOI: 10.1097/SIH.0b013e3182a99dd9

De Niear, M. A. (2016). The Plasticity of Temporal Perception: Perceptual Training Enhances Multisensory Temporal Acuity. Vanderbilt University.

Dean, G. M., & Morris, P. E. (2003). The relationship between self‐reports of imagery and spatial ability. British Journal of Psychology, 94(2), 245-273.

Dennison, M. S., & D'Zmura, M. (2017). Cybersickness without the wobble: Experimental results speak against postural instability theory. Applied ergonomics, 58, 215-223.

Descartes, R. (2017). Meditations on first philosophy: With selections from the objections and replies (J. Cottingham, Trans). Cambridge University Press. (Original work published 1641)

Di Luzio, P., Borgomaneri, S., Sanchioni, S., Tessari, A., & Romei, V. (2021). Exposure to first-person shooter videogames is associated with multisensory temporal precision and migraine incidence. Cortex, 134, 223-238.

Diederich, A. (1992). Probability inequalities for testing separate activation models of divided attention. Perception & Psychophysics, 52(6), 714–716. https://doi.org/10.3758/BF03211708

Diederich, A. (1995). Intersensory facilitation of reaction time: Evaluation of counter and diffusion coactivation models. In Journal of Mathematical Psychology (Vol. 39, Issue 2, pp. 197–215). https://doi.org/10.1006/jmps.1995.1020

Diederich, A., & Colonius, H. (2004). Bimodal and trimodal multisensory enhancement: Effects of stimulus onset and intensity on reaction time. Perception and Psychophysics, 66(8), 1388–1404. https://doi.org/10.3758/BF03195006

Diederich, A., & Colonius, H. (2007). Modeling spatial effects in visual-tactile saccadic reaction time. Perception and Psychophysics, 69(1), 56–67. https://doi.org/10.3758/BF03194453

Diederich, A., & Colonius, H. (2008). Crossmodal interaction in saccadic reaction time: Separating multisensory from warning effects in the time window of integration model. Experimental Brain Research, 186(1), 1–22. https://doi.org/10.1007/s00221-007-1197-4

Diederich, A., & Colonius, H. (2009). Crossmodal interaction in speeded responses: time window of integration model. In Progress in Brain Research. https://doi.org/10.1016/S0079-6123(09)01311-9

Diederich, A., & Colonius, H. (2015). The time window of multisensory integration: Relating reaction times and judgments of temporal order. Psychological Review. https://doi.org/10.1037/a0038696

Dilanchian, A. T., Andringa, R., & Boot, W. R. (2021). A pilot study exploring age differences in presence, workload, and cybersickness in the experience of immersive virtual reality environments. Frontiers in Virtual Reality, 2, 736793.

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. Proceedings - Virtual Reality Annual International Symposium, 222–228. https://doi.org/10.1109/vr.1999.756955

DiZio, P., & Lackner, J. R. (2000). Motion sickness side effects and aftereffects of immersive virtual environments created with helmet-mounted visual displays. *The Capability of Virtual Reality to Meet Military Requirements, 1*, 11-12.

Donohue, S. E., Todisco, A. E., & Woldorff, M. G. (2013). The rapid distraction of attentional resources toward the source of incongruent stimulus input during multisensory conflict. Journal of Cognitive Neuroscience, 25(4), 623-635.

Donohue, S. E., Woldorff, M. G., & Mitroff, S. R. (2010). Video game players show more precise multisensory temporal processing abilities. Attention, Perception, and Psychophysics. https://doi.org/10.3758/APP.72.4.1120

Draper, J. V. (1995). Teleoperators for advanced manufacturing: applications and human factors challenges. The International Journal of Human Factors in Manufacturing, 5(1), 53–85. https://doi.org/10.1002/hfm.4530050105

Dresel, M., Plaumann, J., & Jochems, N. (2023, September). Investigating Age Differences in Passive Haptic Feedback for Immersive Virtual Reality: A Pilot Study on Configuration Tasks. In International Conference on Extended Reality (pp. 210-222). Cham: Springer Nature Switzerland.

Driver, J., & Spence, C. (2000). Multisensory perception: Beyond modularity and convergence. In Current Biology. https://doi.org/10.1016/S0960-9822(00)00740-5

Drugowitsch, J., DeAngelis, G. C., Klier, E. M., Angelaki, D. E., & Pouget, A. (2014). Optimal multisensory decision-making in a reaction-time task. *Elife*, *3*, e03005.

Dużmańska, N., Strojny, P., & Strojny, A. (2018). Can simulator sickness be avoided? A review on temporal aspects of simulator sickness. Frontiers in psychology, 9, 2132.

Eckstein, B., Krapp, E., & Lugrin, B. (2018). Towards serious games and applications in smart substitutional reality. 2018 10th International Conference on Virtual Worlds and Games for Serious Applications, VS-Games 2018 - Proceedings. https://doi.org/10.1109/VS-Games.2018.8493444

Ehrenfeld, S., & Butz, M. V. (2013). The modular modality frame model: continuous body state estimation and plausibility-weighted information fusion. Biological cybernetics, 107, 61-82.

Enders, C. K. (2017). Multiple imputation as a flexible tool for missing data handling in clinical research. *Behaviour Research and Therapy, 98*, 4-18.

Ernst, M. O. (2007). Learning to integrate arbitrary signals from vision and touch. *Journal of Vision, 7*(5), 7-7.

Ernst, M. O. (2012). Optimal multisensory integration: assumptions and limits.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. Nature, 415(January), 429–433.

Eskes, G. A., Klein, R. M., Dove, M. B., Coolican, J., & Shore, D. I. (2007). Comparing temporal order judgments and choice reaction time tasks as indices of exogenous spatial cuing. Journal of neuroscience methods, 166(2), 259-265.

Felnhofer, A., Kothgassner, O. D., Beutl, L., Hlavacs, H., & Kryspin-Exner, I. (2012). Is virtual reality made for men only? Exploring gender differences in the sense of presence. Proceedings of the International Society on presence research, 103-112.

Ferrè, E. R., Walther, L. E., & Haggard, P. (2015). Multisensory interactions between vestibular, visual and somatosensory signals. PLoS One, 10(4), e0124573.

Fetsch, C. R., Turner, A. H., DeAngelis, G. C., & Angelaki, D. E. (2009). Dynamic reweighting of visual and vestibular cues during self-motion perception. Journal of Neuroscience, 29(49), 15601-15612.

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

Fisher, J. A., Garg, A., Singh, K. P., & Wang, W. (2017, March). Designing intentional impossible spaces in virtual reality narratives: A case study. In *2017 IEEE Virtual Reality (VR)* (pp. 379-380). IEEE.

Foley, R. T., Whitwell, R. L., & Goodale, M. A. (2015). The two-visual-systems hypothesis and the perspectival features of visual experience. *Consciousness and Cognition, 35*, 225-233.

Foss-Feig, J. H., Kwakye, L. D., Cascio, C. J., Burnette, C. P., Kadivar, H., Stone, W. L., & Wallace, M. T. (2010). An extended multisensory temporal binding window in autism spectrum disorders. *Experimental brain research*, *203*(2), 381–389. https://doi.org/10.1007/s00221-010-2240-4

Foxe, J. J. (2009). Multisensory Integration: Frequency Tuning of Audio-Tactile Integration. Current Biology, 19(9), R373–R375. https://doi.org/10.1016/j.cub.2009.03.029

Freeman, J., Avons, S. E., Meddis, R., Pearson, D. E., & Ijsselsteijn, W. (2000). Using behavioral realism to estimate presence: A study of the utility of postural responses to motion stimuli. Presence: Teleoperators and Virtual Environments, 9(2), 149–164. https://doi.org/10.1162/105474600566691

Freitag, S., Weyers, B., & Kuhlen, T. W. (2016). Examining rotation gain in CAVE-like virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, *22*(4), 1462-1471.

Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004). Recalibration of audiovisual simultaneity. Nature Neuroscience, 7(7), 773–778. https://doi.org/10.1038/nn1268

Fulvio, J. M., Ji, M., & Rokers, B. (2021). Variations in visual sensitivity predict motion sickness in virtual reality. Entertainment Computing, 38, 100423.

Gibbon, J., & Rutschmann, R. (1969). Temporal order judgment and reaction time. Science, 4, 413–415.

Gibbs, J. K., Gillies, M., & Pan, X. (2022). A comparison of the effects of haptic and visual feedback on presence in virtual reality. *International Journal of Human-Computer Studies, 157*, 102717.

Gielen, S. C. A. M., Schmidt, R. A., & Van Den Heuvel, P. J. M. (1983). On the nature of intersensory facilitation of reaction time. 34(2), 161–168.

Gim, B., Kang, S., Kim, G., Yeo, D., Hwang, S., & Kim, S. (2024, May). Curving the Virtual Route: Applying Redirected Steering Gains for Active Locomotion in In-Car VR. In Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (pp. 1-7).

Glazebrook, C. M., Welsh, T. N., & Tremblay, L. (2016). The processing of visual and auditory information for reaching movements. Psychological Research, 80(5), 757–773. https://doi.org/10.1007/s00426-015-0689-2

Glenberg, A. M. (1997). Mental models, space, and embodied cognition. In Creative thought: An investigation of conceptual structures and processes. https://doi.org/10.1037/10227-018

Glenn, C. P., & Coxon, M. (2024). Individual differences in processing multisensory information predict presence in different virtual reality environments. *Virtual Reality*, *29*(1), 9.

Gondan, M., & Minakata, K. (2016). A tutorial on testing the race model inequality. Attention, Perception, and Psychophysics, 78(3), 723–735. https://doi.org/10.3758/s13414-015-1018-y

Gramann, K. (2013). Embodiment of spatial reference frames and individual differences in reference frame proclivity. Spatial Cognition & Computation, 13(1), 1-25.

Gramann, K., Onton, J., Riccobon, D., Mueller, H. J., Bardins, S., & Makeig, S. (2010). Human brain dynamics accompanying use of egocentric and allocentric reference frames during navigation. *Journal of Cognitive Neuroscience*, *22*(12), 2836-2849.

Grant, K. W., & Seitz, P.-F. (2002). The use of visible speech cues for improving auditory detection of spoken sentences. The Journal of the Acoustical Society of America, 108(3), 1197. https://doi.org/10.1121/1.1288668

Grassini, S., & Laumann, K. (2020). Questionnaire measures and physiological correlates of presence: A systematic review. *Frontiers in Psychology*, *11*, 349.

Grassini, S., Laumann, K., & Luzi, A. K. (2021). Association of individual factors with simulator sickness and sense of presence in virtual reality mediated by head-mounted displays (HMDs). *Multimodal Technologies and Interaction, 5*(3), 7.

Greenwald, A. G. (1976). Within-subjects designs: To use or not to use? *Psychological Bulletin, 83*(2), 314.

Grigoriadou, M., Loukakis, G., & Lekkas, T. (Under Review) The spatial embodied perception of non-Euclidean geometries in video games.

Guldin, W. O., & Grüsser, O. J. (1998). Is there a vestibular cortex?. Trends in neurosciences, 21(6), 254-259.

Gyselinck, V., Picucci, L., Nicolas, S., & Piolino, P. (2006). Construction of a spatial mental model from a verbal description or from navigation in a virtual environment. Cognitive Processing, 7, 46-48.

Hakkinen, J., Vuori, T., & Paakka, M. (2002, October). Postural stability and sickness symptoms after HMD use. In IEEE international conference on systems, man and cybernetics (Vol. 1, pp. 147-152). Yasmine Hammamet: IEEE.

Hartmann, T., Wirth, W., Schramm, H., Klimmt, C., Vorderer, P., Gysbers, A., Böcking, S., Ravaja, N., Laarni, J., Saari, T., Gouveia, F., & Sacau, A. M. (2016). The spatial presence experience scale (SPES): A short self-report measure for diverse media settings. Journal of Media Psychology, 28(1), 1–15. https://doi.org/10.1027/1864-1105/a000137

Hartmann, T., Wirth, W., Vorderer, P., Klimmt, C., Schramm, H., & Böcking, S. (2015). Spatial presence theory: State of the art and challenges ahead. Immersed in media: Telepresence theory, measurement & technology, 115-135.

Hassenstein, V. B., & Reichardt, W. (1956). System theoretical analysis of time, sequence and sign analysis of the motion perception of the snout-beetle. Zeitschrift Für Naturforschung, 11, 513–525.

Hecht, D., & Reiner, M. (2009). Sensory dominance in combinations of audio, visual and haptic stimuli. Experimental brain research, 193, 307-314.

Hecht, D., Reiner, M., & Halevy, G. (2005). Multi-Modal Stimulation, Response Time, and Presence Technion – Israel Institute of Technology The University of Haifa. May, 269–274.

Heeter, C. (1992). Being There: The Subjective Experience of Presence. Presence: Teleoperators and Virtual Environments, 1(2), 262–271. https://doi.org/10.1162/pres.1992.1.2.262

Hein, D., Mai, C., & Hußmann, H. (2018, May). The usage of presence measurements in research: a review. In Proceedings of the International Society for Presence Research Annual Conference (Presence) (pp. 21-22). Prague: The International Society for Presence Research.

Held, R. M., & Durlach, N. I. (1992). Telepresence. Presence: Teleoperators and Virtual Environments. https://doi.org/10.1162/pres.1992.1.1.109

Hershenson, M. (1962). Reaction time as a measure of retroactive inhibition. Journal of Experimental Psychology, 63(3), 289–293. https://doi.org/10.1037/h00395163

Hidi, S. E. (1995). A reexamination of the role of attention in learning from text. *Educational Psychology Review, 7*, 323-350.

Hillis, J. M., Ernst, M. O., Banks, M. S., & Landy, M. S. (2002). Combining sensory information: mandatory fusion within, but not between, senses. Science, 298(5598), 1627-1630.

Hillock, A. R., Powers, A. R., & Wallace, M. T. (2011). Binding of sights and sounds: Age-related changes in multisensory temporal processing. *Neuropsychologia*. 49(3), 461–467. https://doi.org/10.1016/j.neuropsychologia.2010.11.041

Hillock-Dunn, A., & Wallace, M. T. (2012). Developmental changes in the multisensory temporal binding window persist into adolescence. Developmental Science, 15(5), 688–696. https://doi.org/10.1111/j.1467-7687.2012.01171.x

Holmes, N. P. (2007). The law of inverse effectiveness in neurons and behaviour: Multisensory integration versus normal variability. Neuropsychologia, 45(14), 3340–3345. https://doi.org/10.1016/j.neuropsychologia.2007.05.025

Hopstaken, J. F., Van Der Linden, D., Bakker, A. B., & Kompier, M. A. (2015). A multifaceted investigation of the link between mental fatigue and task disengagement. Psychophysiology, 52(3), 305-315.

Howard, M. C., & Van Zandt, E. C. (2021). A meta-analysis of the virtual reality problem: Unequal effects of virtual reality sickness across individual differences. Virtual Reality, 25(4), 1221-1246.

Huang, M., & Alessi, N. (1999). Presence as an Emotional Experience. Medicine Meets Virtual Reality: The Convergence of Physical and Informational Technologies Options for a New Era in Healthcare, eds, 148–153. http://widgets.ebscohost.com/prod/customerspecific/ns000290/authentication/index.php?url=http%3A%2F%2Fsearch.ebscohost.com%2Flogin.aspx%3Fdirect%3Dtrue%26AuthType%3Dip%2Ccookie%2Cshib%2Cuid%26db%3Deric%26AN%3DEJ1072174%26lang%3Dpt-br%26site%3Deds-live%26s

Hutmacher, F. (2019). Why is there so much more research on vision than on any other sensory modality?. *Frontiers in psychology*, *10*, 481030.

Hwang, I. J., Kim, S. C., Chang, E. H., Jeon, H. J., Chun, Y. S., Park, C. H., & Kim, H. T. (2012). Effect of an independent visual foreground on cybersickness in virtual environment. *Korean Journal of Cognitive and Biological Psychology, 24*(3), 251-263.

Hyun, J., Habuchi, Y., Park, A., Ishikawa, T., Kourogi, M., & Kurata, T. (2010). Service-Field Simulator using MR Techniques: Behavior Comparison in Real and Virtual Environments. Proceedings on ICAT2010, 14–21.

Iachini, T., Maffei, L., Masullo, M., Senese, V. P., Rapuano, M., Pascale, A., Sorrentino, F., & Ruggiero, G. (2019). The experience of virtual reality: are individual differences in mental imagery associated with sense of presence? Cognitive Processing, 20(3), 291–298. https://doi.org/10.1007/s10339-018-0897-y

IJsselsteijn, W., De Ridderb, H., Freeman, J., & Avons, S. (1990). Presence: Concept, determinants and measurement. Journal of Banking and Finance, 14(2–3), 649–672. https://doi.org/10.1016/0378-4266(90)90068-D

Insko, B. E. (2003). Measuring presence: Subjective, behavioral, and physiological methods. In G. Riva, F. Davide, & W. A. IJsselsteijn (Eds.), *Being there: Concepts, effects and measurements of user presence in synthetic environments* (pp. 109-119). IOS Press.

Islam, R., Lee, Y., Jaloli, M., Muhammad, I., Zhu, D., & Quarles, J. (2020, March). Automatic detection of cybersickness from physiological signal in a virtual roller coaster simulation. In 2020 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW) (pp. 648-649). IEEE.

Iwaniec, J. (2019). Questionnaires: Implications for effective implementation. In The Routledge handbook of research methods in applied linguistics (pp. 324-335). Routledge.

Jaeger, B. K., & Mourant, R. R. (2001, October). Comparison of simulator sickness using static and dynamic walking simulators. In Proceedings of the human factors and ergonomics society annual meeting (Vol. 45, No. 27, pp. 1896-1900). Sage CA: Los Angeles, CA: SAGE Publications.

Johnson, D. M. (2005). Introduction to and review of simulator sickness research.

Jung, J. Y., Adelstein, B. D., & Ellis, S. R. (2000, July). Discriminability of prediction artifacts in a time-delayed virtual environment. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 44, No. 5, pp. 499-502). Sage CA: SAGE Publications.

Jurnet, I. A., Beciu, C. C., & Maldonado, J. G. (2005). Individual differences in the sense of presence. In *Proceedings of Presence 2005: The 8th International Workshop on Presence* (pp. 133-142). London: University College London.

Kandil, F. I., Diederich, A., & Colonius, H. (2014). Parameter recovery for the time-window-of-integration (twin) model of multisensory integration in focused attention. Journal of Vision, 14(11), 14-14.

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: Teleoperators & Virtual Environments*, *9*(5), 463-472.

Keshavarz, B., Philipp-Muller, A. E., Hemmerich, W., Riecke, B. E., & Campos, J. L. (2019). The effect of visual motion stimulus characteristics on vection and visually induced motion sickness. Displays, 58, 71-81.

Keshavarz, B., Ramkhalawansingh, R., Haycock, B., Shahab, S., & Campos, J. L. (2018). Comparing simulator sickness in younger and older adults during simulated driving under different multisensory conditions. Transportation research part F: traffic psychology and behaviour, 54, 47-62.

Kim, H., & Lee, I. K. (2022). Studying the effects of congruence of auditory and visual stimuli on virtual reality experiences. IEEE Transactions on Visualization and Computer Graphics, 28(5), 2080-2090.

Kim, H., Jeon, S. B., & Lee, I. K. (2024). Locomotion Techniques for Dynamic Environments: Effects on Spatial Knowledge and User Experiences. IEEE Transactions on Visualization and Computer Graphics.

Kim, H., Seo, Y. J., Yang, B., & Chu, H. Y. (2017). Black perception in a transparent OLED display. *Optics Express*, *25*(4), 3954-3965.

Kim, J., Luu, W., & Palmisano, S. (2020). Multisensory integration and the experience of scene instability, presence, and cybersickness in virtual environments. *Computers in Human Behavior, 113*, 106484. https://doi.org/10.1016/j.chb.2020.106484

Kim, J., Palmisano, S., Luu, W., & Iwasaki, S. (2021). Effects of linear visual-vestibular conflict on presence, perceived scene stability and cybersickness in the Oculus Go and Oculus Quest. *Frontiers in Virtual Reality, 2*, 582156.

Klatzky, R. L. (1998). Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. In *Spatial cognition: An interdisciplinary approach to*

representing and processing spatial knowledge (pp. 1–17). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-80378-8_1

Koelewijn, T., Bronkhorst, A., & Theeuwes, J. (2010). Attention and the multiple stages of multisensory integration: A review of audiovisual studies. Acta Psychologica, 134(3), 372–384. https://doi.org/10.1016/j.actpsy.2010.03.010

Kohl, R. L. (1983). Sensory conflict theory of space motion sickness: an anatomical location for the neuroconflict. Aviation, space, and environmental medicine, 54(5), 464-465.

Kostaki, M., & Vatakis, A. (2018). Temporal order and synchrony judgments: a primer for students. In Timing and time perception: Procedures, measures, & applications (pp. 233-262). Brill.

Kourtesis, P., Collina, S., Doumas, L. A., & MacPherson, S. E. (2019). Technological competence is a precondition for effective implementation of virtual reality head-mounted displays in human neuroscience: A technological review and meta-analysis. *Frontiers in Human Neuroscience, 13*, 342. https://doi.org/10.3389/fnhum.2019.00342

Krijn, M., Emmelkamp, P. M. G., Biemond, R., De Wilde De Ligny, C., Schuemie, M. J., & Van Der Mast, C. A. P. G. (2004). Treatment of acrophobia in virtual reality: The role of immersion and presence. Behaviour Research and Therapy, 42(2), 229–239. https://doi.org/10.1016/S0005-7967(03)00139-6

Laarni, J., Ravaja, N., Saari, T., Böcking, S., Hartmann, T., & Schramm, H. (2015). Ways to measure spatial presence: Review and future directions. In Immersed in Media: Telepresence Theory, Measurement and Technology. https://doi.org/10.1007/978-3-319-10190-3_8

Landy, M. S., Banks, M. S., & Knill, D. C. (2012). Ideal-Observer Models of Cue Integration. Sensory Cue Integration, 5–29. https://doi.org/10.1093/acprof:oso/9780195387247.003.0001

Lanier, J. (2017). Dawn of the new everything: A journey through virtual reality. Random House.

Larsson, J., Landy, M. S., & Heeger, D. J. (2006). Orientation-selective adaptation to first- and second-order patterns in human visual cortex. *Journal of Neurophysiology*, *95*(2), 862–881. https://doi.org/10.1152/jn.00641.2005

Lawson, B. (2014). Motion sickness symptomatology and origins. In Hale, K. S., & Stanney, K. M. (Eds.). Handbook of virtual environments: Design, implementation, and applications. CRC Press. (2nd ed., pp. 531–600)

Lawson, B. D., Proietti, P., Peder, S., Timothy, R., Ramy, K., & Marten, B. (2021). Factors impacting cybersickness.

Lawson, B. D., Proietti, P., Peder, S., Timothy, R., Ramy, K., & Marten, B. (2021). Factors impacting cybersickness. *Frontiers in Virtual Reality, 2*, 627250. https://doi.org/10.3389/frvir.2021.627250

Lemmens, J. S., Valkenburg, P. M., & Peter, J. (2011a). Psychosocial causes and consequences of pathological gaming. Computers in Human Behavior. https://doi.org/10.1016/j.chb.2010.07.015

Lemmens, J. S., Valkenburg, P. M., & Peter, J. (2011b). The Effects of Pathological Gaming on Aggressive Behavior. Journal of Youth and Adolescence. https://doi.org/10.1007/s10964-010-9558-x

Leonard, A. S., & Masek, P. (2014). Multisensory integration of colors and scents: Insights from bees and flowers. Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology, 200(6), 463–474. https://doi.org/10.1007/s00359-014-0904-4

Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-sense of presence inventory. Presence: Teleoperators and Virtual Environments, 10(3), 282–297. https://doi.org/10.1162/105474601300343612

Levitin, D. J., Maclean, K., Mathews, M., & Chu, L. (1999). The Perception of Cross-Modal Simultaneity. International Journal of Computing Anticipatory Systems, 6–9.

Lin, J. W., Duh, H. B. L., Parker, D. E., Abi-Rached, H., & Furness, T. A. (2002, March). Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment. In Proceedings ieee virtual reality 2002 (pp. 164-171). IEEE.

Ling, Y., Nefs, H. T., Brinkman, W. P., Qu, C., & Heynderickx, I. (2013). The relationship between individual characteristics and experienced presence. Computers in Human Behavior, 29(4), 1519-1530.

Litleskare, S. (2021). The relationship between postural stability and cybersickness: it's complicated–an experimental trial assessing practical implications of cybersickness etiology. Physiology & Behavior, 236, 113422.

Little, R. J., & Rubin, D. B. (2019). *Statistical analysis with missing data* (Vol. 793). John Wiley & Sons. https://doi.org/10.1002/9781119482260

Liu, C. L., & Uang, S. T. (2012, May). A study of sickness induced within a 3D virtual store and combated with fuzzy control in the elderly. In 2012 9th International Conference on Fuzzy Systems and Knowledge Discovery (pp. 334-338). IEEE.

Lochner, D. C., & Gain, J. E. (2021, November). VR natural walking in impossible spaces. In Proceedings of the 14th ACM SIGGRAPH Conference on Motion, Interaction and Games (pp. 1-9).

Lohmann, J., & Butz, M. V. (2017). Lost in space: multisensory conflict yields adaptation in spatial representations across frames of reference. Cognitive Processing, 18, 211-228.

Lombard, M., & Ditton, T. (1997). At the Heart of It All: The Concept of Presence. Journal of Computer-Mediated Communication, 3(2). https://doi.org/10.1111/j.1083-6101.1997.tb00072.x

Lombard, M., Ditton, T. B., Crane, D., Davis, B., Gil-Egui, G., Horvath, K., & Rossman, J. (2000). Measuring presence: A literature-based approach to the development of a standardized paper-and-pencil instrument. Presence 2000: The Third International Workshop on Presence, January 13. https://doi.org/10.1007/978-3-319-14892-2

Lopez, C., Blanke, O., & Mast, F. W. (2012). The human vestibular cortex revealed by coordinate-based activation likelihood estimation meta-analysis. Neuroscience, 212, 159-179.

Lorenz, M., Brade, J., Klimant, P., Heyde, C. E., & Hammer, N. (2023). Age and gender effects on presence, user experience and usability in virtual environments–first insights. PloS one, 18(3), e0283565.

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. https://doi.org/10.1016/j.displa.2015.02.002

Macaluso, E. (2010). Orienting of spatial attention and the interplay between the senses. Cortex, 46(3), 282-297.

MacDonald, J. (2018). Hearing lips and seeing voices: the origins and development of the ‘McGurk Effect’and reflections on audio–visual speech perception over the last 40 years. Multisensory Research, 31(1-2), 7-18.

Man, D., & Olchawa, R. (2018). The possibilities of using BCI technology in biomedical engineering. In Biomedical Engineering and Neuroscience: Proceedings of the 3rd International Scientific Conference on Brain-Computer Interfaces, BCI 2018, March 13-14, Opole, Poland (pp. 30-37). Springer International Publishing. Doi:10.1007/978-3-319-75025-5_4.

Mania, K., & Chalmers, A. (2001). The effects of levels of immersion on memory and presence in virtual environments: A reality centered approach. Cyberpsychology & behavior, 4(2), 247-264.

Marill, T. (1956). Detection Theory and Psychophysics Thomas Marill. Technology.

Marucci, M., Di Flumeri, G., Borghini, G., Sciaraffa, N., Scandola, M., Pavone, E. F., ... & Aricò, P. (2021). The impact of multisensory integration and perceptual load in virtual reality settings on performance, workload, and presence. *Scientific Reports*, *11*(1), 4831. https://doi.org/10.1038/s41598-021-84231-0

Marzvanyan, A., & Alhawaj, A. F. (2019). Physiology, sensory receptors. [Updated 2023 Aug 14]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2024 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK539861/

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*(5588), 746–748. https://doi.org/10.1038/264746a0

Meehan, M., Insko, B., Whitton, M., & Brooks, F. P. (2002). Physiological measures of presence. ACM Transactions on Graphics, 21(3), 645–652. https://doi.org/10.1145/566654.566630

Mégevand, P., Molholm, S., Nayak, A., & Foxe, J. J. (2013). Recalibration of the Multisensory Temporal Window of Integration Results from Changing Task Demands. PLoS ONE, 8(8). https://doi.org/10.1371/journal.pone.0071608

Meredith, M. A., & Stein, B. E. (1983). Interactions Among Converging Sensory Inputs in the Superior Colliculus. Science, 369(July), 389–391.

Meredith, M. A., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. The Journal of Neuroscience, 7(10), 3215–3229. https://doi.org/citeulike-article-id:409430

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. Cognitive Psychology, 14(2), 247–279. https://doi.org/10.1016/0010-0285(82)90010-X

Miller, J. (1986). Timecourse of coactivation in bimodal divided attention. Perception & Psychophysics, 40(5), 331–343. https://doi.org/10.3758/BF03203025

Miller, J. (2016). Statistical facilitation and the redundant signals effect: What are race and coactivation models? Attention, Perception, and Psychophysics, 78(2), 516–519. https://doi.org/10.3758/s13414-015-1017-z

Miller, J., & Schwarz, W. (2006). Dissociations between reaction times and temporal order judgments: A diffusion model approach. Journal of Experimental Psychology: Human Perception and Performance, 32(2), 394–412. https://doi.org/10.1037/0096-1523.32.2.394

Mindflux (2000), VFX3D, http://www.mindflux.com.au/products/iis/vfx3d.html

Minsky, M. (1980). Telepresence. Omni, June, 45–51. https://doi.org/10.1351/pac198052010233

Moss, J. D., Austin, J., Salley, J., Coats, J., Williams, K., & Muth, E. R. (2011). The effects of display delay on simulator sickness. Displays, 32(4), 159-168.

Muryy, A., & Glennerster, A. (2021). Route selection in non-Euclidean virtual environments. *PloS one, 16*(4), e0247818.

Nardini, M., Jones, P., Bedford, R., & Braddick, O. (2008). Development of cue integration in human navigation. Current biology, 18(9), 689-693. doi: 10.1016/j.cub.2008.04.021.

Navarra, J., Soto-Faraco, S., & Spence, C. (2007). Adaptation to audiotactile asynchrony. Neuroscience Letters, 413(1), 72–76. https://doi.org/10.1016/j.neulet.2006.11.027

Navarra, J., Soto-Faraco, S., & Spence, C. (2007). Adaptation to audiotactile asynchrony. *Neuroscience Letters, 413*(1), 72–76. https://doi.org/10.1016/j.neulet.2006.11.008

Navarra, J., Vatakis, A., Zampini, M., Soto-Faraco, S., Humphreys, W., & Spence, C. (2005). Exposure to asynchronous audiovisual speech extends the temporal window for audiovisual integration. Cognitive Brain Research, 25(2), 499–507. https://doi.org/10.1016/j.cogbrainres.2005.07.009

Nelson, W. T., Roe, M. M., Bolia, R. S., & Morley, R. M. (2000). Assessing simulator sickness in a see-through HMD: Effects of time delay, time on task, and task complexity. Image.

Ng, A. K., Chan, L. K., & Lau, H. Y. (2020). A study of cybersickness and sensory conflict theory using a motion-coupled virtual reality system. *Displays*, *61*, 101922. https://doi.org/10.1016/j.displa.2019.101922

Nguyen-Vo, T., Riecke, B. E., & Stuerzlinger, W. (2018, March). Simulated reference frame: A cost-effective solution to improve spatial orientation in vr. In 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR) (pp. 415-422). IEEE.)

Nidiffer, A. R., Diederich, A., Ramachandran, R., & Wallace, M. T. (2018). Multisensory perception reflects individual differences in processing temporal correlations. Scientific Reports, 8(1), 1–15. https://doi.org/10.1038/s41598-018-32673-y

Nilsson, N. C., Peck, T., Bruder, G., Hodgson, E., Serafin, S., Whitton, M., ... & Rosenberg, E. S. (2018). 15 years of research on redirected walking in immersive virtual environments. *IEEE computer graphics and applications*, *38*(2), 44-56.

Noel, J., Wallace, M. T., & Alais, D. (2015). True and Perceived Synchrony are Preferentially Associated With Particular Sensory Pairings. *Scientific Reports*, *5*(1), 1-8. https://doi.org/10.1038/srep17467

Notter, M. P., & Murray, M. M. (2017). Temporal Binding Window scripts: A lightweight MATLAB tool to analyze the temporal binding window in a multisensory integration study. *Zenodo*. https://doi.org/10.5281/zenodo.815876

Nunez, D. (2007). A Capacity Limited, Cognitive Constructionist Model of Virtual Presence. Dissertation, University of Cape Town. http://pubs.cs.uct.ac.za/archive/00000454/

Nunez, D., & Blake, E. (2003, October). The thematic baseline technique as a means of improving the sensitivity of presence self-report scales. In 6th International Workshop on Presence (pp. 6-8).

Nystad, E., & Sebok, A. (2004). A Comparison of Two Presence Measures Based on Experimental Results. Seventh Annual International Workshop Presence 2004, 266–273.

Nystad, E., & Sebok, A. (2004, October). A comparison of two presence measures based on experimental results. In *Proceedings of the Seventh International Workshop on Presence* (pp. 266–273).

Oruç, I., Maloney, L. T., & Landy, M. S. (2003). Weighted linear cue combination with possibly correlated error. Vision Research, 43(23), 2451–2468. https://doi.org/10.1016/S0042-6989(03)00435-8

Panahi-Shahri, M., Fathi Ashtiani, A., Azad-Fallah, P., & Montazer, G. (2009). Reliability and validity of Igroup Presence Questionnaire (IPQ). *Archive of SID*. Spring, 3(1), 27–34.

Parise, C. V., & Ernst, M. O. (2016). Correlation detection as a general mechanism for multisensory integration. Nature Communications, 7, 1–9. https://doi.org/10.1038/ncomms11543

Park, G. D., Allen, R. W., Fiorentino, D., Rosenthal, T. J., & Cook, M. L. (2006, October). Simulator sickness scores according to symptom susceptibility, age, and gender for an older driver assessment study. In Proceedings of the human factors and ergonomics society annual meeting (Vol. 50, No. 26, pp. 2702-2706). Sage CA: Los Angeles, CA: SAGE Publications.

Parker, D., & Prothero, J. (2003). A Unified Approach to Presence and Motion Sickness. In Virtual and Adaptive Environments. https://doi.org/10.1201/9781410608888.ch3

Pasalar, S., Ro, T., & Beauchamp, M. S. (2010). TMS of posterior parietal cortex disrupts visual tactile multisensory integration. European Journal of Neuroscience, 31(10), 1783–1790. https://doi.org/10.1111/j.1460-9568.2010.07193.x

Payne, J., & Robb, A. (2024). Exploring the effects of self-overlapping spaces on distance perception and action judgments. *ACM Transactions on Applied Perception*, *21*(4), 1-14.

Peck, T. C., Fuchs, H., & Whitton, M. C. (2011, March). An evaluation of navigational ability comparing redirected free exploration with distractors to walking-in-place and joystick locomotio interfaces. In 2011 IEEE Virtual Reality Conference (pp. 55-62). IEEE. https://doi.org/10.1109/VR.2011.5759437

Peer, E., Rothschild, D., Gordon, A., & Damer, E. (2022). Erratum to Peer et al. (2021) Data quality of platforms and panels for online behavioral research. *Behavior Research Methods, 54*(5), 2618–2620. https://doi.org/10.3758/s13428-022-01848-5

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1–2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*. https://doi.org/10.3758/s13428-018-01193-y

Pöhlmann, K. M. T., O'Hare, L., Dickinson, P., Parke, A., & Föcker, J. (2022). Action video game players do not differ in the perception of contrast-based motion illusions but experience more vection and less discomfort in a virtual environment compared to non-action video game players. Journal of Cognitive Enhancement, 1-17.

Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends in Cognitive Sciences, 1*(2), 56–61. https://doi.org/10.1016/S1364-6613(97)01008-5

Pöppel, E. (2009). Pre-semantically defined temporal windows for cognitive processing. Philosophical Transactions of the Royal Society: B, Biological Sciences, 364, 1887-1896.

Pöppel, E. (2009). Pre-semantically defined temporal windows for cognitive processing. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1525), 1887–1896. https://doi.org/10.1098/rstb.2009.0015

Pöppel, E., & Artin, T. T. (1988). *Mindworks: Time and conscious experience.* Harcourt Brace Jovanovich.

Pouget, A., Deneve, S., & Duhamel, J. R. (2002). A computational perspective on the neural basis of multisensory spatial representations. *Nature Reviews Neuroscience, 3*(9), 741–747. https://doi.org/10.1038/nrn914

Powers III, A. R., Hillock-Dunn, A., & Wallace, M. T. (2016). Generalization of multisensory perceptual learning. *Scientific Reports, 6,* 23374. https://doi.org/10.1038/srep23374

Powers, A. R., Hillock, A. R., & Wallace, M. T. (2009). Perceptual Training Narrows the Temporal Window of Multisensory Binding. Journal of Neuroscience, 29(39), 12265–12274. https://doi.org/10.1523/JNEUROSCI.3501-09.2009

Prothero, J. D., Draper, M. H., Furness 3rd, T. A., Parker, D. E., & Wells, M. J. (1999). The use of an independent visual background to reduce simulator side-effects. Aviation, space, and environmental medicine, 70(3 Pt 1), 277-283.

Prsa, M., Gale, S., & Blanke, O. (2012). Self-motion leads to mandatory cue fusion across sensory modalities. Journal of neurophysiology, 108(8), 2282-2291.

Raab, D. (1962). STATISTICAL FACILITATION OF SIMPLE REACTION TIMES. Transactions of the New York Academy of Sciences, 24(5 Series 2), 574–590.

Raaen, K., & Kjellmo, I. (2015). Measuring latency in virtual reality systems. In Entertainment Computing-ICEC 2015: 14th International Conference, ICEC 2015, Trondheim, Norway, September 29-Ocotober 2, 2015, Proceedings 14 (pp. 457-462). Springer International Publishing.

Rajagopalan, A., Kumar, S. S., & Mukkadan, J. K. (2017). Effect of vestibular stimulation on auditory and visual reaction time in relation to stress. Journal of Advanced Pharmaceutical Technology & Research, 8(1), 34-38.

Rangelova, S., Rehm, K., Diefenbach, S., Motus, D., & André, E. (2020). Gender differences in simulation sickness in static vs. moving platform VR automated driving simulation. In HCI in Mobility, Transport, and Automotive Systems. Automated Driving and In-Vehicle Experience Design: Second International Conference, MobiTAS 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part I 22 (pp. 146-165). Springer International Publishing.

Ravaja, N., Salminen, M., Holopainen, J., Saari, T., Laarni, J., & J�rvinen, A. (2004). Emotional response patterns and sense of presence during video games. Proceedings of the Third Nordic Conference on Human-Computer Interaction - NordiCHI ’04, 339–347. https://doi.org/10.1145/1028014.1028068

Razzaque, S., Swapp, D., Slater, M., Whitton, M. C., & Steed, A. (2002, May). Redirected walking in place. In *Egve* (Vol. 2, pp. 123-130).

Reason, J. T. (1975). *Motion sickness*.

Rebenitsch, L., & Owen, C. (2016). Review on cybersickness in applications and visual displays. *Virtual Reality, 20*(2), 101-125. https://doi.org/10.1007/s10055-016-0285-9

Regenbrecht, H. (2002). Measuring presence in augmented reality environments: design and a first test of a questionnaire. Annual International Workshop Presence, 1–4. https://www.google.com/%5Cnpapers3://publication/uuid/BE89B550-B84B-4411-9464-71837C29609F

Reichenbach, A., Bresciani, J. P., Bülthoff, H. H., & Thielscher, A. (2016). Reaching with the sixth sense: Vestibular contributions to voluntary motor control in the human right parietal cortex. *NeuroImage, 124*, 869-875. https://doi.org/10.1016/j.neuroimage.2015.09.046

Rempel, D., Willms, K., Anshel, J., Jaschinski, W., & Sheedy, J. (2007). The effects of visual display distance on eye accommodation, head posture, and vision and neck symptoms. Human factors, 49(5), 830-838.

Riccio, G. E., & Stoffregen, T. A. (1991). An ecological theory of motion sickness and postural instability. *Ecological Psychology, 3*(3), 195-240. https://doi.org/10.1207/s15326969eco0303_2

Riecke, B. E., & Schulte-Pelkum, J. (2013). Perceptual and cognitive factors for self-motion simulation in virtual environments: How can self-motion illusions ("vection") be utilized? In F. Steinicke, Y. Visell, J. Campos, & A. Lécuyer (Eds.), *Human walking in virtual environments: Perception, technology, and applications* (pp. 27-54). Springer. https://doi.org/10.1007/978-1-4419-8432-6_2

Riecke, B., & Von der Heyde, M. (2002). Qualitative modeling of spatial orientation processes using logical propositions: Interconnecting spatial presence, spatial updating, piloting, and spatial cognition. Max Planck Institute for Biological Cybernetics, 100. http://scholar.google.com.proxy.lib.sfu.ca/scholar?as_q=&num=100&btnG=Search+Scholar

&as_epq=&as_oq=&as_eq=&as_occt=any&as_sauthors=%22BE+Riecke%22&as_publication=&as_ylo=&as_yhi=&as_sdt=1.&as_sdtp=on&as_sdts=5&hl=en#5

Riva, G., Davide, F., & IJsselsteijn, W. A. (2003). Measuring presence: Subjective, behavioral and physiological methods. Being there: Concepts, effects and measurement of user presence in synthetic environments, 5, 110-118.

Riva, G., Mantovani, F., Capideville, C. S., Preziosa, A., Morganti, F., Villani, D., ... & Alcañiz, M. (2007). Affective interactions using virtual reality: the link between presence and emotions. Cyberpsychology & behavior, 10(1), 45-56.

Robb, A., & Barwulor, C. (2019, September). Perception of spatial relationships in impossible spaces. In ACM Symposium on Applied Perception 2019 (pp. 1-5).

Robb, A., & Barwulor, C. (2021, September). Spatial judgments in impossible spaces preserve important relative information. In *ACM Symposium on Applied Perception 2021* (pp. 1-9). https://doi.org/10.1145/3474451.3476211

Rosas, P., Wichmann, F. A., & Wagemans, J. (2007). Texture and object motion in slant discrimination: Failure of reliability-based weighting of cues may be evidence for strong fusion. Journal of Vision, 7(6), 3. https://doi.org/10.1167/7.6.3

Ryan, V., & Griffin, R. (2016). An investigation into anxiety in virtual reality following a self-compassion induction. *Annual Review of CyberTherapy and Telemedicine*.

Sadiq, O. (2019). *Psychometric correlates of multisensory integration as potential predictors of cybersickness in virtual reality* (Master's thesis, University of Waterloo). https://uwspace.uwaterloo.ca/handle/10012/15234

Sadiq, O., & Barnett-Cowan, M. (2022). Can the perceived timing of multisensory events predict cybersickness? *Multisensory Research, 35*(7-8), 623-652. https://doi.org/10.1163/22134808-bja10066

Santangelo, V., & Spence, C. (2007). Multisensory Cues Capture Spatial Attention Regardless of Perceptual Load. Journal of Experimental Psychology: Human Perception and Performance, 33(6), 1311–1321. https://doi.org/10.1037/0096-1523.33.6.1311

Saredakis, D., Szpak, A., Birckhead, B., Keage, H. A., Rizzo, A., & Loetscher, T. (2020). Factors associated with virtual reality sickness in head-mounted displays: a systematic review and meta-analysis. Frontiers in human neuroscience, 14, 96.

Schubert, T., & Friedmann, F. (2001). The Experience of Presence: Factor Analytic Insights. 10(3), 266–281. https://doi.org/10.1162/105474601300343603

Schubert, T., Friedmann, F., & Regenbrecht, H. (1999a). Decomposing the sense of presence: Factor analytic insights. Workshop on Presence, 1–5. http://www.temple.edu/ispr/prev_conferences/proceedings/98-99-2000/1999/Schubert.pdf

Schubert, T., Friedmann, F., & Regenbrecht, H. (1999b). Embodied Presence in Virtual Environments. Visual Representations and Interpretations, 269–278. https://doi.org/10.1007/978-1-4471-0563-3_30

Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. Presence: Teleoperators and Virtual Environments. https://doi.org/10.1162/105474601300343603

Schuemie, M. J., van der Straaten, P., Krijn, M., & van der Mast, C. A. P. G. (2001). Research on Presence in Virtual Reality: A Survey. CyberPsychology & Behavior, 4(2), 183–201. https://doi.org/10.1089/109493101300117884

Schwarz, W. (1989). A new model to explain the redundant-signals effect. Perception & Psychophysics, 46(5), 498–500. https://doi.org/10.3758/BF03210867

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

Shayman, C. S., Seo, J. H., Oh, Y., Lewis, R. F., Peterka, R. J., & Hullar, T. E. (2018). Relationship between vestibular sensitivity and multisensory temporal integration. Journal of neurophysiology, 120(4), 1572-1577.

Sheridan, T. B. (1992). Musings on Telepresence and Virtual Presence. Presence: Teleoperators and Virtual Environments. https://doi.org/10.1162/pres.1992.1.1.120

Simon, J., Takács, E., Orosz, G., Berki, B., & Winkler, I. (2020). Short-term cognitive fatigue effect on auditory temporal order judgments. Experimental Brain Research, 238(2), 305-319.

Singer, M. J., & Witmer, B. G. (1999). On selecting the right yardstick. Presence: Teleoperators and Virtual Environments, 8(5), 566–573. https://doi.org/10.1162/105474699566486

Skarbez, R. (2016). Plausibility illusion in virtual environments.

Skarbez, R., Brooks, F. P., & Whitton, M. C. (2017). A survey of presence and related concepts. ACM Computing Surveys. https://doi.org/10.1145/3134301

Skrodzki, M. (2021). Illustrations of non-Euclidean geometry in virtual reality. *Virtual Images: Trilogy of Synthetic Realities I*, *5*, 192.

Slater, M. & M. Usoh (1993) Presence in Immersive Virtual Environments, IEEE Virtual Reality Annual International Symposium (VRAIS), September 18-22, Seattle, Washington, 90-96.

Slater, M. & M. Usoh (1994) Representation Systems, Perceptual Position and Presence in Virtual Environments, Presence: Teleoperators and Virtual Environments,

Slater, M. (1998). <Measuring Presence A Response to the Witmer and Singer Presence Questionnaire.pdf>. 1–13.

Slater, M. (2002). Presence and the sixth sense. Presence: Teleoperators and Virtual Environments, 11(4), 435–439. https://doi.org/10.1162/105474602760204327

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1535), 3549-3557. https://doi.org/10.1098/rstb.2009.0138

Slater, M., & Steed, A. (2000). A Virtual Presence Counter. 413–434.

Slater, M., & Usoh, M. (1993). Representations Systems, Perceptual Position, and Presence in Immersive Virtual Environments. Presence: Teleoperators and Virtual Environments, 2(3), 221–233. https://doi.org/10.1162/pres.1993.2.3.221

Slater, M., & Usoh, M. (2013). An experimental exploration of presence in virtual environments.

Slater, M., Banakou, D., Beacco, A., Gallego, J., Macia-Varela, F., & Oliva, R. (2022). A separate reality: An update on place illusion and plausibility in virtual reality. *Frontiers in Virtual Reality, 3*, 914392. https://doi.org/10.3389/frvir.2022.914392

Slater, M., Usoh, M., & Chrysanthou, Y. (1995). The Influence of Dynamic Shadows on Presence in Immersive Virtual Environments. https://doi.org/10.1246/nikkashi1948.74.305

Slater, M., Usoh, M., & Steed, A. (1994). Depth of Presence in Virtual Environment. 1–33. https://doi.org/10.1162/pres.1994.3.2.130

Slutsky, D. A., & Recanzone, G. H. (2001). Temporal and spatial dependency of the ventriloquism effect. AUDITORY AND VESTIBULAR SYSTEMS, 12(1), 7–10.

So, R. H., Lo, W. T., & Ho, A. T. (2001). Effects of navigation speed on motion sickness caused by an immersive virtual environment. Human factors, 43(3), 452-461.

Spence, C. (2007). Audiovisual multisensory integration. Acoustical Science and Technology, 28(2), 61–70. https://doi.org/10.1250/ast.28.61

Spence, C., & Squire, S. (2003). Multisensory integration: Maintaining the perception of synchrony. Current Biology, 13(13), 519–521. https://doi.org/10.1016/S0960-9822(03)00445-7

Spence, C., Baddeley, R., Zampini, M., James, R., & Shore, D. I. (2003). Multisensory temporal order judgments: When two locations are better than one. Perception and Psychophysics, 65(2), 318–328. https://doi.org/10.3758/BF03194803

Stanney, K. M., Hale, K. S., Nahmens, I., & Kennedy, R. S. (2003). What to expect from immersive virtual environment exposure: Influences of gender, body mass index, and past experience. *Human Factors, 45*(3), 504-520. https://doi.org/10.1518/hfes.45.3.504.27254

Stauffert, J. P., Niebling, F., & Latoschik, M. E. (2018, March). Effects of latency jitter on simulator sickness in a search task. In 2018 IEEE conference on virtual reality and 3D user interfaces (VR) (pp. 121-127). IEEE.

Stauffert, J. P., Niebling, F., & Latoschik, M. E. (2020). Latency and cybersickness: Impact, causes, and measures. A review. Frontiers in Virtual Reality, 1, 582204.

Stecker, G.C. (2018) Temporal binding of auditory spatial information across dynamic binaural events. *Atten Percept Psychophys* 80, 14–20. https://doi.org/10.3758/s13414-017-1436-0

Stein, B. E. (1993). *The merging of the senses*. MIT Press.

Stein, B. E., & Meredith, M. A. (1990). Multisensory integration. *Annals of the New York Academy of Sciences, 608*(1), 51-70. https://doi.org/10.1111/j.1749-6632.1990.tb48894.x

Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: Current issues from the perspective of the single neuron. Nature Reviews Neuroscience, 9(4), 255–266. https://doi.org/10.1038/nrn2331

Steuer, J. (1992). Defining virtual reality: dimensions determining telepresence, Communication in the age of virtual reality. Journal of Communication. https://doi.org/10.1111/j.1460-2466.1992.tb00812

Stevenson, R. A., & Wallace, M. T. (2013). Multisensory temporal integration: Task and stimulus dependencies. Experimental Brain Research. https://doi.org/10.1007/s00221-013-3507-3

Stevenson, R. A., Wilson, M. M., Powers, A. R., & Wallace, M. T. (2013). The effects of visual training on multisensory temporal processing. Experimental Brain Research. https://doi.org/10.1007/s00221-012-3387-y

Stevenson, R. A., Zemtsov, R. K., & Wallace, M. T. (2012). Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusions. Journal of Experimental Psychology: Human Perception and Performance, 38(6), 1517. https://doi.org/10.1037/a0027339

Stoffregen, T. A., Mantel, B., & Bardy, B. G. (2017). The senses considered as one perceptual system. *Ecological Psychology, 29*(3), 165-197.

Su, Y. H. (2014). Content congruency and its interplay with temporal synchrony modulate integration between rhythmic audiovisual streams. Frontiers in Integrative Neuroscience, 8, 92.

Sugita, Y., & Suzuki, Y. (2003). Audiovisual perception: Implicit estimation of sound-arrival time. Nature, 421(6926), 911. https://doi.org/10.1038/421911a

Suma, E. A., Lipps, Z., Finkelstein, S., Krum, D. M., & Bolas, M. (2012). Impossible spaces: Maximizing natural walking in virtual environments with self-overlapping architecture. IEEE Transactions on Visualization and Computer Graphics, 18(4), 555-564.

Tachibana, R., & Matsumiya, K. (2022). Accuracy and precision of visual and auditory stimulus presentation in virtual reality in Python 2 and 3 environments for human behavior research. *Behavior Research Methods, 54*(2), 729-751.

Talsma, & Woldorff, M. G. (2005). Selective attention and multisensory integration: Multiple phases of effects on the evoked brain activity. Journal of Cognitive Neuroscience, 17(7), 1098-1114. https://doi.org/10.1162/0898929054475172

Talsma, Doty, T. J., & Woldorff, M. G. (2007). Selective attention and audiovisual integration: Is attending to both modalities a prerequisite for early integration? Cerebral Cortex. https://doi.org/10.1093/cercor/bhk016

Talsma, Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. In Trends in Cognitive Sciences. https://doi.org/10.1016/j.tics.2010.06.008

Theeuwes, J. (1991). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. Perception & psychophysics, 49(1), 83-90.

Thesen, T., Vibell, J. F., Calvert, G. A., & Osterbauer, R. A. (2004). Neuroimaging of multisensory processing in vision, audition, touch, and olfaction. Cognitive Processing, 5(2), 84–93. https://doi.org/10.1007/s10339-004-0012-4

Tian, N., Lopes, P., & Boulic, R. (2022). A review of cybersickness in head-mounted displays: raising attention to individual susceptibility. Virtual Reality, 26(4), 1409-1441. https://doi.org/10.1007/s10055-022-00638-2

Todd, J. W. (1912). *Reaction to multiple stimuli* (No. 25). Science Press.

Traunmüller, H., & Öhrström, N. (2007). Audiovisual perception of openness and lip rounding in front vowels. Journal of Phonetics, 35(2), 244-258.

Usoh, M., Arthur, K., Whitton, M. C., Bastos, R., Steed, A., Slater, M., & Brooks Jr, F. P. (1999, July). Walking > walking-in-place > flying in virtual environments. In *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques* (pp. 359-364). https://doi.org/10.1145/311535.311589

van de Rijt, L. P. H., Roye, A., Mylanus, E. A. M., Opstal, A. J. van, & van Wanrooij, M. M. (2019). The principle of inverse effectiveness in audiovisual speech perception. BioRxiv, 585182. https://doi.org/10.1101/585182

Van der Burg, E., Alais, D., & Cass, J. (2018). Rapid recalibration to audiovisual asynchrony follows the physical—not the perceived—temporal order. Attention, Perception, and Psychophysics, 33(37), 1–9. https://doi.org/10.3758/s13414-018-1540-9

Van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W., & Theeuwes, J. (2008). Pip and Pop: Nonspatial Auditory Signals Improve Spatial Visual Search. Journal of Experimental

Psychology: Human Perception and Performance, 34(5), 1053–1065. https://doi.org/10.1037/0096-1523.34.5.1053

Van Eijk, R. L. J., Kohlrausch, A., Juola, J. F., & Van De Par, S. (2008). Audiovisual synchrony and temporal order judgments: Effects of experimental method and stimulus type. Perception and Psychophysics, 70(6), 955–968. https://doi.org/10.3758/PP.70.6.955

van Emmerik, M. L., de Vries, S. C., & Bos, J. E. (2011). Internal and external fields of view affect cybersickness. Displays, 32(4), 169-174.

van Wassenhove, V., Grant, K. W., & Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. Neuropsychologia, 45(3), 598–607. https://doi.org/10.1016/j.neuropsychologia.2006.01.001

Van Waveren, J. M. P. (2016, November). The asynchronous time warp for virtual reality on consumer hardware. In Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology (pp. 37-46).

Vasconcelos-Raposo, J., Bessa, M., Melo, M., Barbosa, L., Rodrigues, R., Teixeira, C. M., Cabra, L., & Sousa, A. A. (2016). Adaptation and Validation of the Igroup Presence Questionnaire (IPQ) in a Portuguese Sample. Presence Teleoperators & Virtual Environments, 25(3), 191–203. https://doi.org/10.1162/PRES

Vasylevska, K., & Kaufmann, H. (2015, October). Influence of Path Complexity on Spatial Overlap Perception in Virtual Environments. In ICAT-EGVE (pp. 159-166).

Vatakis, A., & Spence, C. (2007). Crossmodal binding: Evaluating the “unity assumption” using audiovisual speech stimuli. Perception and Psychophysics. https://doi.org/10.3758/BF03193776

Vatakis, A., Navarra, J., Soto-Faraco, S., & Spence, C. (2008). Audiovisual temporal adaptation of speech: Temporal order versus simultaneity judgments. Experimental Brain Research, 185(3), 521–529. https://doi.org/10.1007/s00221-007-1168-9

Vorderer, P, Wirth, W., Saari, T., Gouveia, F. R., Biocca, F., Jäncke, F., Böcking, S., Hartmann, T., Klimmt, C., Schramm, H., Laarni, J., Ravaja, N., Gouveia, L. B., Rebeiro, N., Sacau, A., Baumgartner, T. & Jäncke , P. (2004). Development of the MEC Spatial Presence Questionnaire (MEC-SPQ). Unpublished report to the European Community, Project Presence: MEC (IST-2001-37661). Hannover, Munich, Helsinki, Porto, Zurich.

Vorderer, P., Wirth, W., Gouveia, F. R., Biocca, F., Saari, T., Jäncke, L., Böcking, S., Schramm, H., Gysbers, A., Hartmann, T., Klimmt, C., Ravaja, N., Sacau, A., Baumgartner, T., & Jäncke, P. (2004). MEC Spatial Presence Questionnaire (MEC-SPQ). Short Documentation and Instructions for Application. Report to the European Community, Project Presence: MEC (IST-2001-37661)., June.

Wallace, M. T. (2004). The development of multisensory processes. Cognitive Processing. https://doi.org/10.1007/s10339-004-0017-z

Wallace, M. T., & Stevenson, R. A. (2014). The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. In Neuropsychologia. https://doi.org/10.1016/j.neuropsychologia.2014.08.005

Warburton, M., Mon-Williams, M., Mushtaq, F., & Morehead, J. R. (2023). Measuring motion-to-photon latency for sensorimotor experiments with virtual reality systems. Behavior research methods, 55(7), 3658-3678. doi: 10.3758/s13428-022-01983-5.

Warren, W. H., Rothman, D. B., Schnapp, B. H., & Ericson, J. D. (2017). Wormholes in virtual space: From cognitive maps to cognitive graphs. Cognition, 166, 152-163. https://doi.org/10.1016/j.cognition.2017.05.020

Warwick-Evans, L., Symons, N., Fitch, T., & Burrows, L. (1998). Evaluating sensory conflict and postural instability. Theories of motion sickness. Brain research bulletin, 47(5), 465-469.

Weech, S., Kenny, S., & Barnett-Cowan, M. (2019). Presence and cybersickness in virtual reality are negatively related: A review. *Frontiers in Psychology, 10*, 158. https://doi.org/10.3389/fpsyg.2019.00158

Weigelt, K., & Wiemeyer, J. (2012). Spatial perception and spatial experience in stereoscopic 3D sports broadcasts. 2012 International Conference on 3D Imaging, IC3D 2012 - Proceedings, 1–11. https://doi.org/10.1109/IC3D.2012.6615135

Weißker, T., Kunert, A., Fröhlich, B., & Kulik, A. (2018, March). Spatial updating and simulator sickness during steering and jumping in immersive virtual environments. In 2018 IEEE conference on virtual reality and 3D user interfaces (VR) (pp. 97-104). IEEE.

Welch, R. B. (2003). Adapting to telesystems. In *Virtual and Adaptive Environments* (pp. 129-167). CRC Press.

Westland, J. C. (2010). Lower bounds on sample size in structural equation modeling. Electronic commerce research and applications, 9(6), 476-487.

Wilson, J. R., & Nichols, S. C. (2002). Measurement in virtual environments: another dimension to the objectivity/subjectivity debate. Ergonomics, 45(14), 1031-1036.

Wilson, J. R., Nichols, S., & Haldane, C. (1997). Presence and side effects: Complementary or contradictory? *Advances in Human Factors/Ergonomics, 21*, 889-892.

Wilson, K. A. (2023). Windows on time: unlocking the temporal microstructure of experience. *Review of Philosophy and Psychology*, *14*(4), 1197-1218.

Wirth, W., Hartmann, T., Böcking, S., Vorderer, P., Klimmt, C., Schramm, H., Saari, T., Laarni, J., Ravaja, N., Gouveia, F. R., Biocca, F., Sacau, A., Jäncke, L., Baumgartner, T., & Jäncke, P. (2007). A process model of the formation of spatial presence experiences. Media Psychology, 9(3), 493–525. https://doi.org/10.1080/15213260701283079

Witmer, B. G., & Singer, M. J. (1998). Measuring Presence in Virtual Environments - A Presence Questionaire. Presence: Teleoperators and Virtual Environments, 3(???), 225–240. https://doi.org/https://doi.org/10.1162/105474698565686

Witmer, B. G., Jerome, C. J., & Singer, M. J. (2005). The factor structure of the Presence Questionnaire. Presence: Teleoperators and Virtual Environments, 14(3), 298–312. https://doi.org/10.1162/105474605323384654

Wolff, A., Berberian, N., Golesorkhi, M., Gomez-Pilar, J., Zilio, F., & Northoff, G. (2022). Intrinsic neural timescales: temporal integration and segregation. *Trends in cognitive sciences*, *26*(2), 159-173.

Yao, R., Heath, T., Davies, A., Forsyth, T., Mitchell, N., & Hoberman, P. (2014). Oculus vr best practices guide. Oculus VR, 4, 27-35.

Youngblut, C. (2003). Experience of Presence in Virtual Environments. Institute for Defense Analyses, September, 158.

Zhang, D., & Luo, Y. (2012, June). Single-trial ERPs elicited by visual stimuli at two contrast levels: Analysis of ongoing EEG and latency/amplitude jitters. In 2012 IEEE Symposium on Robotics and Applications (ISRA) (pp. 85-88). IEEE.

Zilka, R., & Bonneh, Y. (2022). Individual differences in audio-visual binding can predict the varied severity of motion sickness. *bioRxiv*. https://doi.org/10.1101/2022.05.

Appendices

Appendix 1: Gaming Experience Questionnaire Adapted from Lemmens et al., (2011)

Time spent on games.

The following questions ask about your previous **experience** in regard to **gaming and virtual reality**. Please answer the following using **NUMERIC VALUES**.

1. How many days a week do you play games on a console/pc/ handheld/mobile phone?

……
……

2. On an average day that you play games on a pc/console/handheld, how much time do you spend playing?

……
……

3. How many times have you used virtual reality equipment before this experiment?

……
……

4. How long did you spend using the virtual reality equipment on average per session? (Please leave blank if above answer is Zero)

……
……

Appendix 2: Important Health and Safety Information

Important Information

Some people report feeling disoriented, experiencing short term motion sickness, and other adverse effects, whilst using the Oculus Rift headset. These side effects are more likely to occur if you have a history of motion sickness (e.g. you get car sick) or you have photosensitive epilepsy. They are also more likely if you are currently hungover or suffering from ill health (e.g. a cold or flu). The Oculus Rift is also not appropriate if you are; currently pregnant, on prescribed medication, have a history of cardiac conditions/high blood pressure, under the age of 13, using medical equipment such as a pacemaker or hearing aids, currently suffering from a contagious condition of the eye or skin, have a binocular vison abnormality or a psychiatric condition, or if you are intending to drive or operate heavy machinery shortly afterwards.

You must advise the researchers if any of the circumstances above apply to you. You do not need to specify which one.

If these *do not* apply to you but you do feel unwell at any point when using the equipment, or shortly afterwards, you must inform the researchers immediately.

It is worth noting that this equipment is still in the early stages of development and all potential ill effects are not currently known. If you wish to you can view the formal risk assessment of the equipment before taking part.

We ask that you sign below to confirm that you are now aware of the above information and have discussed this with the researchers where appropriate.

Name of participant: ……..…….

Signed: ……. Date: ……..

Name of researcher: ……..…….

Signed: ……. Date: ……..

This project has been approved by the Health and Life Sciences Research Ethics Committee

Appendix 3: Adapted MATLAB Scripts from Notter, Murray & Micah (2017)

```
 % This script computes the temporal binding window (TBW) for each subjects
in a

 % multisensory integration study. The computed output will look like the
ones

 % from:

 %     Stevenson, R. A., Zemtsov, R. K., & Wallace, M. T. (2012). Individual

 %     differences in the multisensory temporal binding window predict

 %     susceptibility to audiovisual illusions. Journal of Experimental

 %     Psychology: Human Perception and Performance, 38(6), 1517.

 %     http://dx.doi.org/10.1037/a0027339

 %

 % But in contrast to their approach, the following two changes were done:

 %     First, this script splits the subject data into two equally big parts
to

 %         create the AV and VA condition.

 %     Second, the resulting temporal binding window curve (the curve that

 %         combines the two sigmoid functions, was not resized to reach 100%.

 %

 % For more information about this script see:

 %     https://github.com/miykael/temporal_binding_window

 %

 % Syntax:  [subj] = tbw_subj(xlsxfile, hoi, x_bound)

 %

 % Input parameters:

 %    xlsxfile - Path to the xlsxfile containing subject information such as
```

```
 %               subject_ids, category_id, offsets & simultaneity information

 %    hoi      - (optional) Hight of interest to compute TBW (default 50%)

 %    x_bound  - (optional) Range between which to compute TBW (default -
+750ms)

 %

 % Output parameters:

 %    subj     - This output variable contains the b-values and the binding

 %               points for the AV and VA side of the TBW-curve, the TBW
point

 %               and the temporal bind curve for each category for each
subject.

 %

 % Examples:

 % >> tbw_subj('dataset.xlsx')

 % >> tbw_subj('dataset.xlsx', 0.5, 750)

 %

 % See also: glmfit, glmval



 % Author:   Michael Notter

 % History:  25.05.2017  file created



 function [subj] = tbw_subj1
 ('TOJdataset.xlsx')



 % Set input parameters

 sampling_rate = 0.1;    % sampling rate for TBW-curve

 if ~exist('hoi','var')
```

```
    hoi = 0.5;          % Hight of interest (hoi) to compute TBW

end

if ~exist('x_bound','var')

    x_bound = 750
    ;      % Left and Right bound in ms for the estimation of TBW

end



% Read data from xlsx-file

data = xlsread('TOJdataset.xlsx', 1);

subj_list = data(:, 1);      % Array indicating subject number

categ = data(:, 3);          % Array indicating category number

offset = data(:, 4);         % Array indicating stimuli offset (x-axis)

simultaneity = data(:, 5);   % Array indicating stimuli percept, 0 or 1 (y-
axis)



% Run computation for each subject separately

subjects = unique(subj_list)';

for s = subjects



    % Extract data of current subject

    s_id = subj_list==s;

    sOffset = offset(s_id);

    sSimultaneity = simultaneity(s_id);

    sCateg = categ(s_id);
```

```
    % Run computation for each category separately, and add an additional
    %   category that includes all elements of a category
    categories = [unique(sCateg)', max(sCateg) + 1];
    for c = categories

        % Select elements of current category
        if c == max(sCateg) + 1
            selector = true(length(sCateg), 1);
        else
            selector = sCateg==c;
        end
        cOffset = sOffset(selector);
        cSimultaneity = sSimultaneity(selector);

        % Sort everything by offset
        [cOffset, sort_id] = sort(cOffset);
        cSimultaneity = cSimultaneity(sort_id);

        % Execute glmfit for AV and VA condition (split offset in middle)
        split = ceil(length(cOffset) / 2);
        [b_AV, dev_AV, stats_AV] = glmfit(cOffset(1:split), ...
            cSimultaneity(1:split), 'binomial', 'link', 'logit');
        [b_VA, dev_VA, stats_VA] = glmfit(cOffset(split+1:length(cOffset)),
...
            cSimultaneity(split+1:length(cOffset)), 'binomial', 'link',
'logit');
```

```
% Sampling points for TBW curve
x_line = [-x_bound:sampling_rate:x_bound]';

% Fit sigmoid function to estimated data
y_AV = glmval(b_AV, x_line, 'logit');
y_VA = glmval(b_VA, x_line, 'logit');

% Estimate temporal binding window at hight of interest
[minimum_AV, min_AV_id] = min(abs(y_AV - hoi));
bind_AV = x_line(min_AV_id);

[minimum_VA, min_VA_id] = min(abs(y_VA - hoi));
bind_VA = x_line(min_VA_id);

temp_bind_window = -1 * bind_AV + bind_VA;

% Compute temporal binding curve
cut_id = find(y_AV > y_VA, 1);
temp_bind_curve = [y_AV(1:cut_id) ; y_VA(cut_id+1:end)];

% Store relevant values in subject variable for later use
subj(s).categ(c).b_AV = b_AV;
subj(s).categ(c).b_VA = b_VA;
subj(s).categ(c).bind_AV = bind_AV;
```

```
        subj(s).categ(c).bind_VA = bind_VA;

        subj(s).categ(c).temp_bind_window = temp_bind_window;

        subj(s).categ(c).temp_bind_curve = temp_bind_curve;

    end

    subj(s).x_line = x_line;


    % Create overwiev plot for given subject

    figure;

    set(gcf,'Renderer','OpenGL');

    plot(x_line, [subj(s).categ.temp_bind_curve]);

    title(['Subject: ', num2str(s)]);

    xlabel(sprintf('Stimulus Onset Asynchrony [ms]\n(from AV to VA)'));

    ylabel('%-Perceived Synchronous');

    txt_legend = [sprintf(repmat('cat_%d ',1,length(categories)-1), ...
        categories(1:end-1)), 'all'];

    fig_legend = strsplit(txt_legend);

    legend(fig_legend);

    ylim([0 1]);


    hold on

    % Plot hight of interest

    xlimits = get(gca,'xlim');

    plot(xlimits, [hoi hoi], '--', 'Color',[0.5,0.5,0.5]);

    hold off
```

```
    % Plot information about binding points in figure

    txt_AV = sprintf(['AV: ', repmat('%5d',1,length(categories))], ...
        round([subj(s).categ.bind_AV]));

    txt_VA = sprintf(['VA: ', repmat('%5d',1,length(categories))], ...
        round([subj(s).categ.bind_VA]));

    txt_TBW = sprintf(['TBW:', repmat('%5d',1,length(categories))], ...
        round([subj(s).categ.temp_bind_window]));

    text(xlimits(1) * 0.9, 0.95,['      ', txt_legend], ...
        'HorizontalAlignment','left', 'FontName', 'FixedWidth');

    text(xlimits(1) * 0.9, 0.90,txt_AV, ...
        'HorizontalAlignment','left', 'FontName', 'FixedWidth');

    text(xlimits(1) * 0.9, 0.85,txt_VA, ...
        'HorizontalAlignment','left', 'FontName', 'FixedWidth');

    text(xlimits(1) * 0.9, 0.80,txt_TBW, ...
        'HorizontalAlignment','left', 'FontName', 'FixedWidth');



    % Save Figure

    pbaspect([1.2 1 1])
    print(sprintf('result07_02_sub%.2d', s),'-dpng')
    close()

end
function [group] = tbw_group(subj, threshold)
% Set input parameters
if ~exist('threshold','var')

    threshold = 0;      % Threshold of minimum TBW to include subject

end

if ~exist('hoi','var')

    hoi = 0.5;          % Hight of interest (hoi) to compute TBW

end



nSubj = length(subj);
```

```
nCate = length(subj(1).categ);


% Read data from subj parameter if TBW is above threshold
bind_AV = cell(nCate, 1);
bind_VA = cell(nCate, 1);
temp_bind_window = cell(nCate, 1);
temp_bind_curve = cell(nCate, 1);
subj_used = 0;
for s=1:nSubj


    % Drop subject if smalles TBW is below threshold
    if min([subj(s).categ.temp_bind_window]) > threshold


        for c=1:nCate


            bind_AV{c} = [bind_AV{c} subj(s).categ(c).bind_AV];
            bind_VA{c} = [bind_VA{c} subj(s).categ(c).bind_VA];
            temp_bind_window{c} = [temp_bind_window{c}
subj(s).categ(c).temp_bind_window];
            temp_bind_curve{c} = [temp_bind_curve{c}
subj(s).categ(c).temp_bind_curve];


        end
        subj_used = subj_used + 1;
    end
```

```
end


% Compute temporal binding window plot for each category
for c=1:nCate


    % Get relevant Data
    x = -1 * bind_AV{c}';
    y = bind_VA{c}';


    % Compute correlation statistic
    [r p] = corr(x,y);


    % Plot Figure
    figure;
    set(gcf,'Renderer','OpenGL');
    plot(x, y, 'x')
    if c==nCate
        category = 'all';
    else
        category = num2str(c);
    end
    title(['Temporal Binding Window for Category: ', category]);
    xlabel(sprintf('Left\nTemporal Binding Window [ms]'));
    ylabel(sprintf('Right\nTemporal Binding Window [ms]'));
```

```
    xlimits = get(gca,'xlim');

    ylimits = get(gca,'ylim');

    text(xlimits(1)*1.1, ylimits(2)*0.95,['r = ' num2str(r)], ...
        'HorizontalAlignment','left', 'FontName', 'FixedWidth');

    text(xlimits(1)*1.1, ylimits(2)*0.90,['p = ' num2str(p)], ...
        'HorizontalAlignment','left', 'FontName', 'FixedWidth');

    text(xlimits(1)*1.1, ylimits(2)*0.85,['n = ' num2str(subj_used)], ...
        'HorizontalAlignment','left', 'FontName', 'FixedWidth');

    lsline



    % Plot value and p-value of mean-TBW

    [h,ptest] = ttest(temp_bind_window{c}, 0, 0.05, 'right');

    txt_meanTBW = sprintf(['Mean TBW: ', num2str(mean(temp_bind_window{c})),
...

        '\np-value:  ', num2str(ptest)]);

    text(xlimits(1)*1.1, ylimits(2)*0.25, txt_meanTBW, ...
        'HorizontalAlignment','left', 'FontName', 'FixedWidth');



    pbaspect([1.2 1 1])

    print(sprintf('result_TBW_categ%.2d', c),'-dpng')

    close()



end



% Compute mean temporal binding curve for each category

x_line = subj(1).x_line;

for c=1:nCate
```

```
    % Get curves
    x_curves = temp_bind_curve{c};

    % Plot Figure
    figure;
    set(gcf,'Renderer','OpenGL');
    shadedErrorBar(x_line, x_curves', {@mean, @(x_curves) std(x_curves)},
...
        {'-k', 'LineWidth', 2}, 0);
    if c==nCate
        category = 'all';
    else
        category = num2str(c);
    end
    title(['Mean Temporal Binding Window for Category: ', category]);
    xlabel(sprintf('Stimulus Onset Asynchrony [ms]\n(from AV to VA)'));
    ylabel('%-Perceived Synchronous');
    ylim([0 1]);

    hold on
    % Plot hight of interest
    xlimits = get(gca,'xlim');
    plot(xlimits, [hoi hoi], '--', 'Color',[0.5,0.5,0.5]);
```

```
    % Find TBW of average curve

    tbw_diff = abs(mean(x_curves,2) - hoi);

    minima_id = find(imregionalmin(tbw_diff));

    minima_val = x_line(minima_id);

    plot([minima_val(1), minima_val(1)], [0, hoi], '-', 'Color',[0, 0, 0]);

    plot([minima_val(end), minima_val(end)], [0, hoi], '-', 'Color',[0, 0,
0]);

    text(minima_val(1) * 0.9, 0.25, num2str(minima_val(1)), ...
        'HorizontalAlignment','left');

    text(minima_val(end) * 0.9, 0.25, num2str(minima_val(end)), ...
        'HorizontalAlignment','right');

    hold off



    pbaspect([1.2 1 1])

    print(sprintf('TOJ_TBC_categ%.2d', c),'-dpng')

    close()



 end
```

Appendix 4: MATLAB Scripts from Alcalá-Quintana and García-Pérez Independent Channels Model

```
 function [report] = fit_TOJ(TOJdata, LamBounds, TauBounds, DeltaBounds, ...
       LamT_Start, LamR_Start, Tau_Start, Delta_Start, Err_Start, Bias_Start,
...
       Model, Plot, Disp)

 load('TOJData.mat');

 TF=[];
 RF=[];

 oi = 149
     TF(1, :) = double(table2array(TOJData(oi, 3:14)));
     RF(1, :) = double(table2array(TOJData(oi+1, 3:14)));
     Part(1,:) = double(table2array(TOJData(oi,1)))

 TOJdata = [-300 -250 -200 -150 -100 -50 50 100 150 200 250 300; ...
     TF;...
     RF];


   LamBounds = [1/200 1/3]; TauBounds = [-Inf Inf]; DeltaBounds = [0 Inf];
   LamT_Start = [1/70 1/10]; LamR_Start = [1/70 1/10]; Tau_Start = [-70 70];
Delta_Start = [20 150]; Err_Start = [0.05]; Bias_Start =[0.5]; Model = 1;
Plot =true; Disp = true;



 %
 % Fit model to data from a TOJ task
 %
 % Usage: o = fit_TOJ(SJ3data, LamBounds, TauBounds, DeltaBounds, ...
 %                    LamT_Start, LamR_Start, Tau_Start, ...
 %                    Delta_Start, Err_Start, Bias_Start, Model, Plot, Disp)
 % Arguments:
 %
 % - TOJdata:     3-by-N_TOJ array with data from the TOJ task. The first row
 %                contains the N_SJ2 stimulus levels; the second row contains
the
 %                count of TF responses at each level; the third row contains
the
 %                count of RF responses at each level
 %
 % - LamBounds:   2-element vector with the lower and upper bounds on
lambda_r and
 %                lambda_t
 %
```

```
 % - TauBounds:   2-element vector with the lower and upper bounds on tau
 %
 % - DeltaBounds: 2-element vector with the lower and upper bounds on delta
 %
 % - LamT_Start:  Starting value(s) for lambda_t. Scalar or vector of
positive reals
 %
 % - LamR_Start:  Starting value(s) for lambda_r. Scalar or vector of
positive reals
 %
 % - Tau_Start:   Starting value(s) for tau. Scalar or vector of reals
 %
 % - Delta_Start: Starting value(s) for delta. Scalar or vector of positive
reals
 %
 % - Err_Start:   Starting value(s) for epsilon. Scalar or vector of reals in
[0, 1]
 %
 % - Bias_Start:  Starting value(s) for xi and kappa. Scalar or vector of
reals in [0, 1]
 %
 % - Model:       Choice of model to be fitted. An integer scalar in the
range between 0
 %                and 3 (implying different assumptions about error
parameters; see the
 %                table underneath) or the string 'best' (case insensitive)
to find the
 %                best-fitting model by the BIC criterion
 %
 %                              ___TOJ____
 %                    Model #   e_TF  e_RF
 %                    -------   ----------
 %                          0    x     x
 %                          1    o     o
 %                          2    o     x
 %                          3    x     o
 %                    ----------------------------
 %                              x : parameter is not included in the model
 %                              o : parameter is included in the model
 %
 % - Plot:        Option to plot data and fitted functions. Logical scalar
 %
 % - Disp:        Option to issue warnings or display progress information.
Logical scalar
 %
 %
 % Output: The output is a structure including parameter estimates and
goodness-of-fit
 %         measures and p-values
```

```
 %
 % Further details in:  Alcalá-Quintana, R. & García-Pérez, M.A. (2013).
Fitting model-based
 %                          psychometric functions to simultaneity and
temporal-order judgment
 %                          data: MATLAB and R routines. Behavior Research
Methods, in press.
 %                          http://dx.doi.org/10.3758/s13428-013-0325-2

 % ------------------Check conditions-------------------------------------
 if ndims(TOJdata)>2|| size(TOJdata,1)~=3 || ~isnumeric(TOJdata) ||
~isreal(TOJdata)
   error('Invalid TOJdata (must be a 2D numeric matrix with 3 rows)')
 elseif ~isnumeric(LamBounds) || ~isreal(LamBounds) || any(LamBounds<=0)
   error('Invalid LamBounds (must be positive reals)')
 elseif numel(LamBounds)~=2
   error('Invalid size of LamBounds (must have two components only)')
 elseif LamBounds(1)==LamBounds(2)
   error('Invalid contents of LamBounds (components must have different
values)')
 elseif ~isnumeric(TauBounds) || ~isreal(TauBounds)
   error('Invalid TauBounds (must be reals)')
 elseif numel(TauBounds)~=2
   error('Invalid size of TauBounds (must have two components only)')
 elseif TauBounds(1)==TauBounds(2)
   error('Invalid contents of TauBounds (components must have different
values)')
 elseif ~isnumeric(DeltaBounds) || ~isreal(DeltaBounds) || any(DeltaBounds<0)
   error('Invalid DeltaBounds (must be positive reals)')
 elseif numel(DeltaBounds)~=2
   error('Invalid size of DeltaBounds (must have two components only)')
 elseif DeltaBounds(1)==DeltaBounds(2)
   error('Invalid contents of DeltaBounds (components must have different
values)')
 elseif ~isnumeric(LamT_Start) || ~isreal(LamT_Start)
   error('Invalid LamTStart (must be real)')
 elseif ~isnumeric(LamR_Start) || ~isreal(LamR_Start)
   error('Invalid LamRStart (must be real)')
 elseif ~isnumeric(Tau_Start) || ~isreal(Tau_Start)
   error('Invalid TauStart (must be real)')
 elseif ~isnumeric(Delta_Start) || ~isreal(Delta_Start)
   error('Invalid DeltaStart (must be real)')
 elseif ~isnumeric(Err_Start) || ~isreal(Err_Start)
   error('Invalid ErrStart (must be real)')
 elseif ~isnumeric(Bias_Start) || ~isreal(Bias_Start)
   error('Invalid BiasStart (must be real)')
 elseif any(any(TOJdata(2:3,:)<0))
   error('Negative counts in rows 2-3 of TOJdata')
 elseif ~strcmpi(Model,'best') && (~isscalar(Model) || ~isnumeric(Model))
```

```
   error('Invalid value for Model (must be the string ''best'' or an integer
scalar between 0 and 3)')
 elseif isnumeric(Model) && (~isreal(Model) || mod(Model,1)~=0 || Model<0 ||
Model>3)
   error('Invalid value for Model (must be the string ''best'' or an integer
scalar between 0 and 3)')
 elseif ~islogical(Plot) || ~isscalar(Plot)
   error('Invalid value for Plot (must be a logical scalar)')
 elseif ~islogical(Disp) || ~isscalar(Disp)
   error('Invalid value for Disp (must be a logical scalar)')
 end
 % ------------------Main----------------------------------------------------
 s1 = warning('off','MATLAB:log:logOfZero');
 s2 = warning('off','MATLAB:divideByZero');
 LamBounds = sort(reshape(LamBounds,1,2));
 TauBounds = sort(reshape(TauBounds,1,2));
 DeltaBounds = sort(reshape(DeltaBounds,1,2));
 LamT_Start = reshape(LamT_Start,1,numel(LamT_Start));
 LamR_Start = reshape(LamR_Start,1,numel(LamR_Start));
 Tau_Start = reshape(Tau_Start,1,numel(Tau_Start));
 Delta_Start = reshape(Delta_Start,1,numel(Delta_Start));
 Err_Start = reshape(Err_Start,1,numel(Err_Start));
 Bias_Start = reshape(Bias_Start,1,numel(Bias_Start));
 LevTOJ = TOJdata(1,:); TOJ_TF = TOJdata(2,:); TOJ_RF = TOJdata(3,:);
 NforBIC = sum(TOJ_TF+TOJ_RF>0);
 options = optimset('MaxIter',500, 'MaxFunEvals',2000, ...
                    'TolCon',1e-4, 'TolFun',1e-4, 'TolX',1e-4, ...
                    'LargeScale','off', 'Display','off');
 rel = version('-release'); rel = sscanf(rel,'%i4');
 if rel==2008 || rel==2009, options = optimset(options, 'Algorithm',
'interior-point'); end
 if rel>=2010, options = optimset(options, 'Algorithm', 'sqp'); end
 if Model==0, Err_Start=0; end
 Models = Model; if strcmpi(Model,'best'), Models = 0:3; end;
NumModels=numel(Models);
 NumIter=length(LamT_Start)*length(LamR_Start)*length(Tau_Start)*length(Delta
_Start)* ...
         length(Err_Start)*length(Bias_Start);
 ChosenModel = NaN;
 BICmin = realmax;
 for ModelLoop = Models
   strm = ['Tried ',num2str(ModelLoop+1),' of ',num2str(NumModels),' models;
'];
 % proceed through initial values
   niter = 0;
   iv_pars = NaN(NumIter,21);
   iv_value = NaN(NumIter,1);
   iv_flag = NaN(NumIter,1);
   iv_iterations = NaN(NumIter,1);
```

```
   iv_funcCount = NaN(NumIter,1);
   OnBound = NaN(NumIter,1);
   for LamTInit = LamT_Start
   for LamRInit = LamR_Start
   for TauInit = Tau_Start
   for DeltaInit = Delta_Start
   for ErrInit = Err_Start
   for BiasInit = Bias_Start
     niter = niter+1;
     str = ['Tried ',num2str(niter),'/',num2str(NumIter),' sets of initial
values; '];
     if ModelLoop==0
       Initial = [LamTInit LamRInit TauInit DeltaInit BiasInit];
       lb = [LamBounds(1) LamBounds(1) TauBounds(1) DeltaBounds(1) 0];
       ub = [LamBounds(2) LamBounds(2) TauBounds(2) DeltaBounds(2) 1];
     elseif ModelLoop==1
       Initial = [LamTInit LamRInit TauInit DeltaInit BiasInit ErrInit
ErrInit];
       lb = [LamBounds(1) LamBounds(1) TauBounds(1) DeltaBounds(1) 0 0 0];
       ub = [LamBounds(2) LamBounds(2) TauBounds(2) DeltaBounds(2) 1 1 1];
     elseif ModelLoop==2 || ModelLoop==3
       Initial = [LamTInit LamRInit TauInit DeltaInit BiasInit ErrInit];
       lb = [LamBounds(1) LamBounds(1) TauBounds(1) DeltaBounds(1) 0 0];
       ub = [LamBounds(2) LamBounds(2) TauBounds(2) DeltaBounds(2) 1 1];
     end
     npars = length(Initial);
     % estimate parameters
     [xmin, value, flag, output] = fmincon(@LogLikelihood, Initial, ...
           [], [], [], [], lb, ub, [], options, ...
           TOJ_TF, TOJ_RF, LevTOJ, ModelLoop);
     iv_pars(niter,1:npars) = xmin;
     iv_value(niter) = value;
     iv_flag(niter) = flag;
     iv_iterations(niter) = output.iterations;
     iv_funcCount(niter) = output.funcCount;
     OnBound(niter) = any(xmin(1:2)==LamBounds(2));
     if NumModels==1 && Disp, disp([str,'Obj.func. = ',num2str(value)]), end
   end
   end
   end
   end
   end
   end
   [v1 best] = min(iv_value);
   if OnBound(best) && any(~OnBound)
     iv_value = iv_value.*(OnBound+1);
     [v2 best2] = min(iv_value);
     if v2<=v1*1.005, best = best2; end
   end
```

```
   temp_pars = iv_pars(best,1:npars);
   temp_flag = iv_flag(best);
   temp_iterations = iv_iterations(best);
   temp_funcCount = iv_funcCount(best);
   ObjectiveFunctionMin = iv_value(best);
   BIC = 2*ObjectiveFunctionMin + npars*log(NforBIC);
   if BIC<BICmin
     BICmin=BIC;
     ChosenModel = ModelLoop;
     Final_flag = temp_flag;
     Final_npars = npars;
     Final_pars = temp_pars;
     Final_iterations = temp_iterations;
     Final_funcCount = temp_funcCount;
     if NumModels>1 && Disp
       disp([strm,' BIC = ',num2str(BIC),' ; best thus far ...'])
     end
   elseif NumModels>1 && Disp
     disp([strm,' BIC = ',num2str(BIC)])
   end
 end
 npars = Final_npars; pars = Final_pars; BIC = BICmin;
 % extract estimated parameters
 [lam_T lam_R tau delta_TOJ xi epskap] = GetParVec(pars, ChosenModel);
 if Disp
   msg1 = 'Consider re-running with a smaller lower bound in LamBounds';
   msg2 = 'Potentially non-informative data';
   if lam_T==LamBounds(1), disp(['Warning: Estimated lambda_T at the lower
bound. ' msg1]),end
   if lam_R==LamBounds(1), disp(['Warning: Estimated lambda_R at the lower
bound. ' msg1]),end
   if lam_T==LamBounds(2), disp(['Warning: Estimated lambda_T at the upper
bound. ' msg2]),end
   if lam_R==LamBounds(2), disp(['Warning: Estimated lambda_R at the upper
bound. ' msg2]),end
 end
 % -----------------Goodness of fit test----------------------------------
 % observed frequencies
 O = [TOJ_TF TOJ_RF];
 % number of observations per delay
 n_TOJ = TOJ_TF + TOJ_RF;
 % theoretical probabilities
 P_TOJ_TF = psi_TOJ_TF(LevTOJ, lam_T, lam_R, tau, delta_TOJ, xi, epskap);
 P_TOJ_RF = 1 - P_TOJ_TF;
 % expectations
 E = [P_TOJ_TF.*n_TOJ P_TOJ_RF.*n_TOJ];
 % degrees of freedom
 dof = sum(n_TOJ>0) - npars;
 % index of fit and significance
```

```
 EE = E(E>0); OO = O(E>0);
 G2 = 2*nansum(OO.*log(OO./EE));
 X2 = sum((OO-EE).^2./EE);
 pG2 = 'Insufficient degrees of freedom';
 pX2 = 'Insufficient degrees of freedom';
 if dof > 0
   pG2 = 1 - chi2cdf(G2, dof);
   pX2 = 1 - chi2cdf(X2, dof);
 end
 if Plot, PlotIt(pars, LevTOJ, ChosenModel, TOJ_TF, TOJ_RF); end
 % rearrange parameters to return
 params = [pars(1:4) epskap(1) xi epskap(2)];
 if ChosenModel==0, params(5)=NaN; params(7)=NaN;
 elseif ChosenModel==2, params(7)=NaN;
 elseif ChosenModel==3, params(5)=NaN;
 end
 % create report
 perfmes = [params(1),params(2),params(3),params(4),params(6),-400,400]
 save('perfmes','perfmes')
 report = struct('Problem', ['Separate fit of TOJ data with '
output.algorithm],...
                 'FlagFrom_fmincon', Final_flag, ...
                 'NumIterations', Final_iterations, ...
                 'NumFuncCount', Final_funcCount, ...
                 'UserModel', Model, ...
                 'FittedModel', ChosenModel, ...
                 'NumFreeParameters', npars, ...
                 'NumCells', 2*sum(n_TOJ>0), ...
                 'NumExpBelow5', sum(E<5), ...
                 'NumExpBelow5_ObsAbove0', sum((E<5).*(O>0)), ...
                 'NumExpBelow1', sum(E<1), ...
                 'NumExpBelow1_ObsAbove0', sum((E<1).*(O>0)), ...
                 'DegreesOfFreedom', dof, ...
                 'ChiSquareStatistic', X2, ...
                 'ChiSquareP_value', pX2, ...
                 'LikelihoodRatioStatistic', G2, ...
                 'LikelihoodRatioP_value', pG2, ...
                 'BIC', BIC, ...
                 'LamBounds', LamBounds, ...
                 'TauBounds', TauBounds, ...
                 'DeltaBounds', DeltaBounds, ...
                 'TOJ_LamT', params(1), ...
                 'TOJ_LamR', params(2), ...
                 'TOJ_Tau', params(3), ...
                 'TOJ_Delta', params(4), ...
                 'TOJ_epsilonTF', params(5), ...
                 'TOJ_Xi', params(6), ...
                 'TOJ_epsilonRF', params(7));
 warning(s1); warning(s2);
```

```
% -----------------Calculate Likelihood-----------------------------------
function L = LogLikelihood(x, TOJ_TF, TOJ_RF, LevTOJ, Model)
  [lam_T lam_R tau delta_TOJ xi epskap] = GetParVec(x, Model);
  penalty = 1.0e20;
  if lam_T<=0 || lam_R<=0 || delta_TOJ<0 || any([xi epskap]<0) || any([xi
epskap]>1)
    L = penalty;
  else
    P_TOJ_TF = psi_TOJ_TF(LevTOJ, lam_T, lam_R, tau, delta_TOJ, xi, epskap);
    P_TOJ_RF = 1 - P_TOJ_TF;
    L = -nansum([TOJ_TF.*log(P_TOJ_TF) TOJ_RF.*log(P_TOJ_RF)]);
    if ~isreal(L) || isinf(L), L = penalty; end
  end
% -----------------Psychometric Functions---------------------------------
function p = psi_TOJ_TF(x, lam_T, lam_R, tau, delta, xi, epskap)
  lTF = epskap(1); lRF = epskap(2);
  pTF = difcdf(-delta, x+tau, lam_T, lam_R);
  pRF = 1 - difcdf(delta, x+tau, lam_T, lam_R);
  pS = 1 - pTF - pRF;
  p = (1-lTF)*pTF + (1-xi)*pS + lRF*pRF;
% -----------------Get parameters out of vector of estimates---------------
function [lam_T lam_R tau delta_TOJ xi epskap] = GetParVec(p, Model)
lam_T = p(1);  lam_R = p(2);  tau = p(3); delta_TOJ = p(4); xi = p(5);
if Model==0, epskap = zeros(1,2);
elseif Model==1, epskap = p(6:7);
elseif Model==2, epskap = [p(6) 0];
elseif Model==3, epskap = [0 p(6)];
end
%-----------------Distribution of the arrival-time-difference------------
function d = difcdf(x,shift,lam1,lam2)
  y = x-shift;
  left = min(lam1.*exp(lam2.*y)./(lam1+lam2),1);
  right = max(1-(lam2.*exp(-lam1.*y))./(lam1+lam2),0);
  d = (y<=0).*left + (y>0).*right;
% -----------------Plot Results-------------------------------------------
function PlotIt(pars, LevTOJ, Model, TOJ_TF, TOJ_RF)
[lam_T lam_R tau delta_TOJ xi epskap] = GetParVec(pars, Model);
scrsz = get(0,'ScreenSize');
figure('Position',[scrsz(3)/10 scrsz(4)/3 4*scrsz(3)/15 scrsz(4)/4])
lw = 3; % line width
ms = 8; % marker size
fsa = 12; % font size for axis labels and title
FS = ['\fontsize{',num2str(fsa),'}'];
fsl = 11; % font size for legend
  subplot (1,1,1)
xinf = min(LevTOJ);  xsup = max(LevTOJ);
step = (xsup-xinf)/800;
xval = xinf:step:xsup; % values x-axis
P_TOJ_TF = psi_TOJ_TF(xval, lam_T, lam_R, tau, delta_TOJ, xi, epskap);
```

```
P_TOJ_RF = 1 - P_TOJ_TF;
n = TOJ_TF + TOJ_RF;
plot(xval,P_TOJ_RF,'c','LineWidth',lw); hold on
plot(LevTOJ,5./n,'k -.',LevTOJ,1./n,'k :','LineWidth',1)
plot(LevTOJ,TOJ_RF./n,'cO','MarkerFaceColor','c','MarkerSize',ms); hold off
axis([xinf xsup 0 1])
xlabel([FS ' Delay or SOA, \Delta\itt'])
ylabel([FS ' Probability of response'])
title([FS ' TOJ task'])
text(xinf+0.05*(xsup-xinf),0.9,'\Psi^{*}_{TOJ-RF}',...
     'Color','c','FontWeight','demi','FontSize',fsl)

load('perfmes.mat')
task = 'TOJ';
LamT= perfmes(1,1)
LamR= perfmes(1,2)
Tau = perfmes(1,3)
Delta = perfmes(1,4)
Xi = perfmes(1,5)
PlotRange = [-400 400];

PerformanceMeasures(task, LamT, LamR, Tau, Delta, Xi, PlotRange)

f = gcf;
exportgraphics(f,'Study4.pdf',"Append",true)
```

Appendix 5: MATLAB Scripts for Mapping Data to the Extended Twin Model

```
% Load the data from the .mat file
load('file.mat');

measured_RTs = [];
measured_TOJs = [];
SE_RT = [];
SE_TOJ = [];

param_est_all = zeros(%rows, %cols);

% Extract the desired columns for each participant in a loop
for i = 1:size(file,1)
    SE_RT(1, :) = double(table2array(file (i, %col_SE)));
    measured_RTs(1, :) = double(table2array(file(i, %col_mRTs)));
    SE_TOJ(1, :) = double(table2array(file (i, %col_SE)));
    measured_TOJs(1, :) = double(table2array(file(i, %col_mTOJ)));


    Optimise_Constrained2;
    param_est_all(i, :) = param_est;
    Check_RTs_from_parameters;
    Check_TOJs_from_parameters;
    RT_Scatterplot;
    TOJ_Scatterplot;
end

%Optimise_Constrained2
lb = [20; 20; 0; 0; 0; 0; 0];% Visual Processing Time, Audio Processing
time, Secondary Common processing time, RT Window Width, TOJ Window Width,
Temporal Gain, Bias
ub = [10000; 10000; 10000; 10000; 10000; 10000; 1];% Visual Processing Time,
Audio Processing time, Secondary Common processing time, RT Window Width, TOJ
Window Width, Temporal Gain, Bias
options = optimoptions('fmincon','Display','off','TolFun',1e-1000,'TolX',1e-
1000);
param_initial = [100;100;100;0;0;0;0.5]; % Visual Processing Time, Audio
Processing time, Secondary Common processing time, RT Window Width, TOJ
Window Width, Temporal Gain, Bias
[param_est,fval] = fmincon(@(param)minimization_func(param, measured_RTs,
measured_TOJs,SE_RT,SE_TOJ), param_initial,[], [], [], [], lb, ub, [],
options);
%[param_est,fval] = fmincon(@(param)minimization_func2(param, measured_RTs,
measured_TOJs), param_initial,[], [], [], [], lb, ub, [], options);
save ('param_est','param_est') %do not need for bootstrap
```

```
 function residual = minimization_func(parameters, measured_RTs,
measured_TOJs,SE_RT,SE_TOJ)
     estimated_RTs = estimate_RTs_from_parameters(parameters);
     estimated_TOJs = estimate_TOJs_from_parameters(parameters);
     %Z-standardisation following the orginial parameter mapping technique
     residual = sum(((measured_RTs - estimated_RTs)/SE_RT ).^2 +
((measured_TOJs - estimated_TOJs)/SE_TOJ).^2);
 end

 function estimated_RTs = estimate_RTs_from_parameters(parameters)
     % V, A, mu, omega1, omega2, delta, bias

     V = parameters(1);
     A = parameters(2);
     mu = parameters(3);
     omega1 = parameters(4);
     omega2 = parameters(5);
     delta = parameters(6);
     bias = parameters(7);
     lv = 1/V;
     la = 1/A;
     om = [omega1 omega2];
     tau = [0 50 100 150 200 250 300];
     nt = length(tau);

     %REDUNDANT TARGET (divided attention), any stimulus opens window
     %------------------------------------------------------------------------
---

     %1) VA: Visual first, Auditory second
     %------------------------------------------------------------------------
---
     %Probability of integration
     pi_VA1 = zeros(1,nt);
     pi_VA2 = zeros(1,nt);
     %first stimulus presented wins, here V
     for i = 1: nt
         if tau(i) >= om(1)
             pi_VA1(i) = (exp(-lv*tau(i))*(-1+exp(lv*om(1)))*la)/(lv+la);
         else
             pi_VA1(i) = (lv-exp(la*(-om(1)+tau(i)))*lv+la-exp(-
lv*tau(i))*la)/(lv+la);
         end
     end
     %second stimulus presented wins, here A
     for i = 1:nt
         pi_VA2(i) = (la/(la+lv))*(exp(-lv*tau(i))-exp(-lv*(om(1)+tau(i))));
% L2 wins
     end
```

```
    pi_VA = pi_VA1+pi_VA2;

    %-------------------------------------------------------------------------
---
    %Mean RT
    mrt_VA = zeros(1,nt);
    for i=1:nt
        mrt_VA(i) = 1/lv-exp(-lv*tau(i))*(1/lv-1/(lv+la))+mu-pi_VA(i)*delta;
    end

    %-------------------------------------------------------------------------
---

    %2) AV: Auditory first, Visual second
    %-------------------------------------------------------------------------
---
    %Probability of integration
    pi_AV1 = zeros(1,nt);
    pi_AV2 = zeros(1,nt);
        %first stimulus presented wins, here A
    for i= 1: nt
         if tau(i) >=om(1)
            pi_AV1(i) = (exp(-la*tau(i))*(-1+exp(la*om(1)))*lv)/(la+lv);
        else
            pi_AV1(i) = (la-exp(lv*(-om(1)+tau(i)))*la+lv-exp(-
la*tau(i))*lv)/(la+lv);
        end
    end

    %second stimulus presented wins, here V
    for i = 1:nt
        pi_AV2(i) = (lv/(lv+la))*(exp(-la*tau(i))-exp(-la*(om(1)+tau(i))));
    end

    pi_AV = pi_AV1+pi_AV2;

    %-------------------------------------------------------------------------
---
    %Mean RT
    mrt_AV = zeros(1,nt);
    for i=1:nt
        mrt_AV(i) = 1/la-exp(-la*tau(i))*(1/la-1/(la+lv))+mu-pi_AV(i)*delta;
    end

    % mean of VA and AV
    estimated_RTs = [fliplr(mrt_VA) mrt_AV(2:end)];
 end
```

```
function estimated_TOJs = estimate_TOJs_from_parameters(parameters)
    % V, A, mu, omega1, omega2, delta, bias

    V = parameters(1);
    A = parameters(2);
    mu = parameters(3);
    omega1 = parameters(4);
    omega2 = parameters(5);
    delta = parameters(6);
    beta = parameters(7);
    lv = 1/V;
    la = 1/A;
    om = [omega1 omega2];
    tau = [50 100 150 200 250 300];
    tau_n=-fliplr(tau');
    nt = length(tau);

    % Probs TOJ    : V before A
for i= 1: nt
     if tau(i) >=om(2)
pi_TOJ_VA(i) =  1-(la/(la+lv))*(exp(-lv*(om(2)+tau(i)))*(beta-(beta-
1)*exp(2*lv*om(2))));
      else
          pi_TOJ_VA(i) = (1/(la+lv))*(exp(la*(-om(2)+tau(i)))*lv+beta*(la*(1-
exp(-lv*(om(2)+tau(i))))+lv*(1-exp(la*(-om(2)+tau(i))))));
      end
end
%pi_VA_TOJ=pi_TOJ_VA1+pi_TOJ_VA2;

%-------------------------------------------------------------------------
% Probs TOJ    : A before V


for i=1:nt
    if tau_n(i) +om(2) <= 0;
pi_TOJ_AV(i)=(lv/(lv+la))*(exp(la*(tau_n(i)-om(2)))*(1+beta*(-
1+exp(2*la*(om(2))))));
    else
    pi_TOJ_AV(i)=(1/(la+lv))*(exp(la*(tau_n(i)-om(2)))*lv+beta*(la*(1-exp(-
lv*(om(2)+tau_n(i))))+lv*(1-exp(la*(-om(2)+tau_n(i))))));
    end
end

estimated_TOJs= [0 fliplr(pi_TOJ_AV(2:6)) pi_TOJ_VA 1];

function Check_RTs = Check_RTs_from_parameters(param_est)
    % V, A, mu, omega1, omega2, delta, bias
load("param_est.mat")
```

```
    V = param_est(1);
    A = param_est(2);
    mu = param_est(3);
    omega1 = param_est(4);
    omega2 = param_est(5);
    delta = param_est(6);
    bias = param_est(7);
    lv = 1/V;
    la = 1/A;
    om = [omega1 omega2];
    tau = [0 50 100 150 200 250 300];
    nt = length(tau);

    %REDUNDANT TARGET (divided attention), any stimulus opens window
    %-------------------------------------------------------------------------
---

    %1) VA: Visual first, Auditory second
    %-------------------------------------------------------------------------
---
    %Probability of integration
    pi_VA1 = zeros(1,nt);
    pi_VA2 = zeros(1,nt);
    %first stimulus presented wins, here V
    for i = 1: nt
        if tau(i) >= om(1)
            pi_VA1(i) = (exp(-lv*tau(i))*(-1+exp(lv*om(1)))*la)/(lv+la);
        else
            pi_VA1(i) = (lv-exp(la*(-om(1)+tau(i)))*lv+la-exp(-
lv*tau(i))*la)/(lv+la);
        end
    end
    %second stimulus presented wins, here A
    for i = 1:nt
        pi_VA2(i) = (la/(la+lv))*(exp(-lv*tau(i))-exp(-lv*(om(1)+tau(i))));
% L2 wins
    end

    pi_VA = pi_VA1+pi_VA2;

    %-------------------------------------------------------------------------
---
    %Mean RT
    mrt_VA = zeros(1,nt);
    for i=1:nt
        mrt_VA(i) = 1/lv-exp(-lv*tau(i))*(1/lv-1/(lv+la))+mu-pi_VA(i)*delta;
    end
```

```
    %-------------------------------------------------------------------------
---

    %2) AV: Auditory first, Visual second
    %-------------------------------------------------------------------------
---
    %Probability of integration
    pi_AV1 = zeros(1,nt);
    pi_AV2 = zeros(1,nt);
        %first stimulus presented wins, here A
    for i= 1: nt
         if tau(i) >=om(1)
            pi_AV1(i) = (exp(-la*tau(i))*(-1+exp(la*om(1)))*lv)/(la+lv);
        else
            pi_AV1(i) = (la-exp(lv*(-om(1)+tau(i)))*la+lv-exp(-
la*tau(i))*lv)/(la+lv);
        end
    end

    %second stimulus presented wins, here V
    for i = 1:nt
        pi_AV2(i) = (lv/(lv+la))*(exp(-la*tau(i))-exp(-la*(om(1)+tau(i))));
    end

    pi_AV = pi_AV1+pi_AV2;

    %-------------------------------------------------------------------------
---
    %Mean RT
    mrt_AV = zeros(1,nt);
    for i=1:nt
        mrt_AV(i) = 1/la-exp(-la*tau(i))*(1/la-1/(la+lv))+mu-pi_AV(i)*delta;
    end

    % mean of VA and AV
    estimated_RTs = [fliplr(mrt_VA) mrt_AV(2:end)]

    save('estimated_RTs', 'estimated_RTs');
 end
 function Check_TOJs = Check_TOJs_from_parameters(param_est)
    % V, A, mu, omega1, omega2, delta, bias

 load("param_est.mat")

    V = param_est(1);
    A = param_est(2);
    mu = param_est(3);
    omega1 = param_est(4);
    omega2 = param_est(5);
```

```
    delta = param_est(6);
    beta = param_est(7);
    lv = 1/V;
    la = 1/A;
    om = [omega1 omega2];
    tau = [0 50 100 150 200 250];
    tau_n=-fliplr(tau');
    nt = length(tau);

    % Probs TOJ    : V before A
 for i= 1: nt
     if tau(i) >=om(2)
 pi_TOJ_VA(i) =  1-(la/(la+lv))*(exp(-lv*(om(2)+tau(i)))*(beta-(beta-
1)*exp(2*lv*om(2))));
     else
         pi_TOJ_VA(i) = (1/(la+lv))*(exp(la*(-om(2)+tau(i)))*lv+beta*(la*(1-
exp(-lv*(om(2)+tau(i))))+lv*(1-exp(la*(-om(2)+tau(i))))));
     end
 end
 %pi_VA_TOJ=pi_TOJ_VA1+pi_TOJ_VA2;

 %-------------------------------------------------------------------------
 % Probs TOJ    : A before V


 for i=1:nt
    if tau_n(i) +om(2) <= 0;
 pi_TOJ_AV(i)=(lv/(lv+la))*(exp(la*(tau_n(i)-om(2)))*(1+beta*(-
1+exp(2*la*(om(2))))));
    else
    pi_TOJ_AV(i)=(1/(la+lv))*(exp(la*(tau_n(i)-om(2)))*lv+beta*(la*(1-exp(-
lv*(om(2)+tau_n(i))))+lv*(1-exp(la*(-om(2)+tau_n(i))))));
    end
 end

 estimated_TOJs= [0 fliplr(pi_TOJ_AV(2:6)) pi_TOJ_VA 1]

 save('estimated_TOJs', 'estimated_TOJs');

 %plotRT
 % Define the measured reaction times and SOA values
 file_name = sprintf('TWINRT_participant_%d.png', i);
 file_path = fullfile('C:\Users\c.glenn\OneDrive - York St John
University\Outputs', file_name);

 load ("estimated_RTs.mat")
 SOAs = [-300,-250,-200,-150,-100,-50,0,50,100,150,200,250,300];

 % Create a scatter plot of the data
```

```
 scatter(SOAs, measured_RTs);
 hold on;
 scatter(SOAs, RT_estimated_RTs);

 % Add labels and a title
 xlabel('SOA (ms)');
 ylabel('Reaction Time (s)');
 title('Measured vs. Estimated Reaction Time vs. SOA');

 % Add a legend
 legend('Measured', 'Estimated');

 % Add lines connecting the dots
 plot(SOAs, measured_RTs, '-o');
 plot(SOAs, estimated_RTs, '-o');

 hold off
 % Save the scatter plot as a PNG file
 saveas(gcf, file_path);

 %plotTOJ
 % Define the measured reaction times and SOA values
 file_name = sprintf('TWINTOJ_participant_%d.png', i);
 file_path = fullfile('C:\Users\c.glenn\OneDrive - York St John
University\Outputs', file_name);

 % Define the measured TOJ and SOA values
 load ("estimated_TOJs.mat")
 SOAs = [-300,-250,-200,-150,-100,-50,0,50,100,150,200,250,300];

 % Create a scatter plot of the data
 scatter(SOAs, measured_TOJs);
 hold on;
 scatter(SOAs, estimated_TOJs);

 % Add labels and a title
 xlabel('SOA (ms)');
 ylabel('Frequency of Visual First Judgements (s)');
 title('Measured vs. Estimated TOJs vs. SOA');

 % Add a legend
 legend('Measured', 'Estimated');

 % Add lines connecting the dots
 plot(SOAs, measured_TOJs, '-o');
 plot(SOAs, estimated_TOJs, '-o');

 hold off
 saveas(gcf, file_path);
```

Appendix 6: Moderation of Presence Subscales by RT TBW Across Conditions in Study 4

Table 18
Moderation of Presence Subscales by RT TBW Across Conditions in Study 4

MEC-SPQ Subscale	Condition	*F(df= 1,96)*	*p (moderation)*	*Partial η²*	*B*	*95% CI*	*t*	*p (conditional)*
Attention Allocation	Moderation	4.459	0.037	0.044				
	Stationary				-0.009	[-.018, .000]	-2.035	0.045
	Natural Locomotion				-0.001	[-.006, .004]	-0.242	0.809
Spatial Situational Model	Moderation	5.92	0.017	0.058				
	Stationary				0.147	[.038, .257]	2.674	0.009
	Natural Locomotion				0.047	[-.037, .130]	1.114	0.268
Self-Location	Moderation	9.732	0.002	0.092				
	Stationary				-0.011	[-.021, -.001]	-2.138	0.035
	Natural Locomotion				0.004	[-.002, .010]	1.237	0.219
Perceived Possible Actions	Moderation	6.762	0.011	0.066				
	Stationary				-0.008	[-.017, .001]	-1.725	0.088
	Natural Locomotion				0	[-.005, .006]	0.067	0.946
Higher Cognitive Involvement	Moderation	3.385	0.069	0.034				
	Stationary				-0.012	[-.020, -.003]	-2.642	0.01
	Natural Locomotion				0.003	[-.004, .010]	0.856	0.394
Suspension of Disbelief	Moderation	3.936	0.05	0.039				
	Stationary				0.003	[-.005, .010]	0.748	0.456
	Natural Locomotion				-0.007	[-.015, .001]	-1.671	0.098
Domain Specific Interest	Moderation	5.974	0.016	0.059				
	Stationary				-0.004	[-.014, .006]	-0.753	0.453
	Natural Locomotion				0.002	[-.008, .013]	0.436	0.664
Visual Spatial Imagery	Moderation	1.949	0.166	0.02				
	Stationary				-0.005	[-.015, .005]	-1.021	0.31
	Natural Locomotion				-0.001	[-.011, .009]	-0.248	0.804

Table 18
Moderation of SSQ Subscales by RT TBW across Conditions in Study 4 (Continued)

SSQ Subscale	Condition	*F(df= 1,96)*	*p (moderation)*	*Partial η²*	*B*	*95% CI*	*t*	*p (conditional)*
Nausea	Moderation	7.16	0.009	0.069				
	Stationary				0.02	[.005, .035]	2.65	0.009
	Natural Locomotion				-0.005	[-.014, .004]	-1.083	0.282
Oculomotor	Moderation	6.138	0.015	0.06				
	Stationary				0.023	[.006, .040]	2.721	0.008
	Natural Locomotion				-0.002	[-.010, .006]	-0.568	0.571
Disorientation	Moderation	5.392	0.023	0.053				
	Stationary				0.03	[.010, .050]	2.995	0.004
	Natural Locomotion				-0.005	[-.016, .006]	-0.931	0.354