The Effectiveness of Training in Virtual Environments

María Murcia López

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
of
University College London.

Department of Computer Science
University College London

December 3, 2018

I, María Murcia López, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Abstract

The research presented in this thesis explores the use of consumer virtual reality technology for training, comparing its validity to more traditional training formats. The need to evaluate the effectiveness of training in virtual environments is critical as a wider audience gains access to an array of emerging virtual reality consumer devices. Training is an obvious use case for these devices. This is motivated by the well-known success of domain-specific training simulators, the ability to train in safe, controlled environments and the potential to launch training programs when the physical components required to complete a task are not readily available.

In this thesis, we present four user studies that aim to compare the effectiveness of systems with varying levels of immersion for learning transfer of several tasks, ranging from object location spatial memory to more complex assembly procedures. For every study, evaluation of the effectiveness of training took place in a real-world, physical environment. The first two studies compare geometric and self-motion models in describing human spatial memory through scale distortions of real and virtual environments. The third study examines the effect of level of immersion, self-avatar and environmental fidelity on object location memory in real and virtual environments. The fourth study compares the effectiveness of physical training and virtual training for teaching a bimanual assembly task.

Results highlight the validity of virtual environments for training. The overall conclusion is that virtual training can yield a resulting performance that is superior to other, more traditional training formats. Combined, the outcomes of each of the user studies motivate further study of consumer virtual reality systems in training and suggest considerations for the design of such virtual environments.

Acknowledgements

This research was funded through an EPSRC Industrial Cooperative Award in Science and Technology (Reference EP/L505717/1), with the High Speed Sustainable Manufacturing Institute (HSSMI) as the partner organisation.

First and foremost, I would like to thank Prof. Anthony Steed for absolutely everything. Words cannot capture how fortunate one is to be mentored by such an inspiring, kind and knowledgeable supervisor. I wish to thank my second supervisor, Dr. Simon Julier, for his invaluable advice and project collaborations. To Dr. William Steptoe, for his guidance over the past years. To Prof. Niloy Mitra, for his exceptional feedback during my vivas.

I would like to thank my fellow research group members Alex Bokaris, Ben Congdon, Sebastian Friston, Andrew MacQuarrie, Ye Pan, Jacob Thorn, David Swapp, David Walton, our extended VEIV group and our colleagues from 1ES. To Denis Timm, Dave Twisleton and Nick Turpin, for assisting with the craziest setups. To JJ Giwa, Jeanie Doel, Patricia Fenoy and Lesa Bastien, for taking care of us all. To Dawn Bailey and Wendy Richards, for their support. To Caroline Wardle, for helping with our study designs. I also wish to thank Fiona E. Zisch, Anisa Motala, Jake Greaves, Dr. Hugo J. Spiers and our collaborators, for sharing with me the most extraordinary academic experience. To my HSSMI and Oculus colleagues, for joining along the journey. And to the regulars at my local pub, for all the laughs and British lingo!

To Hanna and Rick, for making London feel like home. To my family, whose love knows no limits. And, most of all, I am eternally thankful to my parents, Victoria and Javier, for always being with me despite any distance.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Research Problem

The availability of affordable consumer Virtual Reality (VR) technology has raised the manufacturing industry's interest in Virtual Environments (VEs) for assembly line operation training. The possibility of training programs to be initiated before physical workstations, parts and tools are available sounds enticing as it could enhance the end-to-end manufacturing process. Moreover, these systems offer unique capabilities related to trainee safety and performance metric gathering as well as the ability to parallelise training by removing dependencies on the aforementioned physical components.

Accordingly, this thesis is concerned with the effectiveness of VEs in training. We focus specificially on level of immersion, understood as the objective fidelity of sensorial stimuli offered by a VR system, as the main parameter that mediates the transfer of knowledge from VEs to real-world scenarios. In other words, in this thesis we explore the ability to use the knowledge that has been acquired in a specific context (the virtual space) in a new or different one (the physical space).

We designed and ran a series of user studies with the goal of comparing the effectiveness of training in systems with different levels of immersion for a series of tasks. These tasks ranged from basic object location memory to more complex, bimanual procedural tasks. Common across all studies is that performance of training was always measured in the real world. Thus, all participants were tested in solving the user study task in a physical environment with physical components. This experimental design choice was informed by the fact that previous studies on the effectiveness of VEs on training have commonly evaluated the effectiveness of the training transfer in computer-generated environments.

1.2 Scope

The work presented in this thesis focuses on level of immersion as the main parameter that mediates the transfer of knowledge from virtual to real environments. For each of the studies presented in this thesis we defined levels of immersion and included physical training with physical components in the real world as a baseline training system against which to compare the validity of training in VEs. We also defined performance metrics relevant to each of the user study tasks such as distance errors for object location memory tasks and assembly time for procedural tasks, amongst others.

Although relevant to the research topic, haptics, locomotion techniques and spatialised audio were not investigated. In addition, studies related to this project did not examine display systems other than consumer Head-Mounted Displays (HMDs) - namely the HTC Vive, the Oculus Rift Development Kit 2 and the Oculus Rift Consumer Version 1 - and desktop computers. In our performance analysis we did not evaluate or compare software and hardware used to build the training VE systems. Additionally, the cost of generating training VEs as well as the deployment and adoption of this training format by industry was not within the scope of this project.

1.3 Contributions

The main contribution in this thesis is the evaluation of the effectiveness of VEs in training for a range of tasks through user studies with participants. The experimental design and method for each of these studies is introduced in the corresponding chapters for replicability. Results encourage further exploration of consumer VR systems in training and highlight their superiority over desktop computer training. We discuss the limitations and include recommendations.

1.3.1 Methodological Contributions

1. An experimental protocol for exploring the effect of spatial distortions on object location memory in physical environments (Chapter 3).

2. An experimental protocol for exploring the effect of spatial distortions on object location memory in VEs (Chapter 4).

3. An experimental protocol for exploring the effect of level of immersion, environmental fidelity and self-avatar on object location memory in VEs (Chapter 5).

4. An experimental protocol for comparing virtual and physical training transfer of bimanual assembly tasks, extending on previous work [2] (Chapter 6).

1.3.2 Substantive Contributions

1. Research findings that highlight the differences and similarities between object location recall in virtual and physical environments after boundary distortions (Chapter 3 and Chapter 4).

2. Research findings that explore the effect of level of immersion, environmental fidelity and self-avatar on object location training in VEs (Chapter 5).

3. Research findings that support the validity of VEs for bimanual assembly tasks and their superiority over other, more traditional training formats (Chapter 6).

1.3.3 Analysis Contributions

1. Two proposed selection criteria for determining how object location models best describe individual participant responses in user studies that explore the effect of spatial transformations of the boundaries of an environment (Chapters 3 and 4).

1.4 Collaborations

The project is based on a series of user studies, two of them being sections of a large collaboration with other University College London (UCL) departments and supported through a James S. McDonnell Foundation Scholar award to Dr. Hugo J. Spiers and a UCL Grand Challenges Small Grant. These user studies are reported in Chapter 3 and Chapter 4. The main project collaborators, afiliations and contributions are described in Table 1.1. Other contributors to these studies include:

- Dominic Zisch, Charles Middleton, Aaron Breuer-Weil, Rowan Haslam, Ludovico Saint Amour di Chanaz, William De Cothi and Thomas Reed, who helped running and piloting the study.
- Derrick Boampong, Tatsuto Suzuki, Nikos Papadosifos and Biao Yang, who provided technical support at the UCL Pedestrian Accessibility Movement Environment Laboratory (PAMELA) facility.
- Simon Julier, who offered technical advice on Three Dimensional (3D) tracking for data collection.

The results presented in Chapter 3 and Chapter 4 are based on our own analysis, performed independently from the rest of collaborators. We also contributed to the design of the research protocol as well as decided on some of the hypotheses.

1.5 Structure

The rest of this thesis is organised as follows. Chapter 2 covers background literature related to the research topic. This chapter introduces relevant research on the parameters of VR systems that mediate learning transfer to the real world as well as on human spatial cognition.

Chapter 3 reports the experimental design, method, results and discussion of a collaborative study. This chapter discusses the plausibility of running a study where participants learn and recall object locations following alterations to the boundaries of a real-world environment.

Table 1.1: Project collaborators, UCL affiliations and contributions (in alphabetical order by first surname).

Collaborator	UCL Affiliation	Contributions
Greaves, Jacob	Brain Sciences	Study design Study execution Data analysis
Motala, Anisa	Biosciences	Study design Study execution Data analysis
Murcia López, María	Computer Science	Study design Study execution Data analysis Technical expertise (3D tracking)
Spiers, Hugo J.	Institute of Behavioural Neuroscience	Supervision Study design
Steed, Anthony	Computer Science	Supervision Study design Technical expertise (3D tracking)
Tyler, Nick	Civil, Environmental and Geomatic Engineering	Supervision Facility support
Zisch, Fiona E.	Institute of Behavioural Neuroscience The Bartlett School of Architecture	Study design Technical expertise (architecture)

Chapter 4 reports the experimental design, method, results and discussion of a study exploring the effect of scale transformations of a VE on spatial memory through object location memory and recall.

Chapter 5 reports the experimental design, method, results and discussion of a study exploring the effect of level of immersion, varied feature fidelity and self-avatar on object location memory.

Chapter 6 reports the experimental design, method, results and discussion of a study comparing real and virtual training of a multi-step bimanual assembly task.

Chapter 7 contains conclusions as well as directions for future work.

Chapter 2

Background

2.1 Motivation

This project has been largely motivated by the well-known success of domain-specific training simulators in medical, military, navigation and pilot training, amongst other fields. Consumer VR systems offer the possibility to train in safe, controlled environments and the potential to launch training programs when the physical components required to complete a task are not readily available. Moreover, these systems are becoming more accessible to a wider audience through a range of affordable consumer devices. VEs could deliver cost-efficient, safe, controlled and potentially effective training. If proven adequate, virtual training would also allow for the completion of operator instruction prior to the installation of physical workstations, tools and components, with optional built-in automatic capture of data relating to system and user performance. This would accelerate the end-to-end manufacturing process and, consequently, increase efficiency of production. However, evidence is needed to ascertain the effectiveness of consumer VR devices for training as opposed to more traditional training formats.

We aim to continue to address the common interest in the fields of neuroscience, experimental psychology and VR for better understanding the way humans perceive, navigate, interpret and recall 3D space [3, 4, 5]. Essential to the survival of motile living species, navigation of environments and recall of specific locations within them highly rely on spatial memory. This is the component of memory responsible for capturing, storing and utilising information about one's external surroundings and spatial orientation. The ability to understand and ultimately predict the behaviour of humans in environments could have a number of benefits: enhanced user experience and layout optimisation for VR training applications, enhanced design of the built environment for urban lifestyles [6, 7] and advanced diagnostic tools for Alzheimer's disease [8], amongst others.

By advancing our knowledge on how humans acquire, store and use spatial representations, we could better inform the design of VEs for spatial training [9]. Various aspects of the user experience design of VR applications, including spatial layout and features of the VEs and objects within it, as well as the user's starting location and facing direction, could be rooted in more complete models that optimise the learning transfer of spatial information.

2.2 Virtual Environments

2.2.1 Level of Immersion

The term *immersion* can be understood as the objective fidelity of sensorial stimuli offered by a VR system. Slater et al. have defined it as "a description of a technology" [10]. Slater suggests to use the term to refer to "what the technology delivers from an objective point of view" [11]. In his later work, he argues that "we describe immersion not by the displays plus tracking, but as a property of the valid actions that are possible within the system" and that "the level of immersion is completely determined by the physical properties of the system" [12]. Under this definition, he claims that "system A is at a higher level of immersion than system B if the valid actions of B form a proper subset of those of system A". Immersion, therefore, can be used to define systems in relation to other systems. Ragan et al. recommend to speak of levels of immersion rather than terms such as nonimmersive and immersive VR [13].

Bowman and McMahan describe immersion as an objective and measurable multidimensional array formed by many components including Field of View (FOV), Field of Regard (FOR), display size, display resolution, stereoscopy, head-based rendering, realism of lighting, frame rate and refresh rate [14]. In the context of this research, we use the concept of *level of immersion* to refer to the different widely available consumer displays and navigation techniques used by the participants to explore the VEs designed for our user studies.

While many applications of VR in training have used desktop environments, more immersive VR systems are now becoming widely available. Highly immersive VR technology potentially increases experimental and environment realism, gives researchers the ability to perform manipulations to an environment, and provides new data sources, such as body tracking, amongst other benefits [15, 16]. When being presented with a stereoscopic view and given access to self-motion cues, participants can respond realistically to situations and events [12, 17].

Previous research has shown that display and interaction fidelity have a strong effect on strategy and performance in a VR first-person shooter game [18]. As technology moves towards augmentation of real world learning by the use of virtual tools, performance in systems with different levels of immersion must be analysed and compared with real world learning.

In this thesis we consider real world learning the highest level of immersion, followed by HMD learning and then desktop learning. We also consider the navigation technique associated with each learning system as an inherent and crucial element of level of immersion. All training systems as well as the corresponding navigation techniques are further detailed in each of the chapters reporting the experiments relating to this thesis. Across all studies we expected the level of immersion to have an effect on training transfer [19, 20].

2.2.2 Environmental Fidelity

When training in a VE it is important to have an understanding of the technological variables that can be sacrificed without degrading learning effectiveness transfer to the real world [16, 20]. One of these variables is *environmental fidelity* which can be understood as the fidelity of mapping from a real-world space to a computer-generated virtual replica. A distinction can be made between two broad types of environmental cues: *geometric*, cues provided by environmental surfaces such as walls, and *featural*, nongeometric cues provided by the environment, such as colour [21, 22, 23]. Previous research has demonstrated an inclination for spatial localisation to be based mainly on geometric properties of an environment, rather than featural cues [22, 24].

Although geometric fidelity of a space can be reproduced using basic 3D objects, such as planes, spheres or cubes, high feature fidelity is not always achievable or may result in the development of computationally expensive systems. Previous studies have assessed the impact of rendering style on distance perception accuracy in virtual replicas of concurrently occupied VEs [25, 26]. These studies suggest that there are no indications of perceived compressed distances in immersive VEs where participants can be certain of them being faithful representations of their occupied space.

Slater et al. explored the effect of visual realism on sense of presence in immersive VEs [12]. Participants were exposed to a VE rendered in two levels of visual realism. They found that subjective presence was higher for the version of the VE with higher visual realism. However, Masahiro Mori's 'Uncanny Valley' hypothesis [27] remains unanswered, since it is not clear whether higher environmental fidelity might result in training enhancement up to a point after which there might be a decrease in performance due to defect magnification.

Based on previous results, in Chapter 5 we directly compare performance resulting from learning object locations in concurrently occupied virtual and real environments. We explore learning and recall of multiple external object locations as subjective measures of spatial perception. We focus on understanding which cues are necessary for the design of virtual spaces that will ensure the optimal transfer of spatial knowledge to the real world.

2.2.3 Self-Avatar

Slater and Usoh have suggested that the sense of presence in VEs, or the subjective feeling of *being there*, can be enhanced by providing users with a virtual self-avatar [28]. Results from several studies have suggested that a self-avatar is also beneficial to performance on interaction tasks in VEs. McManus et al. found that participants with with a self-avatar or who saw a character animation performed behavioural tasks faster and more accurately [29]. Other research has demonstrated that a self-avatar may alleviate participant's cognitive workload in a VE [30]. In their study, Steed et al. found that participants who had an avatar and were allowed to move their hands had significantly higher recall in a task where they had to memorise pairs of letters, perform a spatial rotation exercise, and recall the pairs of letters. They also found that participants who were allowed to move their hands, but could not see their self-avatar, would often not move their hands or stop moving them after a short while. Subjective feedback from participants has also highlighted the utility of a full virtual body as a reference point for spatial tasks [31].

Related work has found that fully tracked, high fidelity virtual avatars can improve distance estimation accuracy in non-photorealistic Ves [32, 33]. Similarly, self-embodiment in highly realistic VEs has been reported to increase accuracy in distance judgements [34]. These results suggest that high fidelity avatars can facilitate enhanced spatial task performance in a VE without compromising the ability for effective information transfer to the real world. A recent study also tested the effect of avatar fidelity on the accuracy of distance estimations in the near-field, comparing with real-world performance [35]. Results showed that estimations were more accurate as visual fidelity of the avatar increased, with accuracy of high fidelity avatars approaching real-world performance.

However, spatial perception enhancement seems to be compromised when using low geometry avatar representation or single point tracking [36]. Other results from studies on egocentric distance estimation indicate that simplified avatar implementations (single-point rather than full body tracking and low fidelity based on rendering small spheres at raw tracking marker locations rather than high fidelity using a textured triangle mesh) are significantly less effective [37]. Moreover, in this study participants who were given simplified avatar representations performed only marginally more accurately than the participants who were given no avatar. Similar results were observed in a study where participants that saw a fully-articulated and tracked representation of themselves made more accurate judgments of absolute egocentric distance to locations than participants who saw no avatar [38]. Nonanimated avatars also improved distance judgments but to a lesser degree. Another study investigated the degree to which self-avatar movement must reflect the actual movements of the participant for accurate distance estimation [38]. Results indicated that experience with an animated avatar, even if the movements of the avatar did not correspond with the participant's body movement, favoured more effective distance estimation.

The study presented in Chapter 5 explores the use of single point tracking virtual avatars based on head tracking with no animations in an object location memory task. The aim is to explore if a low tracking fidelity virtual avatar can enhance performance in an object memory location task, where there is no interaction with the environment and the virtual objects in it, other than unguided, exploratory navigation. We report how results from this study could inform the design of future training systems in which robust avatar motion fidelity involving full body tracking or high fidelity avatars may not be available.

2.2.4 Virtual Environments as Proxies

Our work relates to the overarching theme of visual fidelity in VR training: to what extent does a VE have to look *real* so that the learning and recall of information presented in it is optimal [20, 16, 39]. Findings from relevant spatial cognition studies have highlighted which geometric and featural cues play the most important role and to what degree they are necessary in the training of spatial information [22, 21, 23]. We are also interested in scenarios in which the training VE is potentially different to the environment where the acquired skills are going to be used (the work enviornment). This could happen in cases in which either the layout of the work environment is unknown or it is difficult to replicate in VR.

Similarly, research projects that investigate human spatial cognition can require complicated setups or simulations, difficult or impossible to construct as a physical space [40, 5]. In the study presented in Chapter 3, the specific premise was to build a featureless physical large-scale room (approximately 5 x 5m in surface area) which could change its size and shape. These transformations had to be achieved in a very short period of time to keep experimental task timings within reasonable limits. Solving this involved a very elaborate series of design decisions involving materials, technology, architecture and labour force under restricted research budget. In Chapter 4, we present a study in which we use a virtual replica of this physical setup with a sufficient degree of visual fidelity, where its shape and size can be effortlessly modified. This replica VE can be used to test initial hypotheses and develop models before the laborious and time-consuming task of performing real-world studies and analysis of the real-world data.

In the field of psychology, VEs have been used in lieu of real world environments for several decades [15, 41, 42, 4, 43]. Loomis et al. analyse the benefits and drawbacks of immersive VE technology, as compared with traditional experimental research methods in psychology [15]. On one hand, experimental realism is increased and researchers are provided with the ability to perform alterations that would be impossible, or highly complicated, by other means. New data sources, such as body tracking, can be acquired, providing deeper analysis of body behaviour and navigation. On the other hand, they point out the high complexity of hardware and software as factors that can increase the likelihood of artefacts contaminating results and after-effects such as motion sickness, as disadvantages of the use of this technology in research.

2.2.5 Training in Virtual Environments

Previous research has highlighted the effectiveness of immersive mixed reality training in different disciplines, including military training, medical training and vehicle driving simulators [44, 45], as well as navigation and spatial knowledge training [20, 46], amongst others. Despite the recognised success in the aforementioned fields, studies on immersive virtual training transfer have reported contrasting results.

Several studies have shown that spatial information of the kind required for navigation transfers effectively from virtual to real situations, confirming the potential and benefits of VR technology in spatial training [47]. In particular, this work studied how information about the spatial layouts of virtual buildings acquired from the exploration of 3D computer simulations transfered to their real-world counterpart. Similarly, results to date from several studies have further shown that following several virtual tours of a building, disabled children acquired a considerable degree of spatial competence [48, 49, 50, 51]. Another study examined transfer of spatial learning from a VE to a real-world equivalent environment using a simulation of a shopping centre with elderly participants. Results confirmed the potential of training in VEs for the elderly [52].

Hall and Horwitz compared retention of procedural knowledge of equipment operation in an immersive VE and in a 2D computer environment and found no significant differences [53]. They claimed that VR training may not be superior to conventional electronic media for training certain skills. Gavish et al. evaluated the use of VR and augmented reality technology for industrial maintenance and assembly task training [54]. They concluded that an augmented reality platform was more suitable for training of this type of tasks and encouraged further evaluation of VR-based training.

In a more recent study Gonzalez-Franco et al. compared collaborative conventional face-to-face training with a mixed reality training setup for a manufacturing procedure of an aircraft door [55]. Their results indicated that performance levels yielded by the immersive mixed reality training system were not significantly different from the conventional face-to-face training format. Rose et al. evaluated the transfer from a VE to the real world of a simple sensorimotor task [56]. Overall, virtual training resulted in equivalent or even better real world performance than real or physical training of the task. However, they advise that their findings may not apply to other types of training tasks.

Sowndararajan et al. found an effect of level of immersion in memorising a complex procedure [57]. In their study, participants trained in the system with the higher level of immersion (a large L-shaped projection display) completed tasks significantly faster and with fewer errors than participants trained in the system with lower level of immersion (using a typical laptop display).

Other studies have shown effective learning transfer in VEs with the addition of haptic force-feedback devices. For instance, Adams et al. conducted a study to explore the benefits of haptic feedback for virtual training of a manual task [58]. They reported that force-feedback was a requirement for higher learning transfer in VEs.

Our work preseented in Chapter 6 is inspired by the work of Carlson et al. in 2015 [2], itself motivated by previous work [59, 60, 61]. In a between-subjects experimental design, Carlson et al. compared the effectiveness of virtual bimanual haptic training versus traditional physical training of an assembly task consisting of a six-piece burr puzzle. Their results indicated that physically trained participants initially outperformed virtually trained participants. However, virtually trained participants improved their testing times after two weeks. Results also showed that virtual training was enhanced by using coloured blocks as they helped participants remember the assembly process. We ran a similar task comparing paper- and video-based training with virtual training in the absence of a haptic force-feedback device [62].

We agree with Carlson et al. in that 3D burr puzzles are suitable proxy tasks or abstractions of context-specific manual assembly tasks, such as engine assembly operations at vehicle manufacturing plants. We therefore decided to use the same type of task in our study. Following their reported methods, we complemented the training task with a series of mental rotation tests to distribute participants amongst the condition groups in our between-subjects experimental design [63, 64, 65]. We also decided to colour-code the puzzle blocks and instructions as well as to use a semi-transparent virtual representation of the hands in the VE [66, 67], amongst other recommendations made by the authors which are further explained in Section 6.1.

Our work presented in Chapter 6 extends and builds on previous work by comparing a number of virtual and physical training formats, the latter representing the most common formats (video and paper instructions) in current assembly process training programmes. The main aim of this research is to verify whether exposure to a virtual training environment is sufficient for effective training. We are specifically interested in situations in which haptic devices are not available and when the physical components and tools used in the process are not accessible during training.

2.3 Human Spatial Cognition

2.3.1 Neural Correlates of Spatial Representations

Since the discovery of place cells in the mammalian hippocampal formation of rodents in 1971 [68], the fields of cognitive and behavioural neuroscience interested in space have made substantial advances in understanding how the human brain builds a neural representation of the external environment. Place cells activate when an animal moves through a unique location in an environment, referred to as the place field. Through the overlap of place fields, the activity of multiple place cells can determine the location of the animal in the environment, as well as potentially store other locations within the space [69]. Also identified in bats [70, 71], researchers have attempted to model the firing and network interactions of place cells and other spatially tuned neurons to explain spatial representations in the human brain.

Termed the cognitive map, this internal representation forms the foundation for navigational abilities, as well as enabling feelings of being embodied and embedded in the world [72]. In order to successfully navigate the world, it is crucial for the brain to not only construct an internal representation of spatial geometry and features, but to also represent one's own position, orientation, and movement and likewise objects within the environment.

In 1973 Kaplan suggested the following definition of the cognitive map:

> "The cognitive map is a construct that has been proposed to explain how individuals know their environment. It assumes that people store information about their environment in a simplified form and in relation to other information they already have. It further assumes that this information is coded in a structure which people carry around in their

> heads, and that this structure corresponds, at least to a reasonable degree, to the environment it represents. It is as if an individual carried a map or model of the environment in his head. The map is far from a cartographer's map, however. It is schematic, sketchy, incomplete, distorted, and otherwise simplified and idiosyncratic. It is, after all, a product of experience, not of precise measurement." (pp. 275-6) [73].

The cognitive map is constructed by integrating egocentric and allocentric reference frames [74]. In an egocentric frame of reference, spatial geometry and the location of objects are represented in relation and relative to the self (the head or a limb). Egocentric frameworks exist as a series of visual *snapshots* of the environment [75]. They allow animals to estimate directions and distances to external cues in relation to their own body, allowing them to guide themselves to, reach - or avoid - objects [76]. In an allocentric frame of reference geometric cues are represented in relation to each other in a viewport-independent, quasi absolute manner. This allows for novel route calculation and topography estimation, including salient environmental cues and their spatial relationships, as well as the animal's location within the environment [68]. There is evidence that both egocentric and allocentric representations are necessary for successful navigation of environments [74].

A range of theories and models have been proposed to explain which and how neural processes might parse the environment to allow navigation and action in space. As a basis, these models share an understanding of the cells that are involved in the systematic construction of a map. Place cells in the hippocampus represent locations along the path travelled by laying down place fields in each respectively traversed location [77]. It is thought that, together with medial entorhinal grid cells [78], boundary vector cells and border cells in the medial entorhinal cortex and subiculum [79, 80, 81], and head-direction cells in limbic brain regions (including the presubiculum and entorhinal cortex [82]), these cells form a neural correlate of space: the cognitive map. In order to build a robust and accurate map, the ability to move around the respective environment appears essential. A process known as path integration relates self-motion information to velocity as well as angular and linear direction to pinpoint one's current location, orientation and trajectory [83]. This information is integrated with geometric information from the environment, such as boundaries, and featural information, such as textures or sounds, to form a comprehensive map of one's environment.

2.3.2 Behavioural Studies on Spatial Cognition

Theoretical models of spatial cognition have proven difficult to validate given the complexity of the study of electrical activity of living neurons and signaling in humans as well as the low resolution of current neuroimaging techniques. Novel technologies, such as consumer VR systems, have recently allowed for new viewpoints, combining neural and behavioural observations to advance these models.

The behavioural study of spatial cognition in combination with neuroimaging methodologies has proven valuable in confirming the role of the hippocampus in spatial memory. For example, a study using VR technology with London taxi drivers in 1998 showed that activation of the right hippocampus was strongly associated with accurately knowing the location of places and accurately navigating between them [84].

Behavioural studies on spatial cognition have also helped to discern egocentric and allocentric representations in spatial memory tasks. Mou et al. investigated the frames of reference used in memory to represent the spatial structure of the environment [85]. They found that spatial memories are defined with respect to intrinsic frames of reference selected on the basis of egocentric experience and environmental cues.

However, contrasting results have been reported through behavioural studies with human participants with regards to how humans learn and recall locations within an environment. Previous findings suggest that the human brain might combine mechanisms based on geometric properties of the environment with self-motion information [86, 87, 88, 5, 89]. Moreover, it is not clear whether these strategies are the same when encountering real world and VEs with varying levels of immersion.

Previous research has looked at the effects of landmark configuration on search behaviour. Using a desktop VR system, Spetch et al. analysed the effect of expansions of an array of landmarks on the locus of search for an object presented in a location equidistant to the landmarks [86, 87]. Their results indicated that humans focus on locations that preserve all angles to the landmarks, preserving ratios of distances between landmarks rather than distances.

Waller et al. reported contrasting results when exploring the role of metric distances and angular information of landmarks on location learning in immersive VR [89]. Participants observed a cued location in relation to three landmarks in an immersive VE. They were then asked to return to the location during testing. Landmark configuration was modified between learning and testing to differentiate the effects of distance and inter-landmark angular information. They found that, overall, participants relied more on distance information than angular information.

There is also evidence for spatial updating of egocentric representations [88, 90, 91]. Wang & Simons showed that locations of objects on a circular table can be better remembered if participants navigate around it, rather than using the equivalent rotation of the table [88]. These results highlight the role of proprioceptive and vestibular inputs during self-motion.

In 1996 O'Keefe and Burgess designed and conducted a study aimed at identifying the environmental features controlling the location and shape of the place fields of the place cells of rodents [69]. Extending this work, Hartley et al. used a desktop VE to investigate the effect of manipulations to spatial boundaries of on object location during learning and recall with human participants [5]. The aim of this study was to conceptually replicate previous studies with rodents where alterations to the environmental geometry caused changes in place cell firing. Participants were presented with an object in a rectangular arena, with distant features to help orient themselves. After the learning stage and a brief delay, they re-entered the arena and were asked to mark the location where the object had been in the learning stage. The geometry of the environment was altered between the stages of learning and recall in some of the trials. Response data was compared with a series of spatial distributions predicted by various geometric models described in Section 2.3.3. The experiment found that responses that maintained fixed distances from nearby walls were more common after expansions of the arena and for locations nearer to the boundaries, whereas responses that preserved fixed ratios between opposing walls were more common after contractions of the arena and for locations nearer to the centre. A model derived from response properties of place cells in the rat hippocampus which matches distances of the cue to the four boundaries of the arena was the best fit for their results. Hartley et al. concluded that their results were consistent with the neural representation of location in the hippocampus [79, 5].

Building on the idea that the cognitive map is an exact and corresponding representation of space, the map should distort in precise correlation with changes to spatial geometry. Through two behavioural user studies (presented in Chapter 3 and Chapter 4) we relate our observations to neuronal knowledge derived from single cell recordings from rodent studies, and test the validity of the different models of spatial representation, crucially focusing on the role of physical immersion (and resulting embodiment and embeddedness) in an immersive VR experience. In contrast to earlier, albeit desktop based, studies that found evidence for geometric computations built on distances and ratios to environmental boundaries [86, 5], our hypothesis was informed by observations made in a pilot experiment. We propose and present evidence that when self-motion is available, models combining geometric properties and path integration could hold greater validity for object location memory.

2.3.3 Object Location Models

Previous studies have looked at the effects of altering the geometry of an environment on object location learning and placement during learning and recall [5]. These studies have compared the spatial distribution of participant responses with locations predicted by different geometric models, some derived from previous neurophysiological experiments [92, 79]. The relevant models, described below, were selected and used in our analysis in Chapter 3 and Chapter 4. The predicted locations following transformations of the spatial boundaries in our study are based on these models and were compared with participant response locations. All model definitions are restricted to environments consisting of four main walls and are further illustrated in Figure 2.1.

- Fixed Ratio Allocentric (FRA) model: an object's location is represented as the ratio of distances between opposing walls [5].

- Fixed Ratio Egocentric (FRE) model: an object's location is represented as the ratio of distances between opposing walls [5]. The participant reorients to the initial facing direction (see Figure 2.1).

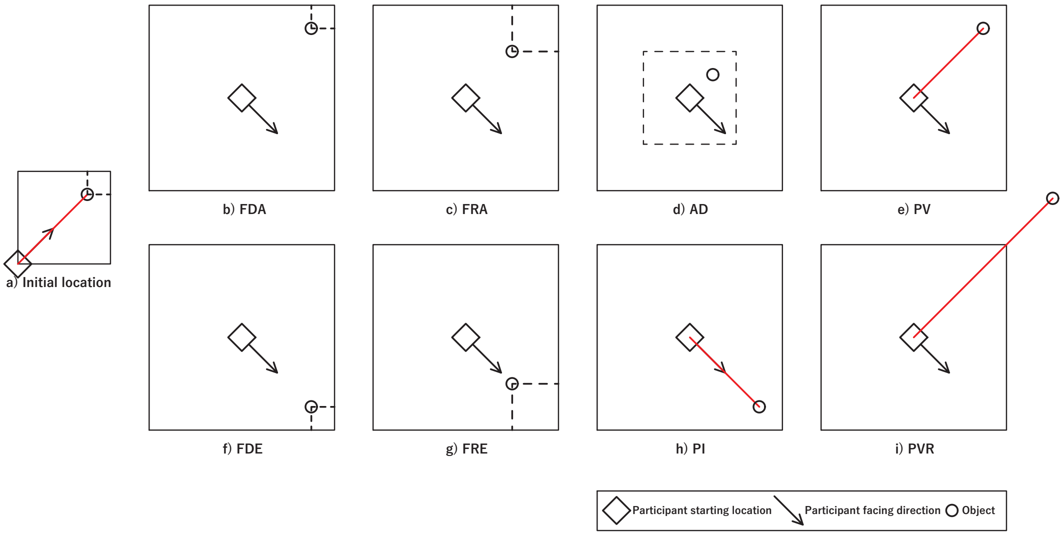
- Fixed Distance Allocentric (FDA) model: an object location is represented as the perpendicular distance from it to its two closest walls [5].

- Fixed Distance Egocentric (FDE) model: an object's location is represented as the perpendicular distance from it to the two closest walls [5]. The participant reorients to the initial facing direction (see Figure 2.1).

- Absolute Distance (AD) model: an object's location is represented by the absolute distance to its original location in world coordinates, regardless of any changes in the environment geometry.

- Path Integration (PI) model: an object location is represented as the vector resulting from a cumulative record of the movements made by the participant from an initial location to the object [93]. The participant replicates the movement maintaining the current facing direction.

- Path Vector (PV) model: based on the PI model, an object's location is represented as the vector resulting from a cumulative record of the movements made by the participant from an initial location to the object [93]. The participant replicates the movement by reorienting to the initial facing direction (see Figure 2.1).



a) Initial location

b) FDA

c) FRA

d) AD

e) PV

f) FDE

g) FRE

h) PI

i) PVR

Figure 2.1: 2D top views of the models in an example scenario for a room transition where the length of the room walls is scaled by a factor of two between learning and placement. Room a) shows the object's initial location (circle) and participant starting location and facing direction (box and arrow). The rest of the sketches show the predicted object location for each model for a new given participant starting location and facing direction (box and arrow). For the Absolute Distance model (d), the initial environment has been overlaid to illustrate that the object XY coordinates for the model coincide with the initial object location. Please note that, under this specific configuration, FDA and PV predictions are the same. These would be different if the participant starting location was different. Also note that in this example and under this configuration the PVR model (i) provides with a predicted location that falls outside the boundaries of the placing room, which in our study would be considered a null model for a given change in environment geometry.

- Path Vector Ratio (PVR) model: an object's location is represented as a scaled vector resulting from a cumulative record of the movements made by the participant from an initial location to the object [93]. The vector is scaled in a way that linearly matches the environment transformation between learning and testing. The participant replicates the movement by reorienting to the initial direction (see Figure 2.1).

2.4 Summary

This chapter has been divided into three main sections. Section 2.1 introduces the motivation for the work presented in this thesis. We discuss the potential benefits of virtual training. We also introduce the common interest in the fields of neuroscience, experimental psychology and VR to advance our knowledge on how humans understand, acquire, store and use representations of space.

Section 2.2 discusses related work on the parameters that mediate the transfer of knowledge from virtual to real environments. We review definitions on the concept of *immersion* and describe how this term is used throughout the thesis. Across all studies, we consider real world learning as the highest level of immersion, followed by HMD learning and then desktop learning. We discuss previous work on *environmental fidelity* and the distinction between two broad types of environmental cues: geometric and featural. We introduce self-avatars and their role in distance estimation accuracy in VEs, noting that previous work has shown that low geometry avatar representation or single point tracking can degrade spatial perception. The work listed on environmental fidelity and self-avatar was used to design the experiment presented in Chapter 5. We review the benefits of using VEs as proxies in research. We also examine past studies on the use of VEs in training, which informed the design of the study presented in Chapter 6.

Section 2.3 contains an overview of human spatial cognition. We discuss the neural correlates of spatial representations, introducing the cognitive map as well as egocentric and allocentric reference frames. We review behavioural studies on spatial cognition, which have proven valuable in advancing our understanding of human spatial cognition given the difficulty of validating theoretical models. Finally, we introduce a series of object location models derived from previous neurophysiological experiments. These models were compared with the spatial distribution of participant responses in the studies presented in Chapter 3 and Chapter 4.

Chapter 3

Experiment: Distorting Physical Space

The question of how humans remember space and the objects within it is crucial in the design of VEs for spatial training. In this chapter we present a study on spatial memory in physical space. Participants were asked to complete a simple spatial memory task: to collect an object in a room, exit the room, re-enter the room and then place the object back where they had found it. The room was geometrically transformed between collection and placement of the object. The participants' responses were compared with a set of models derived from previous neurophysiological experiments with rodents and desktop VR studies in which geometry was manipulated between exposures to the environment (see Section 2.3.3). Results suggest that models which combine memory for geometry and self-motion may hold greater validity in describing human spatial memory.

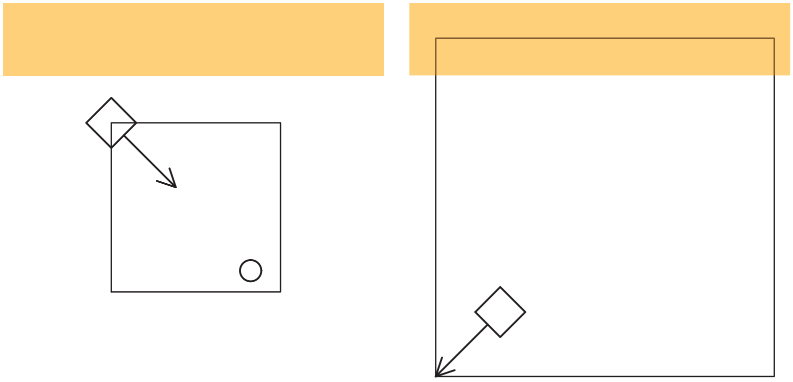
3.1 Experimental Design and Hypotheses

Participants entered a room in one of the three spatial configurations and were asked to collect an object. After a period of time in a separate physical waiting area, they entered a different configuration of the room and were asked to place the object back where they had initially found it. This process was repeated twice. The experimental trials and room dimensions are shown in Table 3.1. Room configuration, object location and participant starting location as well as facing

Figure 3.1: Photograph of the physical reconfigurable room on the UCL PAMELA research facility platform.

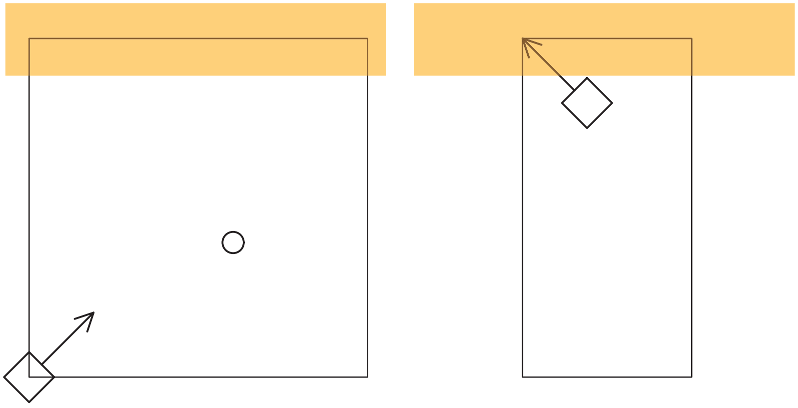
direction for each trial are shown in Figure 3.2. Figure 3.1 shows a photograph of the physical reconfigurable room on the UCL PAMELA research facility platform. Figure 3.3 shows a photograph of the physical reconfigurable room and object in the $2.4 \times 2.4m$ configuration (corresponding to the learning stage of the first trial). Figure 3.4 shows a photograph of the physical reconfigurable room and object in the $4.8 \times 4.8m$ configuration (corresponding to the learning stage of the second trial).

Participants could navigate the room by physically walking around it. We recorded participant navigation in the room as well as the location of the object after being placed by the participant in the placement stage of each trial. We then compared participant behaviours with the different models (see Section 2.3.3). This was achieved by calculating the Euclidean distance between object location as placed by the participants and the location predicted by these models for each trial. Two selection criteria, one based on quadrants and the other based on distance, are also presented to further illustrate our results. These are described in Section 3.3.



a) Trial 1 - Learning stage

a) Trial 1 - Placing stage



c) Trial 2 - Learning stage

d) Trial 2 - Placing stage

◇ Participant starting location ↘ Participant facing direction ○ Object

Figure 3.2: 2D top views of the room layouts for each of the trials. Circles represent the object's location during the learning stage for each trial. Box and arrows indicate the participant's initial location and facing direction. The amber rectangle represents the two rows of amber LED lights suspended over the room.

Since no salient landmark cues were available in the room, it was decided to use light as an orientation cue. This would allow participants to reorient themselves when starting trial stages at different starting locations and with different facing directions. The in situ LED lighting system suspended from the ceiling at UCL PAMELA facility was used for this and all other light sources were turned off. All lights were set to white except for two rows of amber light on the north side of the room (see Figure 3.5). The row of amber lights remained constant throughout the experiment and provided a non-geometric and non-landmark cue. A blackout

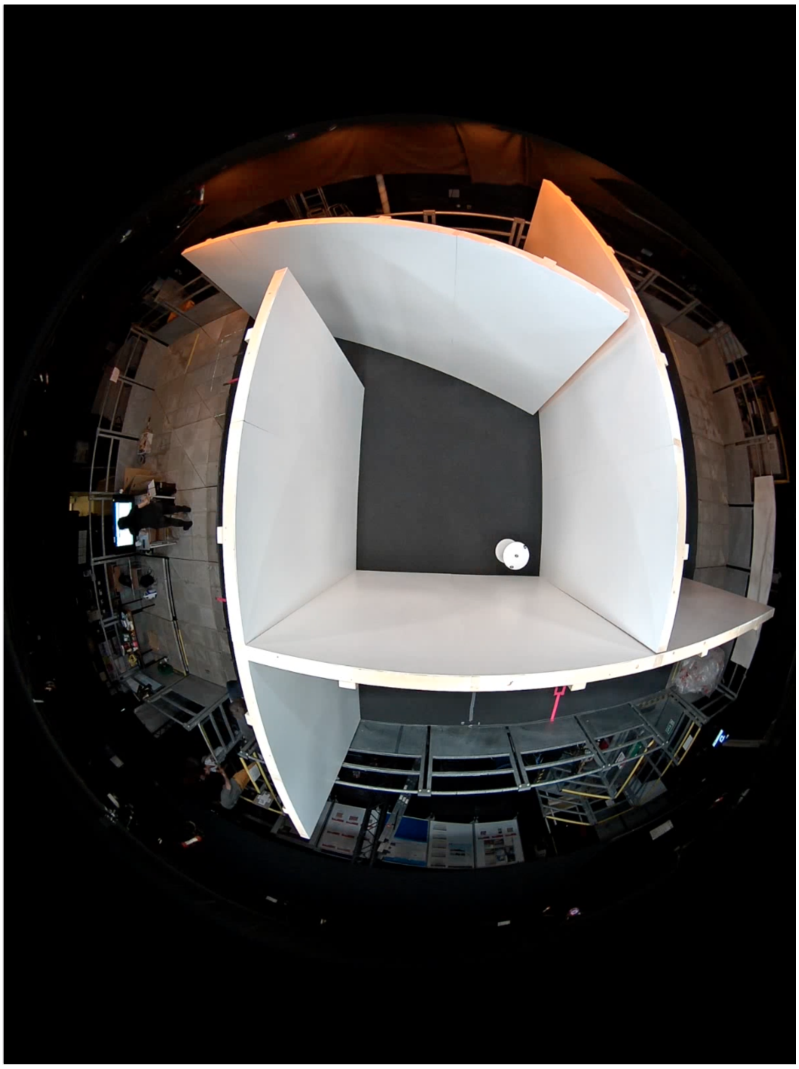


Figure 3.3: Photograph of the physical reconfigurable room and object in the 2.4 × 2.4m configuration (corresponding to the learning stage of the first trial).

Table 3.1: Experimental trials, stages and room dimensions (in m).

Trial	Stage	Room Dimensions (m)
1	Learn	2.4×2.4
	Place	4.8×4.8
2	Learn	4.8×4.8
	Place	2.4×4.8

curtain was suspended from the lighting grid above the platform around the room so participants could not see any external cues within the UCL PAMELA facility.

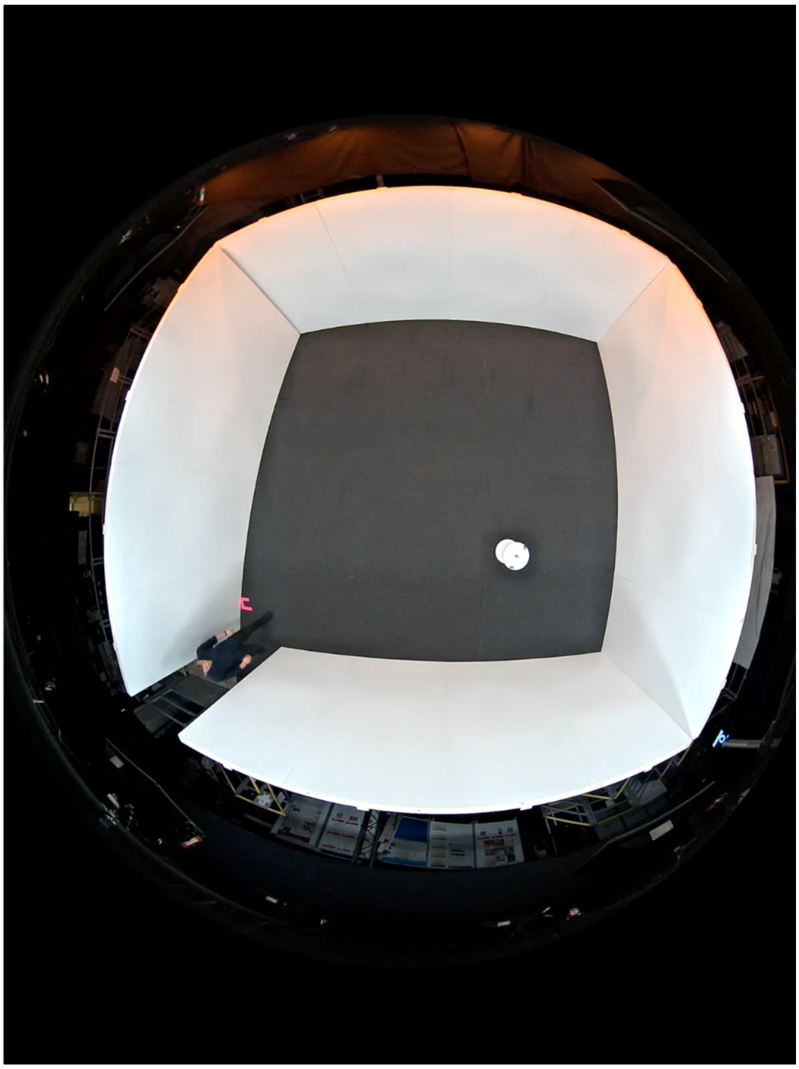


Figure 3.4: Photograph of the physical reconfigurable room and object in the 4.8 × 4.8m configuration (corresponding to the learning stage of the second trial).

Groups of six to nine participants were recruited for each lab session. They were gathered together in the waiting room to complete individually each step of the experiment, one at a time. Experimenters ensured that the waiting times between steps were kept constant throughout the entire study, regardless of the number of participants present at each session.

The study was conducted by three experimenters, each responsible for three distinct tasks: chaperoning the waiting room, overseeing the testing room (containing the reconfigurable room) and escorting between the two rooms. The experimenter chaperoning in the waiting room was responsible for distributing

Figure 3.5: Photograph of the real world reconfigurable environment showing the two rows of amber LED lights used as an orientation cue.

and collecting the different questionnaires to the groups of participants. The experimenter in the testing room was responsible for the technical equipment as well as positioning participants in the correct starting locations and facing directions for each trial. The third experimenter was in charge of escorting participants from and to the waiting room, as well as helping the other experimenters as needed.

The experimental design, data collection and preliminary data analysis exploring a subset of the models, as well as a pilot study were reported in two unpublished MSc student theses [94, 95]. The main purpose of the pilot study was to test the setup and logistics of the physical room transitions. The experience of running the pilot study informed the experimental design of the study presented in this chapter, with the goal of minimising the number of physical room transitions in each session.

Based on preliminary results obtained from the pilot experiment and in contrast to earlier desktop VR studies [86, 87, 89, 5], we hypothesised that responses maintaining fixed distances from nearby walls would be more common after expansions of the room and for objects closer to the boundaries of the room, and fixed ratios between opposing walls would be more common after contractions of the environment and for locations closer to the centre of the environment.

3.2 Method

3.2.1 Materials

A physical room was built consisting of a reconfigurable four-walled room, erected on the moveable platform at the UCL PAMELA research facility. The platform at PAMELA is made of 58 1.2m x 1.2m modules. These are controlled wirelessly and provide interchangeable surfaces. The lighting system at the facility allows the simulation of numerous conditions, ranging from daylight to darkness, ambient to direct light, or variously coloured light scenarios. Each of the four walls was custom-made from four plywood panels measuring $1.2 \times 2.4m$. The largest room was 4.8 wide $\times 2.4m$ long, the smallest room was 2.4 wide $\times 2.4m$ long. All walls were painted white and a grey carpet was placed to cover the original flooring.

It was necessary to provide a way for participants to orientate themselves globally, as they were starting in a different corner each time. Hartley et al.'s experimental design used a mountain range projected at infinity as a way of providing participants with a directional cue [79]. In a physical setup this mode of operation is not feasible. Projecting a panorama onto the wall would provide landmark rather than distal cues; resulting in an environmental feature rather than a constant allowing participants to orientate within the environment. We therefore used PAMELA's programmable LED lighting to light one side of the room in amber as an orientation cue (as detailed in Figure 3.1).

The object used for both learning and recall was a white Tam Tam plastic stool from Habitat (shown in Figure 3.6). The stool is easy to hold when blindfolded, lightweight and provides no misleading (or otherwise) orientation cue to the participant. This object was selected due to its specific rotational symmetry along the vertical axis, eliminating the question of object orientation. The stool has a diameter of $0.31m$ at the widest section and a height of $0.45m$.

Figure 3.6: Image of the object used in the study.

Twelve Optitrack cameras were suspended from the UCL PAMELA lighting structure and used to track the final location of the stool as well as the participant navigation in the room. To enable object tracking we placed three retro-reflective markers on the top of the object. The cameras tracked the location of both the object reflecting infrared light. The data was recorded using the OptiTrack Motive software. As a backup to the tracking system failing a video camera was placed above the center of the room so that the final object location could be estimated in the event of tracking failure.

3.2.2 Participants

A total of 29 participants (18 female, 11 male; average age 44.8 years, SD = 15.5) were recruited from the student and staff population at UCL. Participants were required to be aged between 18 and 65 and have been based in London for at least five years as recent findings indicate cultural variation in spatial navigation strategies [96]. One participant was older than the required age range and was excluded from the study. Another participant was partially sighted and therefore was also excluded. All participants signed a consent form and the study was approved by the UCL Research Ethics Committee (Project ID: CPB/2013/015). Participants were paid £10 per hour for participation. The experimental task lasted approximately 1.5 hours.

3.2.3 Procedure

The study was conducted in two different rooms at UCL PAMELA: a waiting room and a testing room, which contained the room setup for the study. Each session of the study began with groups of participants (between six and nine) gathering in the waiting room for an induction. They were asked to sign a paper copy of the consent form and read an information sheet with written instructions describing the experimental task. One of the experimenters then proceeded to explain the task and participants had the chance to ask questions. An outline of the task was also displayed on a white board in both the waiting and testing rooms. Participants were also told that the experimenters would remind them throughout the procedure which stage would come next. They were also shown a physical version of the Tam Tam stool from Habitat for reference.

The experiment consisted of two trials, specified in Table 3.1. Each trial involved two stages: a learning stage and a placing stage. In the learning stage participants were asked to collect the object. In the placement stage they were asked to place the object back where they had found it in the learning stage. Participants completed the learning stange and the placing stage of each trial individually. Table 3.2 contains an outline of the experimental task with the steps followed by participants between the waiting and testing rooms.

Before accessing the testing room in each of the trial stages, participants were blindfolded and asked to grab the two ends of a cardboard tube with their hands. The facilitator escorting participants then grabed the middle of the pole and guided participants into the testing room. Along the way, the facilitator would disorient and guide the participant to the starting location for the corresponding stage (using figures of eight). Participants were disoriented to stop them from finding correspondences between the starting position during the learning stage and the starting position during the placement stage.

When participants reached the testing room, the facilitator guided them to the task starting position within the four reconfigurable walls through an open corner. The facilitator would leave space within the four reconfigurable walls

Table 3.2: Experimental task outline with steps followed by participants. Individual stages were completed by all participants, one at a time following the same order for each stage. All transitions between waiting and testing involved an experimenter escorting the participant between these rooms.

Step	Room	Trial	Stage	Group/Individual
1	Waiting	Induction	-	Group
2	Testing	Practice (VR version) Trial 1	- Learn	Individual
3	Waiting	Questionnaires	-	Group
4	Testing	Trial 1 Trial 2	Place Learn	Individual
5	Waiting	Questionnaires	-	Group
6	Testing	Trial 2	Place	Individual
7	Waiting	Payment	-	Group

and experimental assistants would close them, with the participants inside. The facilitator would then indicate to the participants that they could remove the blindfold and complete the task. They were also asked to indicate to the experimenter when they had completed each stage. Once the task was complete, participants were asked to place their blindfold on again and wait for the facilitator to hand over the cardboard tube in order to be escorted to the next stage of the study.

Participants were asked to complete all stages as quickly and as accurately as possible. They were not informed about the changes in room configuration between learning and placement stages and variations of starting location and facing direction. They were advised that there was no correct response and to try their best if they were in doubt as to where to place the object.

While participants waited for their turn in the waiting room, participants were asked to complete a Santa Barbara Sense-of-Direction Scale as well as a Myers Briggs Personality Test [97, 98]. Participant payment was processed once all experimental trials and questionnaires were completed.

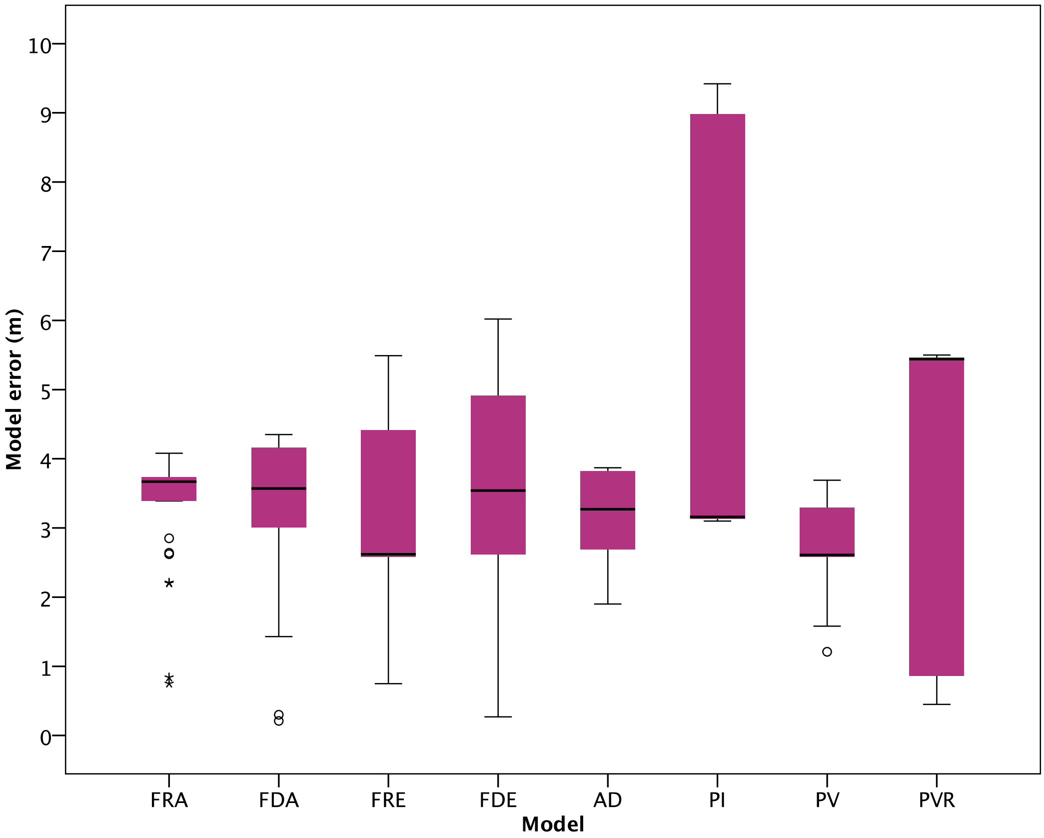


Figure 3.7: Boxplot for the distance between object positions as placed by participants in the placement stage and the original and distorted locations calculated from the different models for each of the trials (model errors) for Trial 1. Medians are shown as dark horizontal lines. Boxes represent the IQR. Whiskers represent either the extreme data points or extend to 1.5 x IQR. Outliers (data points outside the whiskers) are shown by circles. A value, X, is an outlier if X < lower quartile − 1.5 x interquartile range or if X < upper quartile + 1.5 x IQR. See Table 3.4 for pairwise interactions.

3.3 Results

Participant response placement data was used to calculate the Euclidean distance between object positions as placed by participants and locations calculated from the different models for each of the trials (see Section 2.3.3). We label this distance as the model error and use it to quantify how accurate a model is in predicting participant responses for the given trials. Figure 3.7 and Figure 3.8 show boxplots for mean model errors for Trial 1 and Trial 2, respectively. Figure 3.9 shows cluster heat maps and scatter plots showing participant response XY placement during the testing stage each trial.

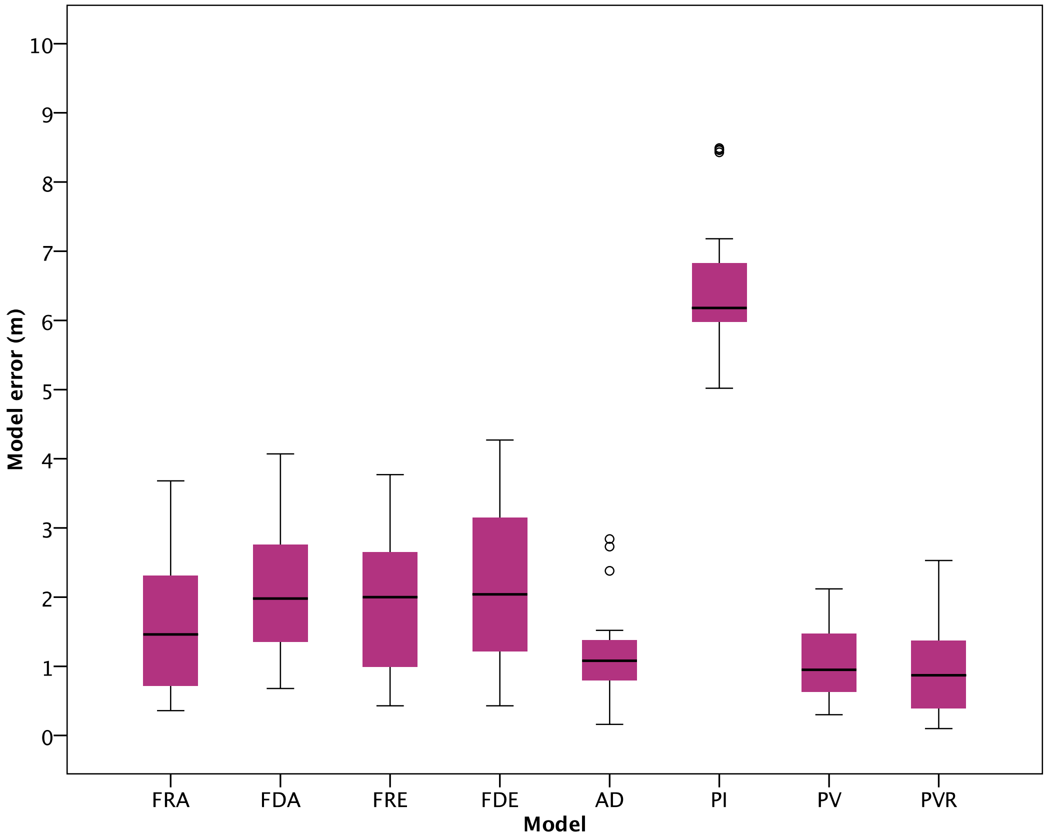


Figure 3.8: Boxplot for the distance between object positions as placed by participants in the placement stage and the original and distorted locations calculated from the different models for each of the trials (model errors) for Trial 2. Medians are shown as dark horizontal lines. Boxes represent the IQR. Whiskers represent either the extreme data points or extend to 1.5 x IQR. Outliers (data points outside the whiskers) are shown by circles. A value, X, is an outlier if $X <$ lower quartile $-$ 1.5 x interquartile range or if $X <$ upper quartile $+$ 1.5 x IQR. See Table 3.5 for pairwise interactions.

Given that no standard procedure for analysing this type of data exists, for informative purposes we defined two different selection criteria to determine which models or subset of models could best describe individual participant responses for each trial (see Section 2.3.3). This was done in order to filter out outliers or participant responses that lay far from the predicted locations. The first selection criteria, refered to as Quadrant Criteria (QC), divided each room into four quadrants. These quadrants were defined by dividing the room into four equal sections with two conceptual lines perpendicular to the walls, intersecting at the centroid of the room. Following this selection criteria, only responses falling in the same quadrant as the

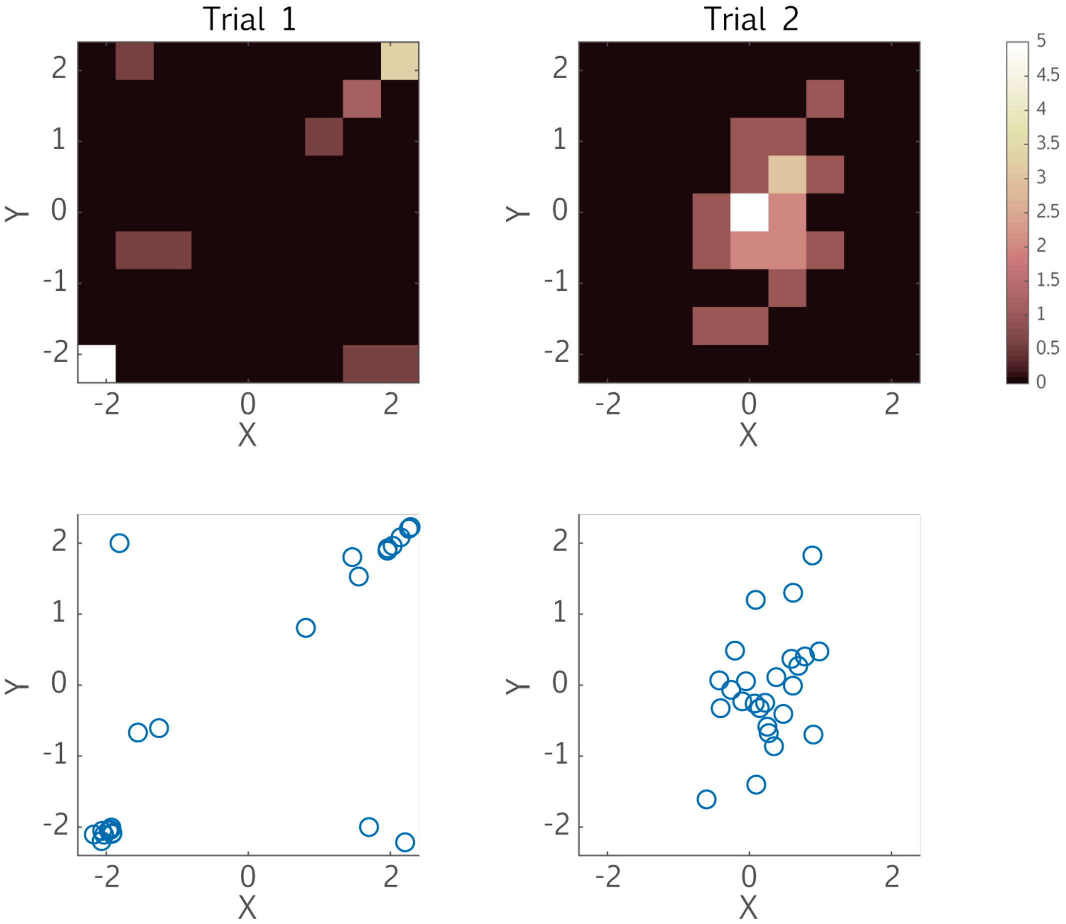


Figure 3.9: Cluster heat maps and scatter plots showing participant response placement during the testing stage for trial 1 (left) and trial 2 (right).The columns represent X-axis and the rows represent Y-axis positions (in m). Each cell is colorized based on the level of response counts in each region.

predicted location given by the corresponding model were considered. The second criteria, labelled Distance Criteria (DC), was defined such that only responses falling within 0.5m from the predicted location given by the corresponding model were computed. Table 3.3 shows the number of participant responses that met each of the selection criteria for each model in each of the trials and overall. We also included results of the union of both criteria.

Table 3.4 and Table 3.5 show pairwise comparisons between models using a paired-samples t-test with Bonferroni corrections for Trial 1 and Trial 2, respectively, without excluding outliers.

For Trial 1, we found a statistically significant difference between the FRE model $(M = 3.11, SD = 1.66)$ and the FDE model $(M = 3.48, SD = 1.90)$. We also

Table 3.3: Participant responses that met each of the selection criteria for each model in each of the trials and overall (both trials combined). The last row indicates the total number of participant responses (N). Note that model PI is a null model (the predicted object location falls outside the boundaries of the room) and therefore does not belong to any of the four valid quadrants. Both the QC and the DC are not applicable in this case.

	QC			**DC**			**QC ∪ DC**		
	Trial 1	Trial 2	Overall	Trial 1	Trial 2	Overall	Trial 1	Trial 2	Overall
FRA	2	8	10	0	5	5	0	2	2
FRE	2	3	5	2	0	2	2	0	2
FDA	9	4	13	0	1	1	0	0	0
FDE	9	14	23	5	1	6	5	1	6
AD	2	8	10	0	4	4	0	2	2
PI	-	-	-	-	-	-	-	-	-
PV	18	14	32	0	3	3	0	2	2
PVR	9	4	13	1	8	9	1	3	4
N	29	29	58	29	29	58	29	29	58

Table 3.4: Pairwise comparison t and p values between all models for Trial 1. The last row shows mean model error values for each model (*M*). Interaction is not significant unless it is explicitly indicated as specified in the legend.

	FRA	**FRE**	**FDA**	**FDE**	**AD**	**PI**	**PV**	**PVR**
FRA		0.63	-0.12	-0.35	1.79	-3.53[a]	2.68[b]	-1.15
FRE			-0.68	-3.24[b]	-0.16	-3.11[b]	0.87	-2.10[b]
FDA				-0.31	0.98	-3.74[a]	2.54[b]	-0.96
FDE					-0.84	-2.32[b]	1.61	-1.33
AD						-3.78[a]	2.14[b]	-1.80
PI							5.86[a]	1.73
PV								-2.13[b]
M	3.34	3.11	3.35	3.48	3.16	5.37	2.80	3.81

[a] interaction is significant at the 0.001 level (two-tailed)

[b] interaction is significant at the 0.05 level (two-tailed)

Table 3.5: Pairwise comparison t and p values between all models for Trial 2. The last row shows mean model error values for each model (*M*). Interaction is not significant unless it is explicitly indicated as specified in the legend.

	FRA	**FRE**	**FDA**	**FDE**	**AD**	**PI**	**PV**	**PVR**
FRA		-1.73	-11.26[a]	-2.78[b]	3.17[b]	-18.01[a]	2.09[b]	2.06[b]
FRE			-0.36	-4.75[a]	3.86[a]	-25.24[a]	5.04[a]	4.68[a]
FDA				-0.77	6.39[a]	-15.68[a]	4.55[a]	4.16[a]
FDE					4.64[a]	-19.72[a]	5.71[a]	5.20[a]
AD						-28.01[a]	-0.10	0.52
PI							32.67[a]	37.08[a]
PV								0.80
M	1.53	1.96	2.05	2.27	1.12	6.46	1.13	1.05

[a] interaction is significant at the 0.001 level (two-tailed)

[b] interaction is significant at the 0.05 level (two-tailed)

found that model PI $(M = 5.37, SD = 2.77)$ had a significantly higher model error than all other models, except model PVR $(M = 3.81, SD = 2.13)$. This was due to the fact that the predicted location from this model fell outside the room, making this model null for this trial. We found that, overall, model PV $(M = 2.80, SD = 0.63)$ had the lowest model error, and was significantly lower than model FRA $(M = 3.34, SD = 0.87)$, model FDA $(M = 3.35, SD = 1.10)$, model AD $(M = 3.16, SD = 0.66)$, the null model PI $(M = 5.37, SD = 2.77)$ and model PVR $(M = 3.81, SD = 2.13)$, but not significantly different from the egocentric models (models FRE and FDE). This model was also the one with the highest number of paricipant responses meeting the QC for Trial 1, but not the DC (see Table 3.3). This is due to a high overall model error across all participants, higher than 0.5m.

For Trial 2, we found that the lowest model error corresponded to model AD $(M = 1.12, SD = 0.65)$, model PV $(M = 1.13, SD = 0.59)$ and model PVR $(M = 1.05, SD = 0.77)$. These models were significantly lower from all other models, but showed no significant difference amongst themselves. Note that, similar to Trial 1, the location predicted by the PI model fell outside the room, making this model null for this trial. For model AD distances are preserved in world coordinates, regardless of any changes in the environment geometry. This suggests that participants may have attempted to map the object location in the room using static, absolute cues (such as marks on the carpet). Other significant interactions were found and are shown in Table 3.5. Similar to the results of Trial 1, model PV was one of the two models with the highest number of participant responses meeting the QC for Trail 1 (see Table 3.3). This was not the case for model AD and model PVR. Due to the high overall model error across all participants, the DC is not very relevant in explaining participant responses for this trial.

Results for the Santa Barbara Sense-of-Direction Scale and the Myers Briggs Personality Test, and their relation with object location memory results are not reported as part of this thesis. Separate statistical analysis showed no effect of gender or age on model error.

3.4 Discussion

Overall, results highlight strong differences in participant behaviour, with contrasting models describing different response object locations. No single model can account for all participant behaviour. Models based on self-motion seem to hold greater accuracy in describing part of our responses, whereas the model that preserves the absolute distance to its original location in world coordinates, regardless of any changes in the environment geometry, seemed to work best in explaining other participant responses. More data would be needed to understand what makes participants behave in ways best represented by different models.

Constructing a reconfigurable, featureless large-scale room in the real world is a complex exercise. Progressive wear and tear caused by room reconfiguration can inevitably create cues on the walls and floor that participants could be using to learn and recall the object's location. This could nullify the intended effect of the experimental design. Similarly, lack of control over external sounds can provide participants with strong directional cues when taking part in the different trials.

The relevant models decribed and explored in our study represent some individual participant behaviour. Unlike Hartley et al. [5], we did not find strong evidence to suggest that responses maintaining fixed distances from nearby walls were more common after expansions of the room and for objects closer to the boundaries of the room, and fixed ratios between opposing walls were more common after contractions of the environment and for locations closer to the centre of the environment. Models based on self-motion seemed to better explain part of our participant responses. This highlights the need to explore more complex paradigms which combine models based on spatial geometry and models based on self-motion or navigation.

The relative success of the AD model in Trial 2, where an object location is represented by the absolute distance to its original location in world coordinates, regardless of any changes to the environment geometry, could have several explanations. On the one hand, participants could have used some of the cues produced by wear and tear of the physical walls or the carpet. This includes a line between the seams of two pieces of carpet as well as the LED light grid on the ceiling of the room. These did not change between trial stages, providing participants with cues that remained constant and would have aided in placing the object in its original location in world coordinates.

Chapter 4

Experiment: Distorting Virtual Space

In this chapter we present the VR equivalent study to the one presented in Chapter 3. Results from our study highlight the role of spatial layout as well as the user's starting location and facing direction, which have a strong effect on participant behaviour. Similar to the study presented in Chapter 3, results suggest that models which combine memory for geometry and self-motion may be better at describing object location memory in immersive VEs. All in all, our VR study on spatial cognition offers promising outcomes, further illustrating its potential as a fundamental research tool in this and similar fields of study.

4.1 Experimental Design and Hypotheses

The experimental design of our immersive VR study is based on Hartley et al.'s 2004 desktop VR study [5]. It is part of a larger research project, which includes a real world version of the study, presented in Chapter 3. In the study presented in this chapter we generated the VEs from a series of 3D scans of the physical environments, where four white wall panels made of four flats each were used to build a space in three different configurations (see Figure 4.1).

Participants used a HMD to enter the VE with the reconstructed room in one of the three configurations and were asked to collect a virtual object. All 3D scans were performed under the same lighting conditions, providing very similar virtual

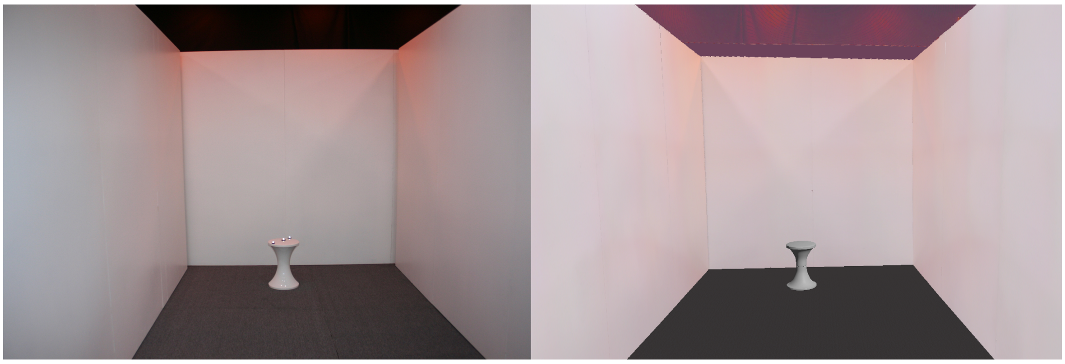


Figure 4.1: Photograph of the real world reconfigurable environment and object (left) and render of the VE and virtual object (right) used in the study. The real world environment was 3D scanned to construct the VEs. A 3D virtual replica of the object was used in the study. Three retroreflective markers are shown in the left image on the object since these were used to 3D track its location in the real world. Markers were not necessary in our virtual version of the task.

replicas of the real world environments. After a period of time in a separate physical waiting room, they entered a different configuration of the VE and were asked to place the object back where they had initially found it. This process was repeated twice. The experimental trials are shown in Table 3.1. Room configuration, object location and participant starting location as well as facing direction for each trial are shown in Figure 3.2.

Participants could navigate the VEs by physically moving around them within the tracked 3D space. We recorded participant navigation around the VEs as well as the location of the virtual object after being placed by the participant in the placement stage of each trial. We then compared participant behaviours with the different models (see Section 2.3.3). This was achieved by calculating the Euclidean distance between object location as placed by the participants and the location predicted by these models for each trial. Two selection criteria, one based on quadrants and the other based on distance, are also presented to further illustrate our results. These are described in Section 3.3.

Groups of six to nine participants were recruited for each lab session. They were gathered together in a waiting room. They completed individually each step of the experiment, one at a time. Experimenters ensured that the waiting times between steps were kept constant throughout the entire study, regardless of the number of participants present at each session.

The study was conducted by three experimenters, each responsible for three distinct tasks: chaperoning the waiting room, overseeing the testing room (containing the VR system) and escorting between the two rooms. The experimenter chaperoning in the waiting room was responsible for distributing and collecting the different questionnaires to the groups of participants. The experimenter in the testing room was responsible for the technical equipment as well as positioning participants in the correct starting locations and facing directions for each trial. The third experimenter was in charge of escorting participants from and to the waiting room, as well as helping the other experimenters as needed.

Based on preliminary results obtained from our real world pilot experiment and in contrast to earlier desktop VR studies [86, 87, 89, 5], we hypothesised that, by providing participants with idiothetic cues in both a physical and an immersive VE set-up, models based on self-motion would hold the highest validity for object location memory [95].

4.2 Method

4.2.1 Materials

The three VEs consisted of high fidelity point clouds obtained from 3D laser scans of the real environments, rendered with a bespoke GPU-based point cloud renderer. The real world environment consisted of a reconfigurable four-walled room, built at UCL PAMELA. Each of the four walls was made from four plywood panels and was 4.8m wide x 2.4m tall. All walls were painted white and a grey carpet was placed to cover the UCL PAMELA facility's flooring. 3D scanning was performed with a Faro Focus 3D S120 laser scanner. The scanned floor was substituted by a texturized plane to fill in the missing points from the scanner's dead spot beneath it.

The VEs were rendered in a HTC Vive Developer Edition at 1:1 scale in Unity at 90FPS with a vertical FOV of 60 degrees. The computer had an Intel Core i7-6700 CPU @ 3.40GHz, with 32GB RAM and an Nvidia GTX 980 Ti GPU running Windows 10 Enterprise. The HTC Vive Developer Edition base stations were placed in opposite corners of the 4.8m x 4.8m surface area (corresponding to the largest room configuration).

Participants used one HTC Vive Developer Edition wireless controller throughout the study. In each trial, the application would always begin with an empty virtual space where only the virtual controller was visible. This enabled the experimenters to make sure that the controller was not out of battery and fully functioning. Participants could then begin the corresponding stage of the experiment by pressing the front red system button. During each stage, participants could grab the object by pressing the rear trigger whenever the controller was within a 5cm range of the object. By releasing the trigger button, the object would then fall to the floor level and rotate to a vertical position.

The object used in the study was a virtual replica of a white Tam Tam plastic stool from Habitat, detailed in Section 3.2.1 and shown in Figure 4.1.

4.2.2 Participants

A total of 39 participants (14 female, 25 male; average age 30.8 years, SD = 10.9) were recruited from the student and staff population at UCL. Participants were required to be aged between 18 and 65 and have been based in London for at least five years as recent findings indicate cultural variation in spatial navigation strategies [96]. All participants signed a consent form and the study was approved by the UCL Research Ethics Committee (Project ID: CPB/2013/015). Participants were paid £10 per hour for participation. The experimental task lasted approximately 1.5 hours.

4.2.3 Procedure

The study was conducted in two different rooms at UCL: a waiting room and a testing room, which contained the VR setup. Each session of the study began with groups of participants (between six and nine) gathering in the waiting room for an induction. They were asked to sign a paper copy of the consent form and read an information sheet with written instructions describing the experimental task. One of the experimenters then proceeded to explain the task and participants had the chance to ask questions. An outline of the task was also displayed on a white board in both the waiting and testing rooms. Participants were also told that the experimenters would remind them throughout the procedure which step would come next.

Participants were informed that they would be able to see the SteamVR chaperone grid when using the VR system. The chaperone grid is a security system embedded in SteamVR that renders a virtual blue grid over the VE when the user approaches the boundaries of the tracking space. SteamVR does not allow for deactivation of the chaperone grid for security reasons. Participants were told that when they saw the blue grid this meant they were approximately 40cm away from the physical limits of the room and asked to ignore it when completing the experimental task. They were also shown a physical version of the Tam Tam stool from Habitat for reference.

The experiment consisted of two trials, specified in Table 3.1. Each trial involved two stages: a learning stage and a placing stage. In the learning stage participants were asked to collect the virtual object. In the placement stage they were asked to place the virtual object back where they had found it in the learning stage. Participants completed the learning stange and the placing stage of each trial individually. Table 3.2 contains an outline of the experimental task with the steps followed by participants between the waiting and testing rooms.

For each stage in the testing room, the participant was asked to wear the HTC Vive Developer Edition HMD as well as to hold a controller with his or her preferred hand. The experimenter would take hold of the opposite side of the controller and guide as well as disorient the participant to the starting location for the corresponding stage (using the HMD as a blindfold). Participants were disoriented to stop them from finding correspondences between the physical testing room and the VEs. Starting locations were marked with tape in the testing room to help the experimenters place participants at the starting location and in the correct facing direction for each stage.

Participants were also asked to indicate to the experimenter when they had completed each stage. Once the task was complete, participants were asked to continue placing the controller to be escorted to the next stage of the study. Participants were asked to indicate to the experimenter when they had completed each stage.

Prior to the two experimental trials, participants completed a practice trial. This was to ensure that they had a chance to familiarise themselves with the HTC Vive controller, as well as grabbing and placing the virtual object. The practice VE consisted of an empty space with a floor and a sphere (radius = 0.25m) to practice interacting with the controller. Participants were asked to navigate to the sphere. They were then asked to grab and release it as many times as needed until they felt comfortable with the interaction.

Participants were asked to complete all stages as quickly and as accurately as possible. They were not informed about the changes in VE configuration between learning and placement stages and variations of starting location and facing direction. They were advised that there was no correct response and to try their best if they were in doubt as to where to place the virtual object.

While participants waited for their turn in the waiting room, participants were asked to complete a Santa Barbara Sense-of-Direction Scale as well as a Myers Briggs Personality Test [97, 98]. Participant payment was processed once all experimental trials and questionnaires were completed.

4.3 Results

Participant response placement data was used to calculate the Euclidean distance between object positions as placed by participants and locations calculated from the different models for each of the trials (see Figure 4.5). We label this distance as the model error and use it to quantify how accurate a model is in predicting participant responses for the given trials. Figure 4.2 and Figure 4.3 show boxplots for mean model errors for Trial 1 and Trial 2, respectively. PI, PV and PVR models were calculated for each participant from their 3D tracked starting location and facing direction, as this data slightly varied from participant to participant. This was due to the experimental procedure, which involved the experimenter physically

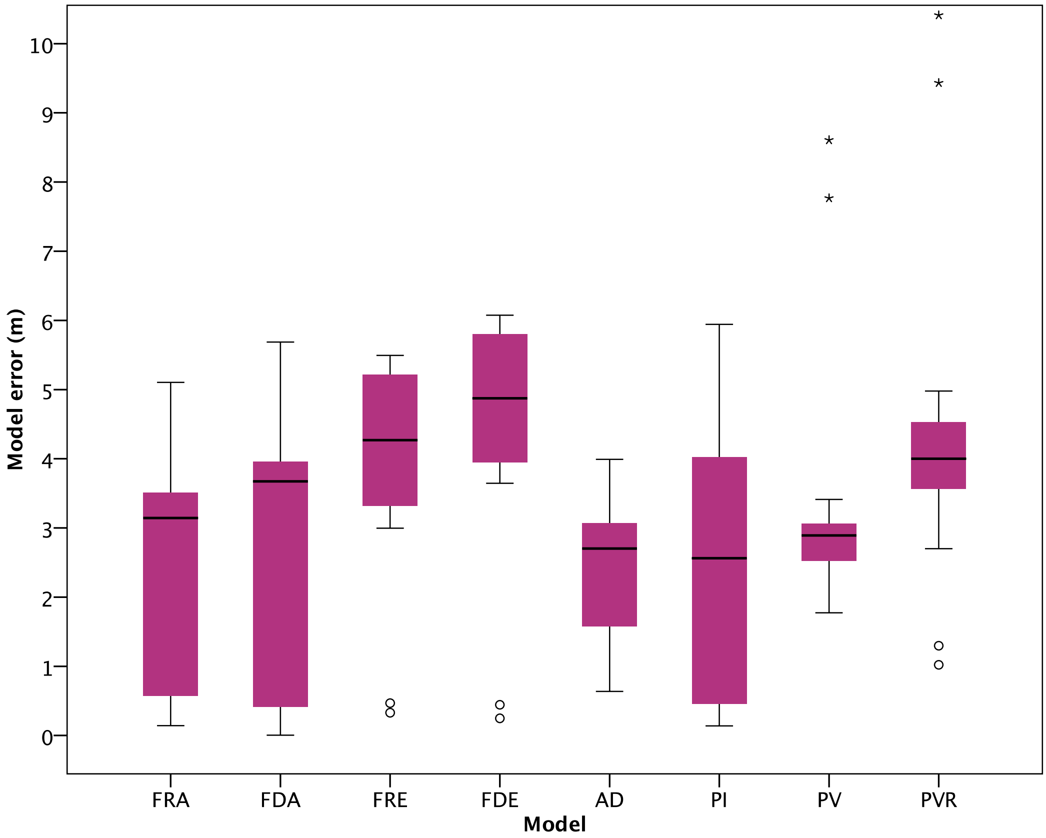


Figure 4.2: Boxplot for the distance between object positions as placed by participants in the placement stage and the original and distorted locations calculated from the different models for each of the trials (model errors) for Trial 1. Medians are shown as dark horizontal lines. Boxes represent the IQR. Whiskers represent either the extreme data points or extend to 1.5 x IQR. Outliers (data points outside the whiskers) are shown by circles. A value, X, is an outlier if X < lower quartile − 1.5 x interquartile range or if X < upper quartile + 1.5 x IQR. See Table 4.2 for pairwise interactions.

guiding and placing participants at the starting location and correct facing direction for every trial stage. Placement data from two participants in Trial 2 was not logged correctly and has therefore not been included in the analysis. Figure 4.4 shows cluster heat maps and scatter plots showing participant response XY placement during the testing stage each trial.

Given that no standard procedure for analysing this type of data exists, for informative purposes we defined two different selection criteria to determine which model or subset of models could best describe individual participant responses for each trial (see Section 2.3.3). The criteria are described in Section 3.3). Table 4.1

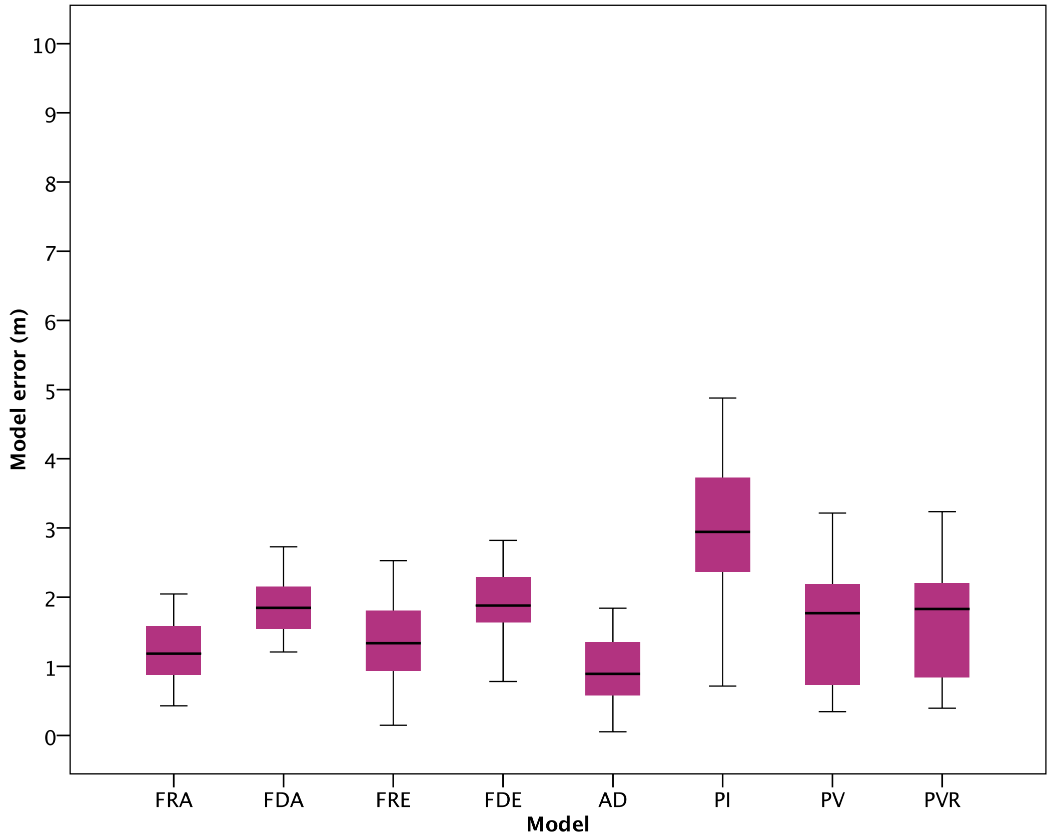


Figure 4.3: Boxplot for the distance between object positions as placed by participants in the placement stage and the original and distorted locations calculated from the different models for each of the trials (model errors) for Trial 2. Medians are shown as dark horizontal lines. Boxes represent the IQR. Whiskers represent either the extreme data points or extend to 1.5 x IQR. Outliers (data points outside the whiskers) are shown by circles. A value, X, is an outlier if $X <$ lower quartile $-$ 1.5 x interquartile range or if $X <$ upper quartile $+$ 1.5 x IQR. See Table 4.3 for pairwise interactions.

shows the number of participant responses that met each of the selection criteria for each model in each of the trials and overall. We also included results of the union of both criteria.

Table 4.2 and Table 4.3 show pairwise comparisons between models using a paired-samples t-test with Bonferroni corrections for Trial 1 and Trial 2, respectively, without excluding outliers. Results show a large number of significant pairwise interactions at the $p<0.05$ level, but we will only report in detail the ones we consider most relevant.

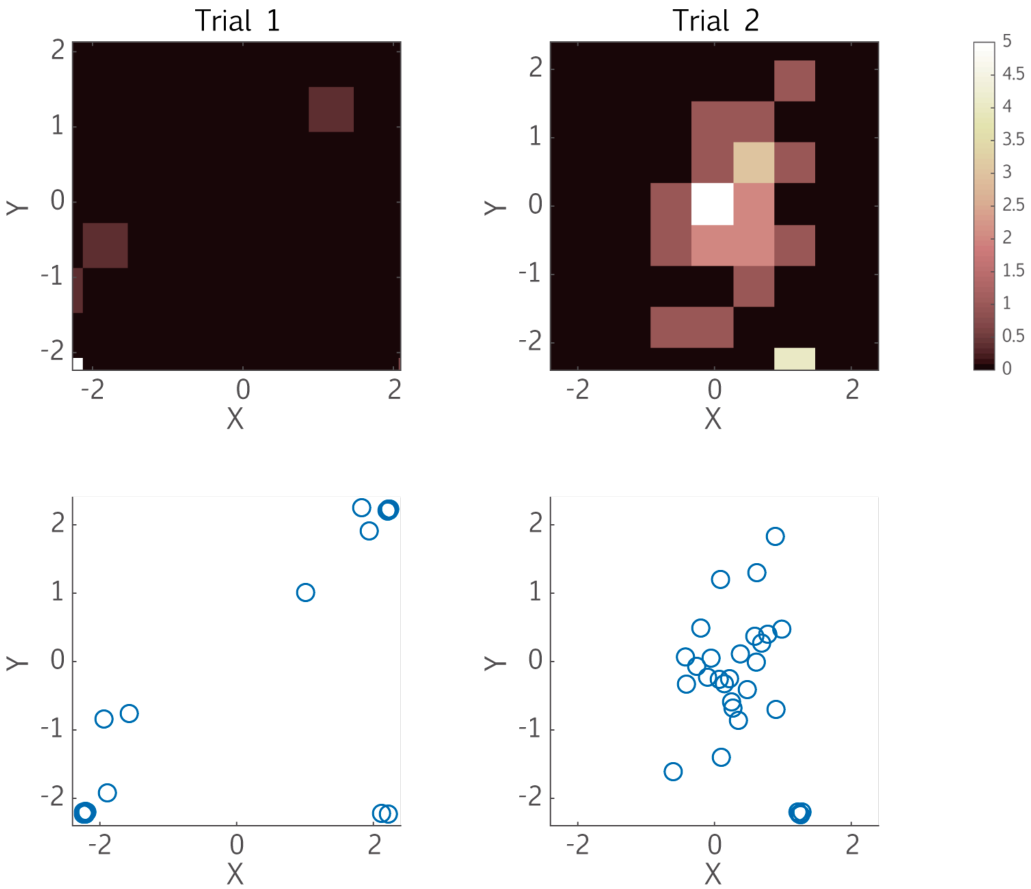
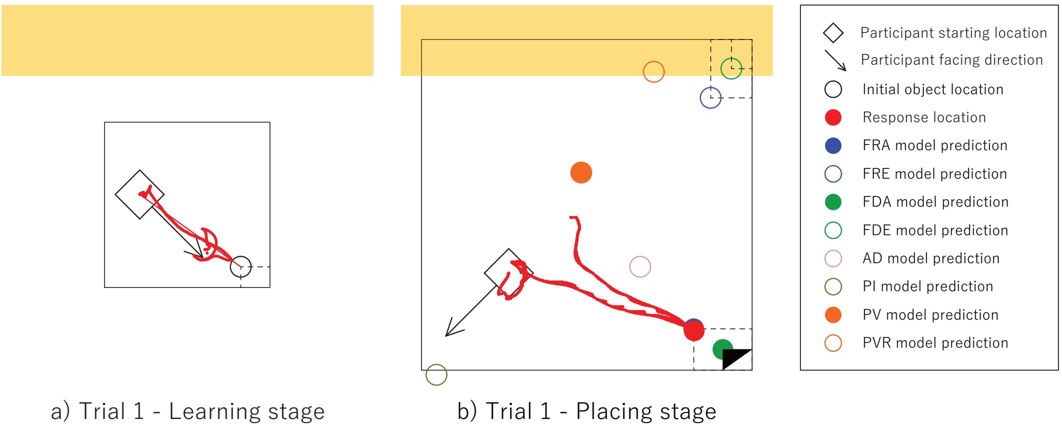


Figure 4.4: Cluster heat maps and scatter plots showing participant response placement during the testing stage for trial 1 (left) and trial 2 (right).The columns represent X-axis and the rows represent Y-axis positions (in m). Each cell is colorized based on the level of response counts in each region.

For Trial 1, there was a statistically significant difference in model error for the FDE model and all other models, with the highest mean model error $(M = 4.68, SD = 1.37)$, followed by the FRE model $(M = 4.13, SD = 1.27)$ and the PVR model $(M = 4.17, SD = 1.61)$. No statistically significant difference was found between the FRA $(M = 2.35, SD = 1.54)$, AD $(M = 2.38, SD = 0.90)$ and PI $(M = 2.37, SD = 1.93)$ models, with the lowest mean model error. The rest of pairwise interactions are contained in Table 4.2. In this trial, models based on egocentric reference frames (FRE, FDE, PV and PVR) had a statistically higher model error, indicating that most participants did not reorient to the initial facing direction during the learning stage. Also, the models with statistically significant lowest distance error (FRA and AD) were also the models with highest number of



a) Trial 1 - Learning stage b) Trial 1 - Placing stage

Figure 4.5: Trial 1 result plot for Participant 8 showing learning and placement stage navigation data (red path) as well as object placement response location (red circle) and all predicted locations (see legend). In this particular example the participant's reponse falls 0.03cm away from the FRA prediction. The amber rectangle represents the two rows of amber LED lights suspended over the room.

participant responses meeting the selection criteria (see Table 4.1). Note that the location predicted by the PI model fell outside the room, making this model null for this trial.

For Trial 2, there was a statistically significant difference in model error for the PI model $(M = 2.99, SD = 0.91)$, with the highest mean model error, and the rest of the models. This is due to the fact that the location predicted by the PI model fell outside the room, making this model null for this trial. There was also a statistically significant difference in model error for the FRA model $(M = 0.87, SD = 0.51)$, with the lowest mean model error, and the rest of the models, but no significant difference with the FRE model $(M = 1.07, SD = 0.60)$. No statistically significant difference was found between the AD $(M = 0.95, SD =)$, FDA $(M = 0.99, SD = 0.45)$ and FRE $(M = 1.07, SD = 0.60)$ models. The rest of pairwise interactions are contained in Table 4.3. Results indicate that this trial triggered object location memory based on the geometry of the envionment (FRE, FRA, FDE, FDA, AD), rather than self motion models (PV and PVR).

Overall, results highlight strong differences in participant behaviour, with contrasting models describing different response object locations. No single model can account for all participant behaviour: models based on spatial geometry seem

Table 4.1: Participant responses that met each of the selection criteria for each model in each of the trials and overall (both trials combined). The last row indicates the total number of participant responses (N). Note that model PI is a null model (the predicted object location falls outside the boundaries of the room) and therefore does not belong to any of the four valid quadrants. Both the QC and the DC are not applicable in this case.

	QC			**DC**			**QC ∪ DC**		
	Trial 1	Trial 2	Overall	Trial 1	Trial 2	Overall	Trial 1	Trial 2	Overall
FRA	15	11	26	7	1	8	15	11	26
FRE	2	7	9	2	1	3	2	7	9
FDA	15	11	26	12	0	12	15	11	26
FDE	2	7	9	2	0	2	2	7	9
AD	15	12	27	0	8	8	15	12	27
PI	-	-	-	-	-	-	-	-	-
PV	9	7	16	0	6	6	9	9	18
PVR	2	7	9	0	2	2	2	7	9
N	39	37	76	39	37	76	39	37	76

to hold greater accuracy in describing part of our responses, whereas models based on self-motion best portray the rest of responses. However, more data would be needed to understand what makes participants behave in ways best represented by different models. Results also indicate that the light cue, designed to help participants reorient back to the initial facing direction, was not salient enough in Trial 1. This cue might have been more obvious in Trial 2, as participants had become more familar with the VE's features at this point in the experimental task.

Results for the Santa Barbara Sense-of-Direction Scale and the Myers Briggs Personality Test, and their relation with object location memory results are not reported as part of this thesis. Separate statistical analysis showed no effect of gender or age on model error.

4.4 Discussion

The limitations listed in Section 3.4 can be conveniently solved using immersive VR. Using this technology, the experimenter can have a higher level of control over the experimental setup and provide ecologically valid stimuli without the noise that is introduced in the real world. This includes easily and rapidly transforming

Table 4.2: Pairwise comparison t and p values between all models for Trial 1. The last row shows mean model error values for each model (M). Interaction is not significant unless it is explicitly indicated as specified in the legend.

	FRA	**FRE**	**FDA**	**FDE**	**AD**	**PI**	**PV**	**PVR**
FRA		-7.59[a]	-4.46[a]	-8.90[a]	-0.24	0.23	-1.66	-5.76[a]
FRE			-5.71[a]	-19.37[a]	9.81[a]	3.86[a]	6.79[a]	2.33[b]
FDA				-7.11[a]	1.45	0.75	-0.49	-4.08[a]
FDE					11.20[a]	4.83[a]	8.88[a]	7.23[a]
AD						0.36	-2.77[b]	-8.66[a]
PI							-1.756	-3.96[a]
PV								-9.24[a]
M	2.35	4.13	2.63	4.68	2.38	2.37	3.04	4.17

[a] interaction is significant at the 0.001 level (two-tailed)

[b] interaction is significant at the 0.05 level (two-tailed)

Table 4.3: Pairwise comparison t and p values between all models for Trial 2. The last row shows mean model error values for each model (M). Interaction is not significant unless it is explicitly indicated as specified in the legend.

	FRA	**FRE**	**FDA**	**FDE**	**AD**	**PI**	**PV**	**PVR**
FRA		-1.47	-2.55[b]	-2.76[b]	5.60[a]	-16.53[a]	-3.60[a]	-3.95[a]
FRE			0.64[b]	-3.19[b]	0.88	-8.27[a]	-7.58[a]	-9.13[a]
FDA				-1.78	0.71	-15.63[a]	-3.20[b]	-3.60[a]
FDE					2.14[b]	-8.16[a]	-3.58[a]	-4.30[a]
AD						-15.84[a]	-3.14[b]	-3.48[a]
PI							5.34[a]	5.28[a]
PV								-3.89[a]
M	0.87	1.07	0.99	1.21	0.95	2.99	1.54	1.59

[a] interaction is significant at the 0.001 level (two-tailed)

[b] interaction is significant at the 0.05 level (two-tailed)

the environment in a variety of ways such as scale transformations. It also allows to automatically capture participant behaviour, including object placement and navigation data, which are generally more difficult to acquire in the real world and require the use of additional tracking devices and software.

The relevant models decribed and explored in our study represent some individual participant behaviour. Unlike Hartley et al. [5], we did not find strong evidence to suggest that responses maintaining fixed distances from nearby walls were more common after expansions of the room and for objects closer to the boundaries of the room, and fixed ratios between opposing walls were more common after contractions of the environment and for locations closer to the centre of the environment. Our results indicate that object location and participant starting location as well as facing direction have a crucial impact on model error in a VE with a low number of available landmarks. They also illustrate that the change in boundary geometry has a crucial impact on participant object location memory. In this study, the design of Trial 1 triggered reponses that were best modelled by the FRA model, as well as the AD model. However, Trial 2 prompted responses best modelled by models based on the geometry of the environment (FRE, FRA, FDE, FDA, AD). This highlights the need to explore more complex paradigms which combine models based on spatial geometry and models based on self-motion or navigation.

Similar to the results from Chapter 3, we observed a relative success of the AD model, where an object location is represented by the absolute distance to its original location in world coordinates, regardless of any changes to the environment geometry. An alternative hypothesis is that participants were attempting to find a correspondence between the physical testing room and the virtual room. This could have happen when participants entered the testing room and were asked to put on the HMD. However, participants were disoriented before being escorted to the starting location for each trial stage (wearing the HMD as a blindfold). Therefore, this hypothesis would only hold for cases in which participants were exceptional at mentally tracking the disorienting path. This encourages further work to understand to what extent the physical environment in which a VR simulation takes place has an impact on spatial learning and recall of the VE.

Results from this studies combined with its real world counterpart motivate the need for further analysis on behavioural differences in real and virtual spaces, and the factors that affect these differences. Furthermore, the availability of novel consumer VR systems could alleviate the need to build complex real world enviornments to carry out specialist research on spatial cognition once these differences are better understood.

Chapter 5

Experiment: The Effect of Environmental Features, Self-Avatar and Level of Immersion on Object Location Memory in Virtual Environments

In this chapter we present a user study on spatial memory based on the work presented in Chapter 3 and Chapter 4. Using a modified version of the task presented in these chapters, we explore the effect of varied environmental feature fidelity of VEs, the use of self-avatars, and the level of immersion of a system on object location learning and recall. Following a between-subjects experimental design, participants were asked to learn the location of three identical objects by navigating one of the three environments: a physical laboratory or low and high detail VE replicas of this laboratory. Participants who experienced the VEs could use either a HMD or a desktop computer. Half of the participants learning in the HMD and desktop systems were assigned a virtual body. Participants were then asked to place physical versions of the three objects in the physical laboratory in the same configuration. We tracked participant movement, measured object placement, and administered a questionnaire related to aspects of the experience.

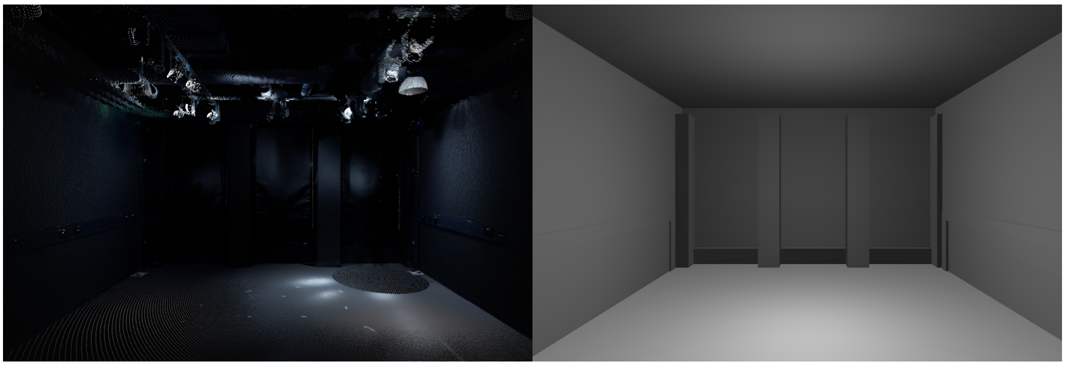


Figure 5.1: Screen captures of the high detail VE (left) and low detail VE (right).

HMD learning resulted in statistically significant higher performance than desktop learning. Results indicate that, when learning in low detail VEs, there is no difference in performance between participants using HMD and desktop systems. Overall, providing the participant with a virtual body had a negative impact on performance. Preliminary inspection of navigation data indicates that spatial learning strategies are different in systems with varying levels of immersion.

5.1 Experimental Design and Hypotheses

In this study we aimed to explore the effect of level of immersion, the presence or absence of a virtual body and the role of environmental features on object location memory. We compared placement accuracy when object locations were learnt in the real world and object locations were learnt in two distinct virtual replicas of the environment: a high detail 3D scan, where colour, environmental and geometric features are available, and a low detail non-photorealistic replica of the shape of the room, where only geometric features were accessible. Participants learnt the position of three identical objects in one of the three environments as shown in Figure 5.1. Once learning was complete and after a short period of time, participants were asked to place the three objects in the real room in their original positions (see Figure 5.2).

Participants observed the VEs and learnt objects positions in different systems following a 2 × 2 × 2 design, with fidelity (high detail, low detail) as a within-subjects factor and avatar (body, no body) and level of immersion (HMD, desktop

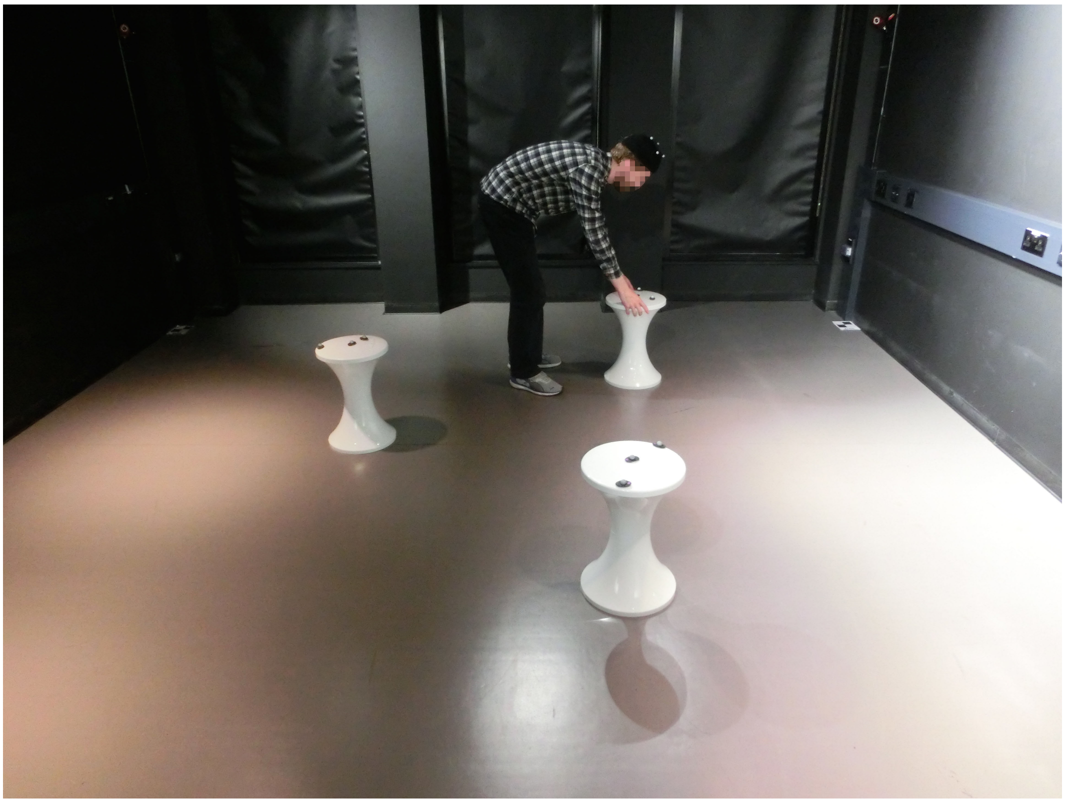


Figure 5.2: Participant placing the three objects in the recall stage. Plastic stools were used as objects for the study. Three retroreflective markers were attached to each stool for optical tracking.

learning system) as between-subjects factors. Real world learning in the real environment with physical objects was treated as an additional learning system. Table 5.1 contains a summary of the mixed design experimental conditions.

Participants in the learning system conditions with a virtual body were assigned a single point tracking avatar model based on head tracking. In other words, a fixed mannequin was placed underneath the participant's head position, with no other reference points or animated movements. Participants learning in the real world and in the HMD learning system conditions were able to explore the space by physically walking around the room. Participants learning in the desktop system condition were able to navigate the room by using keyboard and mouse control, to change position and view, respectively. All participants completed the learning stage in one of the three learning systems and then placed the physical objects in the real world (see Subsection 3.3). In addition to the between-subjects

Table 5.1: Mixed design experimental conditions.

	Learning System				
	Desktop body	Desktop no body	HMD body	HMD no body	Real world
Low detail VE	Desktop body Low detail VE	Desktop no body Low detail VE	HMD body Low detail VE	HMD no body Low detail VE	–
High detail VE	Desktop body High detail VE	Desktop no body High detail VE	HMD body High detail VE	HMD no body High detail VE	–
Real environment	–	–	–	–	Real world

learning system variable, one variable was manipulated within participants learning in the desktop and HMD systems: VE fidelity. Participants in the desktop and HMD learning conditions repeated the task twice, once in the low detail VE and once in the high detail VE. The order in which participants experienced the low detail and high detail VEs was altered, ensuring that the two possible combinations were tested equally. Participants learning in the real world repeated the same task twice, always in the real environment. The dependent variable was placement error, or the absolute distance between participant response and original object position, based on x- and y-coordinates, in meters. We also recorded the navigation paths of all participants when learning and recalling object locations.

We hypothesised that providing optic flow information, natural locomotion, and access to idiothetic cues in a HMD would promote higher similarity with real world learning in terms of placement accuracy and navigation. Previous results have indicated that training in a VEs of relatively low fidelity allows people to develop useful representations of large-scale navigable space [20], contrary to the thought that increasing overall fidelity of a simulator will lead to increases in transfer [99]. Regarding the presence and absence of a single point tracked avatar, we intend to further replicate and verify the results of previous studies in which this type of low motion fidelity virtual body has degraded performance [37]. Because of the availablity of geometric as well as environmental cues, we expected learning in the high detail VE to result in greater accuracy than learning in the low detail VE when placing the objects in their original positions. We predicted that spatial learning and recall in systems with higher level of immersion would result in performance comparable to real world learning.

5.2 Method

5.2.1 Materials

The experiment was conducted in a research laboratory at UCL. The laboratory consisted of a $6m$ long $\times$ $4m$ wide $\times$ $3m$ high open space. The high detail VE was comprised of a high fidelity 3D laser scan point cloud of the room with textures derived from photographs, rendered with a GPU-based point cloud renderer. 3D scanning was performed with a Faro Focus 3D S120 laser scanner. The low detail VE was modeled using diffuse shaded planes to reproduce the geometric shape of the laboratory. Figure 5.1 shows screen captures of the low detail VE, high detail VE and real room from the same viewport. All environments were rendered at scale 1:1 in Unity at $60FPS$ without VSync and a vertical FOV of 60 degrees for the desktop system and $60FPS$ in each eye for the HMD on an Intel Xeon E7 CPU, with 16GB RAM and Nvidia GTX 680 GPU running Windows 7. During the user study the physical room contained a table and a computer that was not included in the scanned virtual 3D model. The table and computer had not been in the room when the 3D scanning took place but were necessary to support the experimental setup.

Head tracking and object positional data was logged with a NaturalPoint OptiTrack motion capture system using twelve Flex 3 cameras and retroreflective markers, at a sampling rate of $60Hz$. The measured mean tracking error was $3mm$. A 27 inch Dell U2713HM monitor and an Oculus Rift Development Kit 2 (DK2) were used as displays for the desktop and HMD learning conditions, respectively. High fidelity single point tracking virtual avatars, based on head tracking, were used in the corresponding desktop body and HMD body conditions. A female and male avatar model were obtained from the Rocketbox® Library [100]. These were preprocessed to remove the heads before being included in the virtual scene. The avatars were not animated and remained in an idle position throughout the task. An Epson EB-585Wi projector was mounted in the ceiling of the laboratory, aligned

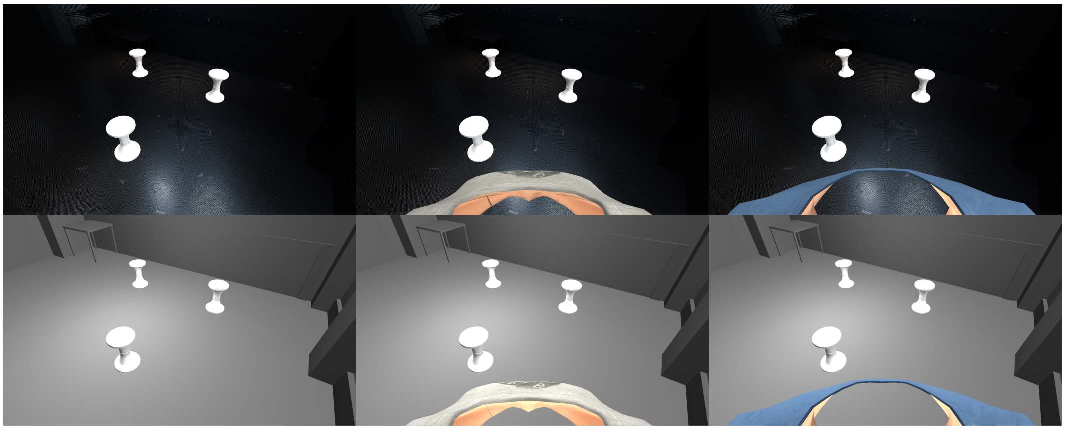


Figure 5.3: Renders of the VE built for the user study. The top row shows the high detail environment with no self-avatar, a female self-avatar and a male self-avatar (from left to right). The bottom row shows the low detail environment with no self-avatar, a female self-avatar and a male self-avatar (from left to right).

with the room and projecting onto the ground. It was used to place the physical versions of the objects in their corresponding positions for real world learning. The objects used in the study were three identical white Tam Tam plastic stools from Habitat, detailed in Section 3.2.1. Figure 5.3 contains renders of the high detail and low detail environments built for the study, with and without the self-avatar representations.

5.2.2 Participants

A total of 20 participants (9 female, 11 male; average age 26 years, $SD = 5.3$) were recruited from the student and staff population at University College London. All participants signed a consent form and the study was approved by the University College London Research Ethics Committee (Project ID: 6708/002). Participants were paid £10 for participation. They were assigned to the different experimental conditions based on individual results for a standard spatial ability test to avoid any possible bias between groups [63].

5.2.3 Procedure

The experimental task consisted of two phases, before and during the lab session. Figure 5.4 shows an overview of the experimental task. Participants performed all their trials in the same learning system condition. Before the lab session,

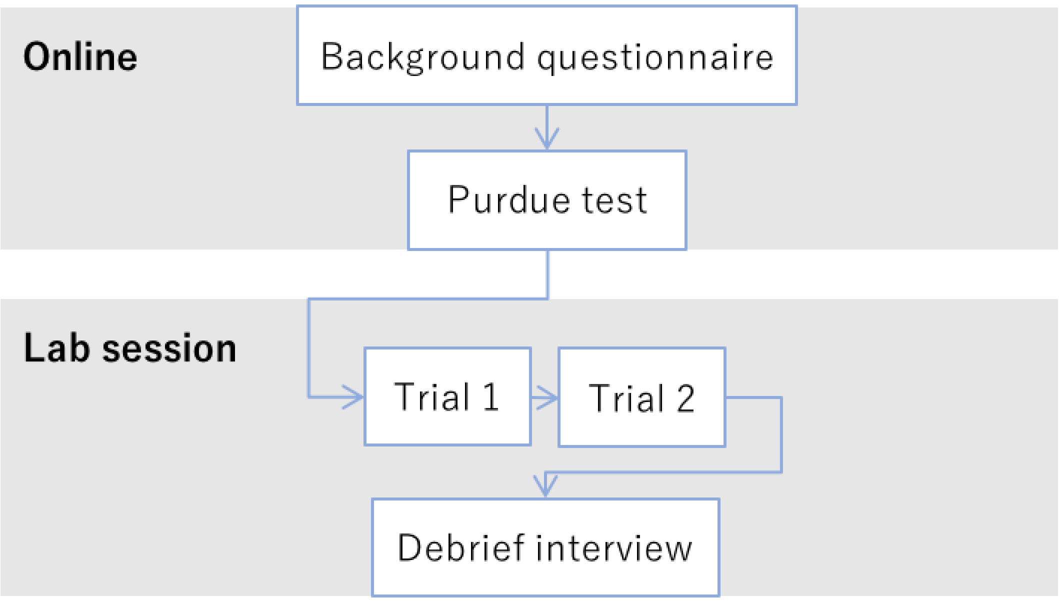


Figure 5.4: Overview of the experimental task.

participants were asked to read an online information sheet that introduced the experimental task. They were asked to read and sign an online informed consent form and asked to complete a digital version of a standard spatial ability test as well as a background questionnaire.

During the lab session, participants were asked to sign a paper copy of the consent form and asked to read an information sheet with written instructions describing the experimental task. Participants were asked to switch off their mobile phones and were introduced in the lab. No practice trials were done and participants were not given feedback on their performance throughout the experiment.

The experimental task consisted of two trials, each with a learning and a recall stage. The learning stage involved viewing the three virtual objects in the real room or one of the low and high detail VEs in one of the three learning system conditions: real world, desktop or HMD. In the recall stage participants were asked to place the three physical objects as they remembered them from the learning stage into the real room. No further information was given and participants were asked to try their best if they were in doubt as to where the object's original position was. There was no time limit for the learning and recall stages, and participants were able to freely navigate the environment. Participants could navigate through all objects of

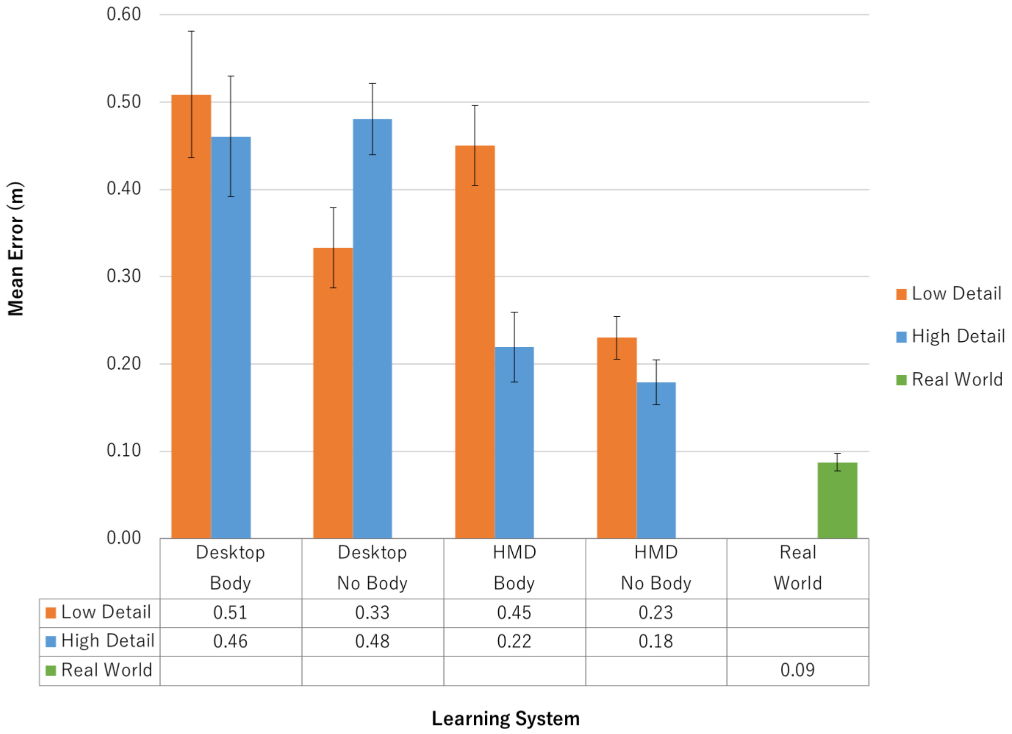
Table 5.2: Post-trial questionnaire related to several aspects of the experience: examination, confidence, difficulty, movement, application and observation. Responses were recorded on a 1-5 Likert scale with varying vocabulary anchoring the low and high ends of the scale, respectively.

Variable	Question	Likert Scale Range
Examination	The learning environment allowed me to closely examine the objects.	1:Poorly - 5:Very Well
Confidence	I am confident that I performed the task well.	1:Unconfident - 5:Confident
Difficulty	The placement task was...	1:Easy - 5:Difficult
Movement	I could move around the learning environment as I wanted.	1:Disagree - 5:Agree
Application	I could directly apply what I learned in the learning environment. when placing the objects in the real room	1:Disagree - 5:Agree
Observation	The learning environment allowed me to naturally observe and learn the object positions.	1:Disagree - 5:Agree

the environment, but not through the environment boundaries. An experimenter was present at all times during the experimental task to manage cables and provide guidance on the different experimental stages.

Participants learning in the HMD and desktop learning systems (16 participants) performed the two trials, each corresponding to one of the two versions of the VE in the learning stage: high detail and low detail. Participants experienced the two VEs in different orders, ensuring that the two possible combinations were tested equally. Participants learning in the real world (4 participants) performed the same trial twice, always learning in the real room. In each trial, and for each participant, all three objects were randomly arranged on a conceptual 5×5 grid, avoiding straight line configurations. Participants could not see the grid in the environment and were asked to ignore retroreflective markers on the stools, which were used to track and identify the stools for data collection.

After each trial, participants were asked to complete a short online questionnaire measuring examination, confidence, difficulty, movement, application and observation (see Table 5.2). After the two trials were completed, they were interviewed regarding individual strategies used throughout the experimental task.



	Desktop Body	Desktop No Body	HMD Body	HMD No Body	Real World
🟧 Low Detail	0.51	0.33	0.45	0.23	
🟦 High Detail	0.46	0.48	0.22	0.18	
🟩 Real World					0.09

Figure 5.5: Mean placement errors in all learning system conditions for real world (green), high detail (blue) and low detail (orange) VEs in m. Error bars show standard errors.

5.3 Results

5.3.1 Object Placement

Tracked object placement data was used to calculate the Euclidean distance, referred to as placement error, between object positions as placed by participants in the recall stage and original object positions. Figure 5.5 shows mean placement errors for all learning system conditions. For statistical analysis the mean placement error was calculated from the error of each of the three objects for all trials.

A three-way mixed Analysis of Variance (ANOVA) with fidelity (high detail, low detail) as a within-subjects factor and avatar (body, no body) and level of immersion (HMD, desktop learning system) as between-subjects factors was run. There were no outliers in the data, as assessed by inspection of a boxplot. There was homogeneity of variances for both high detail placement errors $(p = .257)$ and low detail placement errors $(p = .143)$, as assessed by Levene's test for equality of variances. Results showed a statistically significant two-way interaction between fidelity and level of immersion, $F(1,12) = 6.3, p = .027$, and fidelity and avatar, $F(1,12) = 6.3, p = .027$. Separate statistical analysis showed no effect of gender, age or other background information such as videogame experience on placement error.

Statistical significance of simple main effects was accepted at a Bonferroni-adjusted alpha level of .025. There was a statistically significant simple main effect of avatar for the low detail environment, $F(1,12) = 6.453, p = .026$, but not for the high detail environment, $F(1,12) = .017, p = .899$. All pairwise comparisons were performed for statistically significant simple main effects. Bonferroni corrections were made with comparisons within each simple main effect considered a family of comparisons. Adjusted p-values are reported. Mean placement error was lower when an avatar was present than when an avatar was absent when learning in the low detail environment, with a mean difference of -0.20 $(95\%CI, -0.372$ to $-0.028), p = .026$. There was a statistically significant simple main effect of learning system for the high detail environment, $F(1,12) = 16.423, p = .002$, but not for the low detail environment, $F(1,12) = 1.098, p = .315$. Mean placement error was lower when learning with an HMD system than with a desktop system, when learning in the high detail environment, with a mean difference of -0.083 $(95\%CI, -0.254$ to $-0.089), p = .002$.

A Kruskal-Wallis H test showed that there was an overall statistically significant difference in placement error between the different learning systems, $\chi 2(2) = 56.452, p < .001$, with a mean rank placement error score of 84.15 for desktop learning, 57.53 for HMD learning and 19.15 for Real World learning. When comparing the three system conditions, Real World learning resulted in statistically significant lower placement error $(M = 0.09, SD = 0.04)$, followed by HMD learning $(M = 0.27, SD = 0.16)$ and Desktop learning $(M = 0.45, SD = 0.21)$, respectively. No statistically significant differences were found between the two trials.

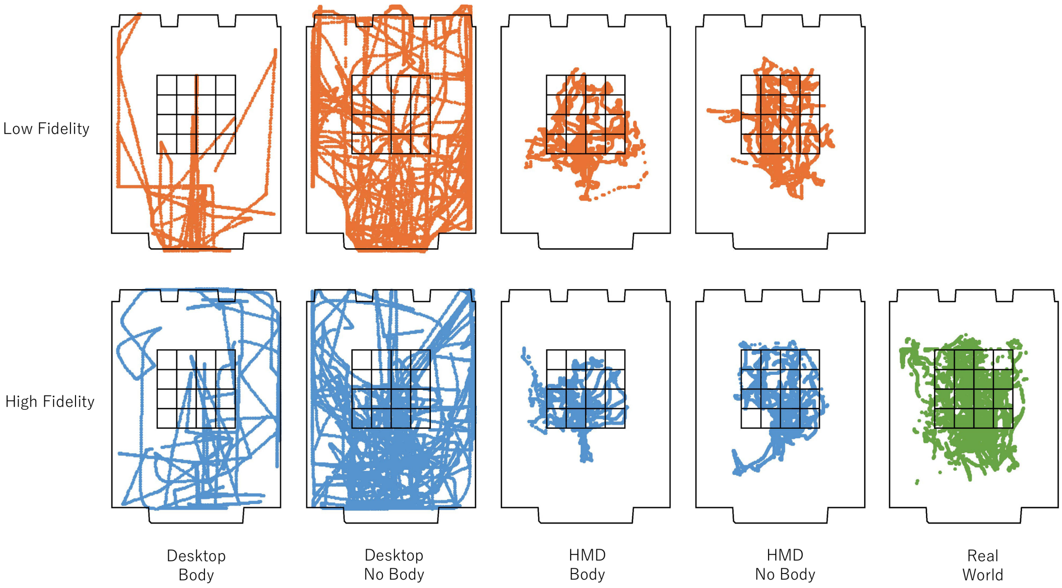


Figure 5.6: Learning stage 2D (XY plane) tracked navigation trajectories for all participants in each learning condition for Real World (green), High Detail (blue) and Low Detail (orange) VEs. Each point represents an XY position at a sampling rate of $60Hz$. All interesections in the 5×5 conceptual grid represent a possible object position. The conceptual 5×5 object grid was invisible to participants.

5.3.2 Questionnaire

A one-way between subjects ANOVA was performed on questionnaire responses for desktop body, desktop no body, HMD body and HMD no body learning system conditions, for high and low detail VEs. Results show a large number of mixed significant interactions with no overarching trend due to the limited number of repetitions.

5.3.3 Navigation

Tracking results, shown in Figure 5.6, indicate contrasting movement patterns in Real World, HMD and Desktop learning system conditions. Qualitative inspection of data suggests that participants learning in the real world and HMD systems primarily navigated areas within the boundaries of the conceptual 5×5 object grid whereas participants learning in the desktop computer mainly navigated areas outside the boundaries of the conceptual 5×5 object grid. The mean percentage of time spent navigating inside and ouside the conceptual 5×5 object grid was calculated for each learning system and is shown in Figure 5.7.

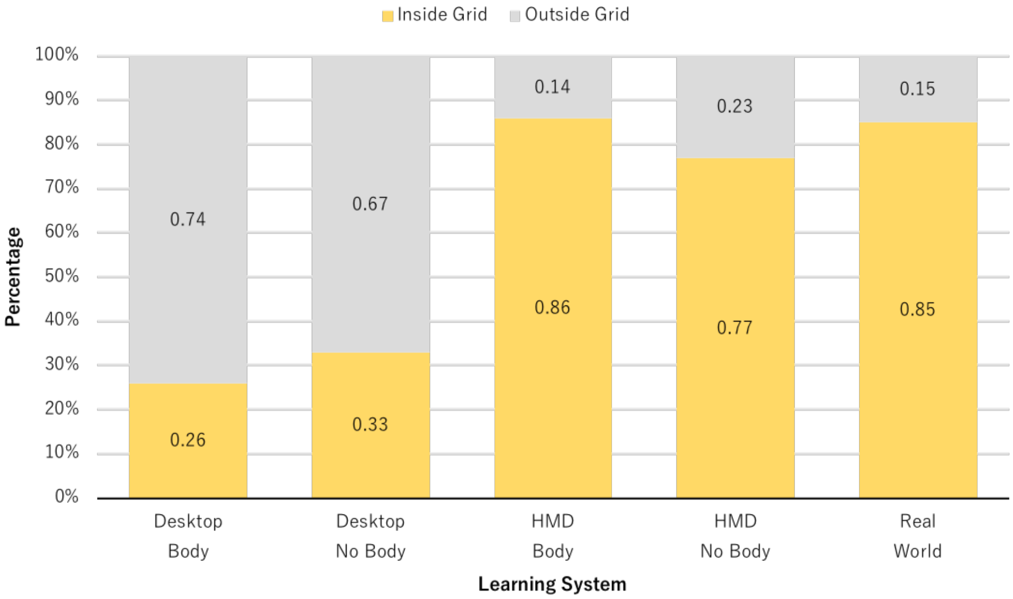


Figure 5.7: Mean percentage of navigation time spent outside (grey) and inside (yellow) the conceptual 5×5 object grid for all participants in each learning condition during the learning stage.

A one-way between subjects ANOVA was conducted to compare the effect of learning system on the percentage of time spent navigating inside the conceptual 5×5 object grid in desktop, HMD and real world learning system conditions. There was a significant effect of learning system on percentage time spent navigating inside the conceptual 5×5 object grid at the $p < .05$ level $[F(2,39) = 371.991, p < .001]$. Post hoc comparisons using the Tukey HSD test indicated that the mean percentage spent navigating inside the conceptual 5×5 object grid for desktop learning $(M = 0.24, SD = 0.06)$ was significantly lower than the mean percentage spent navigating inside the conceptual 5×5 object grid for HMD learning $(M = 0.80, SD = 0.08)$ and real world learning $(M = 0.78, SD = 0.04)$. No significant difference was found between HMD and real world learning.

To further illustrate differences in navigation strategies, we created cluster heat maps of the time spent in each region of the room for each of the system conditions: Desktop (left), HMD (middle) and Real World (right), shown in Figure 5.8. These results show different spatial navigation strategies between desktop and HMD learning strategies, where the former tended to access areas towards the far end

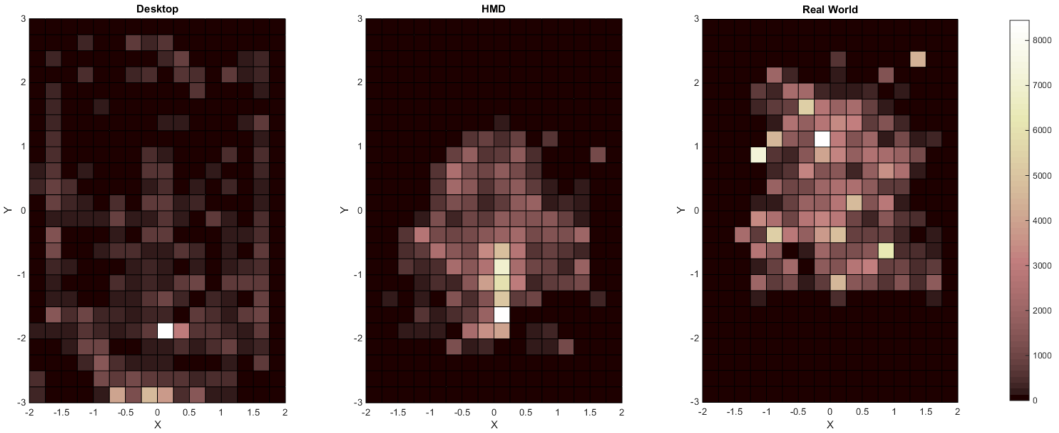


Figure 5.8: Cluster heat maps of the time spent in each region of the room for each of the system conditions: Desktop (left), HMD (middle) and real world (right). The columns represent X-axis and the rows represent Y-axis positions (in m). Each cell is colorized based on the level of counts of the head location (for all participants) in each region during learning.

of the room and the latter tended to navigate areas clustered within the object grid along the Y axis. Along the X axis, the range of positions accessed by participants learning in the desktop system was wider than the range of movement performed by participants learning in the HMD system. HMD navigation was not only different from desktop navigation, but also qualitatively very similar to real world navigation during learning.

5.4 Discussion

This study analyses object location memory transfer from VR to the real world. It extends previous work on spatial perception in VEs [101, 102, 103, 104, 38, 26, 32, 34, 105, 20] by suggesting an experimental task in which participants are asked to learn and recall a series of object configurations in concurrently occupied virtual and real environments.

Our results illustrate that HMD learning resulted in statistically significant higher performance followed by desktop learning. Our analysis suggests that availability of environmental features in VEs can enhance object location memory under certain setups. The overall negative effect of the self-avatar indicates that single point tracked virtual bodies may not be sufficient to increase performance in this experimental task. Specifically, the use of self-avatar in HMD body learning impaired placement accuracy. Single point tracking caused the virtual self-avatar to appear in front of the participant's real body if they leaned forward, partially occluding some of the available environmental features. The degradation in performance might have been because the virtual body occluded features in the environment that the participant could have attended to. This might then have forced a change to a different strategy for learning one or more object placements. Moreover, the lack of motion fidelity provided by single point virtual bodies might interfere with presence in VEs.

The results on navigation strategies seem promising. Similar to participants learning in the real world, participants learning in the HMD system mainly navigated areas within the boundaries of the conceptual 5×5 object grid, whereas participants learning in the desktop system primarily explored areas outside the boundaries of the conceptual 5×5 object grid. This may suggest that, when learning object locations in less immersive systems, users navigate towards the environment boundaries to obtain more global views of the scene. In addition, the range of areas of the room accessed by participants learning in the desktop system was wider than the range of areas of the room participants learning in the real world and HMD system in the X and Y axis. Although differences in navigation in systems with varying levels of immersion have been reported [106], further exploration is required to understand the trajectories selected by users when learning object locations.

One of the limitations of the work presented here is the relatively low number of participants. A larger population sample is needed to further validate our results as well as to explore the effect of more complex self-avatars with higher motion fidelity on spatial memory. It would also allow us to analyse navigation trajectories in more detail, exploring the regions visited by participants in relation to the object locations and features of the environment. Other experimental tasks comparing object location memory in systems with varying levels of immersion are required to confirm whether our results are generalisable.

Chapter 6

Experiment: A Comparison of Virtual and Physical Training Transfer of Bimanual Assembly Tasks

In this chapter we present a study that explores the effect of level of immersion on training of a more complex procedural task, compared with the studies presented in the previous chapter which looked at object location memory. For this, we compare the effectiveness of virtual training and physical training for teaching a bimanual assembly task. In a between-subjects experiment, 60 participants were trained to solve three 3D burr puzzles in one of six conditions comprised of virtual and physical training elements. In the four physical conditions, training was delivered via paper- and video-based instructions, with or without the physical puzzles to practice with. In the two virtual conditions, participants learnt to assemble the puzzles in an interactive VE, with or without 3D animations showing the assembly process. After training, we conducted immediate tests in which participants were asked to solve a physical version of the puzzles. We measured performance through success rates and assembly completion testing times. We also measured training times as well as subjective ratings on several aspects of the experience. Our results show that the performance of virtually trained participants was promising.

Table 6.1: Experimental condition types, acronyms and definitions. Please note the choice of acronym V_I to represent *video* and V_E to represent *virtual environment* to avoid any confusion in making reference to the experimental conditions throughout the thesis.

Type	**Acronym**	**Definition**
Physical	P	Paper instructions
	PB	Paper instructions and physical blocks
	PV_I	Paper instructions and assembly process video
	PV_IB	Paper instructions, assembly process video and physical blocks
Virtual	V_E	Virtual paper instructions and virtual blocks
	V_EA	Virtual paper instructions and virtual blocks, with assembly process animations

Figure 6.1: One of the three 3D printed burr puzzles used in the study.

A statistically significant difference was not found between virtual training with animated instructions and the best performing physical condition (in which physical blocks were available during training) for the last and most complex puzzle in terms of success rates and testing times. Performance in retention tests two weeks after training was generally not as good as expected for all experimental conditions. We discuss the implications of the results and highlight the validity of virtual reality systems in training.

6.1 Experimental Design and Hypotheses

Inspired by previous research [2], in our study we used three different colour-coded versions of a six-piece burr puzzle for the assembly task (see Figure 6.1). Burr puzzles have been commonly used for assembly task training studies in the past because they provide a recognisable and adequately complex model in which participants must follow a specific procedure in order to solve them [59, 2].

Table 6.2: Classification of the experimental conditions according to instruction type (static or static and animated) and block availability (no blocks, physical blocks or virtual blocks) during training. See Table 6.1 for experimental condition types, acronyms and definitions.

	Physical		**Virtual**
	No blocks	Physical blocks	Virtual blocks
Static instructions	P	PB	Virtual Environment (V_E)
Static and animated instructions	PV_I	PV_IB	Virtual Environment with Animations (V_EA)

However, our study differs from previous work in that no haptic devices were used. In addition, we are interested in whether consumer virtual reality systems are sufficient for effective training.

In our study, participants were trained and tested in assembling three versions of a six-piece burr puzzle. To provide increasing difficulty, the first three blocks had been preassembled for the first puzzle, the first two for the second and none for the third. This meant that participants had to remember a higher number of steps in the assembly process over the course of the experimental task for each puzzle.

Following a between-subjects experimental design, participants were trained to solve each puzzle by adding the corresponding unassembled blocks in one of six experimental conditions (see Table 6.1). Experimental conditions were designed to account for scenarios in which blocks are not available (P and PV_I), physical blocks are available (PB and PV_IB) or virtual blocks are available (V_E and V_EA) during training (see Table 6.2 for a classification of the experimental conditions). The physical experimental conditions (P, PB, PV_I and PV_IB) were designed to encompass combinations of paper- and video- based instructions. The virtual experimental conditions (V_E and V_EA) involved a virtual version of the paper instructions, with or without 3D animations showing how to correctly assemble the puzzle, and always with virtual blocks to practice during training. All instructions (static and animated) were colour-coded to match the physical puzzle blocks.

Following training and after a short break, participants were asked to assemble a 3D printed physical version of the corresponding puzzle within a given time. Participants were asked to attend a retention session, two weeks after the training, in which they were asked to solve the same puzzles in the same order and within the same time constraints. We measured success rates as well as training and testing times. Sessions were complemented by a series of mental rotations tests as well as questionnaires and debrief interviews.

As part of their recommendations for future work, Carlson et al. suggested adding a snap-to-fit function or constraint system [107] to alleviate the time that virtually trained participants spent attempting to fit and assemble the virtual blocks [2]. We followed this recommendation and added such functionality in the virtual training environment. We also followed their recommendation to make the selection of a block in the VEs cause a change of colour instead of just causing a change in transparency, as participants in their study reported that it was difficult to discern transparent pieces against the transparent virtual representation of the glove. In their discussion they mentioned individual differences for interaction between the two hands, as some participants showed a preference for the haptic device or the glove for predominant use. We therefore decided to make interaction ambidextrous, meaning all operations were designed to be performed equally by both hands. We made the following hypotheses:

H1: The conditions in which the physical blocks were available during training (PB and PV_IB) would yield a higher number of successful puzzle completions during immediate and retention testing. This relates to the experience (or lack of) built around manipulating and assembling the physical blocks during training.

H2: The conditions in which static and animated instructions (video or 3D animations) were available during training (PV_I, PV_IB and V_EA) would result in lower assembly times during immediate and retention testing, as participants would have received richer visualisation on how to assemble the blocks during training.

H3: Condition PV_IB, with physical blocks and animated instructions (video), would yield the best performance as measured by immediate and retention success rates and assembly testing times. This hypothesis is based on H1 and H2.

Figure 6.2: Physical lab where the experiment took place (left) and analogous VEs (right).

6.2 Method

6.2.1 Materials

The user study was conducted in a lab at UCL. The room consisted of a 3.1m long × 2.7m wide × 4.0m high room. A virtual replica of the laboratory was modeled for the VE used in the virtual experimental conditions. Figure 6.2 contains images of the physical room and analogous VE. An Oculus Rift Consumer Version 1, two Oculus Touch controllers and two Oculus sensors were used for the virtual experimental conditions. The VEs was rendered at scale 1:1 in Unity 5.6.0 without VSync at 90FPS in each eye on an Intel Core i7-4770K CPU @ 3.50GHz, with 16GB RAM and Nvidia GeForce GTX 1080 GPU running Windows 8.1 Pro. The Oculus Avatar SDK 1.15.0 [108] was used to include hand presence and interaction for the Oculus Touch controllers. The Burr Tools 0.6.3 software was used to digitally create and solve the three versions of the six-piece burr puzzles as well as to generate the paper instructions and assembly process videos [109]. The puzzle blocks were 3D printed with a Ultimaker 2+ 3D printer with a 0.4mm nozzle and standard settings, with PLA 3D printing material. 3D models of the burr puzzles used in the study are available to download at `https://vr.cs.ucl.ac.uk/research/virtual-training`. Preassembled blocks for the first and second puzzles were glued together. Paper instructions were printed on A3 paper and attached to 5mm A3 foamboards. Assembly videos were presented using VLC 2.2.3 on a 13-inch mid 2014 MacBook Pro laptop running macOS 10.12.2.

6.2.2 Participants

A total of 60 participants (30 female, 30 male; average age 26.51 years, ($SD = 6.47$)) were recruited from the student and staff population at UCL. All participants signed a consent form and the study was approved by the UCL Research Ethics Committee (Project ID: 6708/004). Participants were paid £15 for participation.

A screener questionnaire was used to filter out potential participants who enjoy solving 3D puzzles or who have any type of colour-blindness. Eligible participants were assigned to the different experimental conditions based on individual results for Purdue's Visualisation of Rotations Test [63] to avoid any possible bias between groups, with a similar mean score for the test in each of the experimental condition groups. Likewise, an equal number of females and males were assigned to each group.

6.2.3 Physical Training Environment

Participants assigned to the physical experimental conditions (P, PB, $\mathrm{PV_I}$ and $\mathrm{PV_IB}$) were seated on a stool in front of the table in the lab on which the blocks had been placed in the correct initial configuration for each puzzle. Participants were seated facing the table and were told that they could adjust the distance to it if they wished to.

Paper instructions were designed to show the initial configuration of the blocks at the top and the assembly process steps at the bottom (see Figure 6.3). For the first two puzzles, blocks that had been preassembled and the corresponding steps in the assembly process were faded out. The orientation of the images of the blocks in the instructions was randomly selected for each puzzle. For those experimental conditions involving paper instructions, these were placed against the wall on the table in front of the participant. Assembly process videos were generated using Burr Tools [109] and showed a step-by-step animation of the assembly process from the perspective matching the one in the paper instructions. The laptop was placed on the table in front of the participant. Participants could interact with the video (play, pause, stop, rewind, and fast forward) using the VLC user interface.

For those experimental conditions in which the physical blocks were available during training (PB and PV_IB) these were initially placed on the table following the same configuration as the paper instructions. Preassembled puzzles were placed behind the blocks.

6.2.4 Virtual Training Environment

Participants assigned to the virtual experimental conditions (V_E and V_EA) were seated on a stool in the center of the lab. They were then asked to put on the Oculus Rift and hold the two Oculus Touch controllers with the experimenter's help. The VE showed the virtual replica of the room and table used in the physical environment in front of them, with the blocks for the corresponding puzzle arranged in the correct configuration. Participants were seated facing the virtual table and were told that they could adjust the distance to it if they wished to. For the first two puzzles (in which two or three of the blocks had been preassembled) participants could see the preassembled puzzle hovering over the table in front of them. Virtual paper instructions were presented against the wall on the table in the same location as the physical paper instructions were presented in the physical training environment.

Using the Oculus Avatar SDK 1.15.0 [108], virtual hands were rendered using the default shader (see Figure 6.4). Participants could then manipulate the 3D environment by grabbing the virtual puzzle blocks. They could hold the trigger button to grab unassembled puzzle blocks and the grip button to move and rotate assembled blocks as a single unit. Participants could grab any block at any given time, but only the correct block in the assembly process could be attached to the puzzle. No physics constraints were added to the blocks meaning they could be moved through each other and through the virtual hands and table.

Visual feedback was provided to aid participants in learning the assembly process during training. When participants grabbed the correct block in the assembly process, a blue transparent preview block was shown in the puzzle,

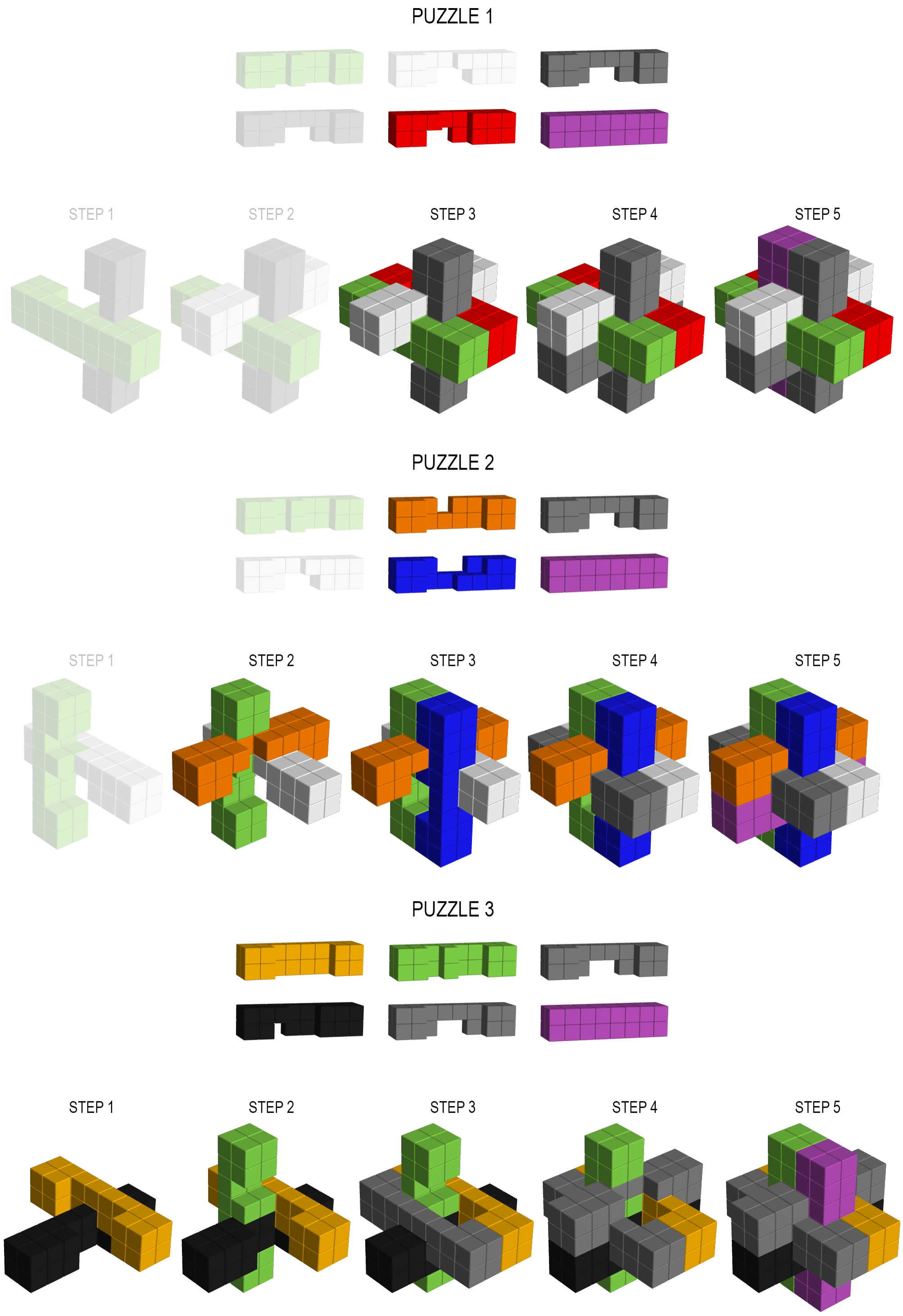


Figure 6.3: Assembly instruction sheet for each of the three burr puzzles used in the study. Each instruction sheet contains a diagram of the six pieces and five ordered steps needed to solve the puzzle. Preassembled pieces and steps for Puzzles 1 and 2 were faded out.

indicating where the block had to be assembled. Participants had the option to deactivate the block preview. A blue highlight was used to indicate what the next block in the assembly process was. This highlight would then turn to red when the block collided with the preview block, indicating that the piece was near its correct location but in the wrong orientation. The highlight would turn green when the block was within an angle of twenty degrees from the correct orientation. If the participant released the trigger when the block showed a green highlight, it would snap into the correct location and the participant could move on to assemble the next piece or reset the puzzle. No audio or vibration feedback was used in the experience.

A user interface with virtual buttons was added on the right-hand side of the virtual table. Buttons were represented by blue spheres which the participant could interact with by touching them, after which they would turn to grey and back to blue to indicate that the interaction was successful. For participants in the V_E and V_EA conditions, two buttons were available: RESET and HELP ON/OFF. Interacting with the RESET button would immediately relocate all blocks in their initial positions so participants could restart the assembly process whenever they wished. The HELP ON/OFF button acted as a toggle to activate and deactivate the blue transparent preview of the block in the puzzle so participants could practice assembling the puzzle with and without the visual aid.

For participants in the V_EA condition, two more buttons were added: NEXT STEP and REPLAY LAST STEP. The NEXT STEP button would trigger the animation of the assembly of the next block in the process. The REPLAY LAST STEP would reposition the last block assembled in its original location on the table and animate its assembly onto the puzzle.

All interactions in the virtual training environment could be equally carried out using either hand and participants could concurrently complete one interaction with each hand. For example, a participant could grab and rotate the assembled pieces with one hand and grab the next block to attach with the other hand.

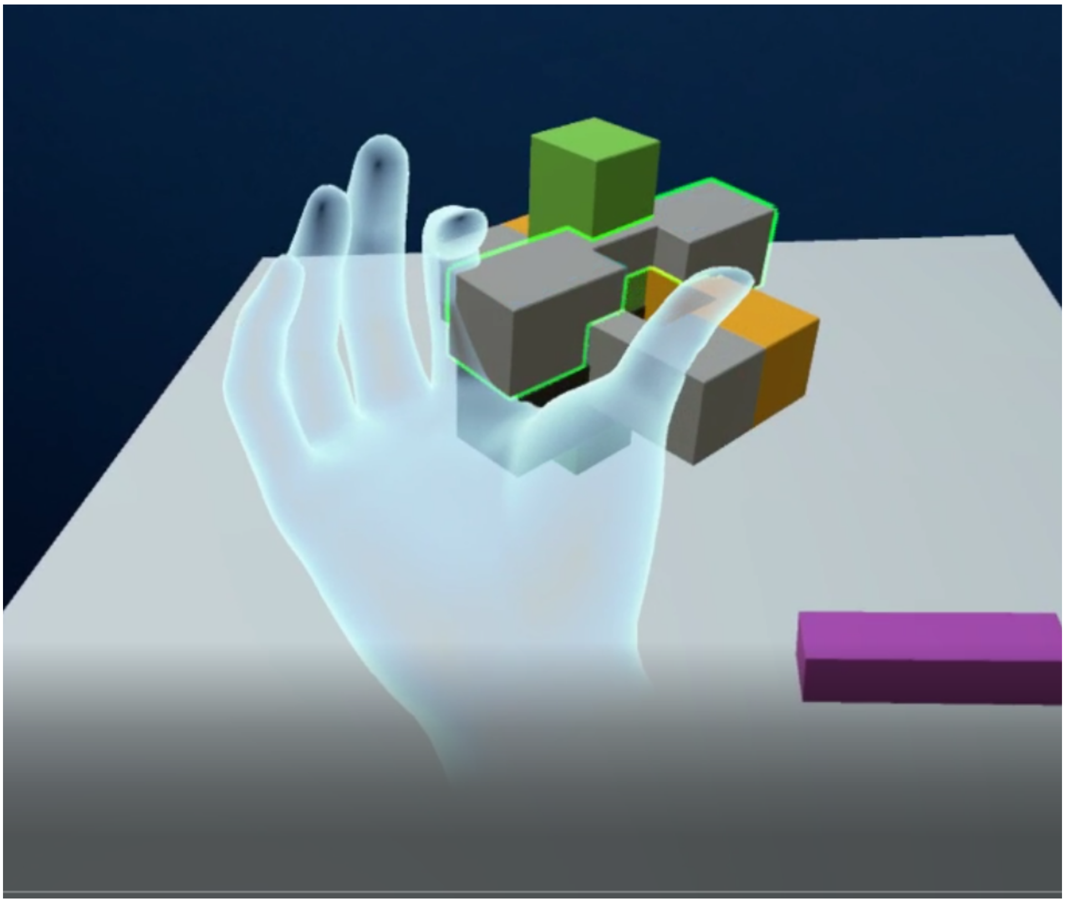


Figure 6.4: Screenshot of a participant grabbing a virtual block and assembling it onto the 3D puzzle. The green highlight indicates on the block is colliding with its preview block and within twenty degrees from the correct orientation. By releasing the trigger button of the Oculus Touch controller the virtual block would snap into its correct location.

6.2.5 Procedure

The experimental task consisted of two lab sessions. The first session comprised training and immediate testing. The second session, two weeks after the first, comprised retention testing. Figure 6.5 shows an outline of the experimental task. Before the first lab session, participants were asked to read and sign an online informed consent form and answer a digital version of Purdue's Visualisation of Rotations Test [63] used to pre-allocate participants to the experimental conditions. Participants also answered a background questionnaire with a specific focus on prior experience with videogames, 3D modelling software and VEs.

During the first lab session, participants were asked to sign a paper copy of the consent form and asked to read an information sheet with written instructions describing the experimental task. In this session participants completed a familiarisation task and three trials, each with a training and a testing stage. The three trials corresponded with each of the three burr puzzles in increasing order of difficulty. During the familiarisation task participants were introduced to the physical or virtual training environment depending on the experimental condition they had been assigned to. A sample assembly task involving piling up rectangular blocks was used and participants were able to familiarise themselves with the paper instruction format, the video player and the interactive VEs, accordingly.

For each of the trials, the training stage involved learning to assemble the corresponding puzzle in one of the six experimental conditions in a maximum time of eight minutes. During the testing stage participants were asked to assemble the physical 3D puzzle in a maximum of three minutes. Time limits for training and testing were defined through piloting of the experimental task. In each trial participants completed the training stage and, after a thirty second break, the testing stage. They then completed a questionnaire at the end of each trial (see Table 6.3). For both training and testing participants were told what the limit times were and were advised that they could end the stage before the time expired if they wished to. Participants were also told that the initial configuration of the blocks on the table during training would match the initial configuration of the blocks during testing and the paper instructions.

Participants were asked to try their best if they were in doubt as to how to assemble the puzzles during testing. An experimenter was present at all times during the experimental task to manage cables for those particpiants in the virtual experimental conditions and provide guidance on the different phases of the experimental task. After completing all trials participants were interviewed regarding the strategies used throughout the sessions.

After a waiting period of two weeks, participants returned to the lab for the second session. In this session participants were asked to complete a paper version

Table 6.3: Post-trial questionnaire related to several aspects of the experience: difficulty, ease of use and seriousness. Responses were recorded on a 1-5 Likert scale with varying vocabulary anchoring the low and high ends of the scale, respectively.

Variable	Question	Likert Scale Range
Difficulty	Please rate the difficulty of the task you just completed.	1: Very difficult - 5: Very easy
Ease of use	Please rate the ease of use in assembling parts in the training environment.	1: Very difficult - 5: Very easy
Seriousness	Please rate how seriously you took the task.	1: Very unseriously - 5: Very seriously

of the Vandenberg and Kuse Mental Rotations Test [65]. They then completed the retention test for each of the three puzzles, in which they were asked to solve the three burr puzzles from the first session without a training phase, in the same order and in a maximum of three minutes. They completed the same questionnaire from the first session at the end of each retention trial (see Table 6.3). After completing all retention trials they were interviewed regarding strategies used throughout the session.

6.3 Results

6.3.1 Types of Errors

Unsuccessful puzzle completions during immediate and retention testing were due to one of two reasons. In most cases, participants did not complete the 3D puzzles within the given maximum time (180s). On the other hand, a low number of participants decided to stop the time before the upper limit thinking that they had successfully solved the puzzle. However, close inspection showed that they had not correctly assembled the pieces. Completion time values for both immediate and retention testing were corrected by assigning the upper time limit (180s) to all unsuccessful attempts.

6.3.2 Immediate Testing

6.3.2.1 Training times

Boxplots with training times for each of the puzzles are shown in Figure 6.6. Non-parametric statistical analysis was performed for training times because our data was not normally distributed as shown by a Shapiro-Wilk test.

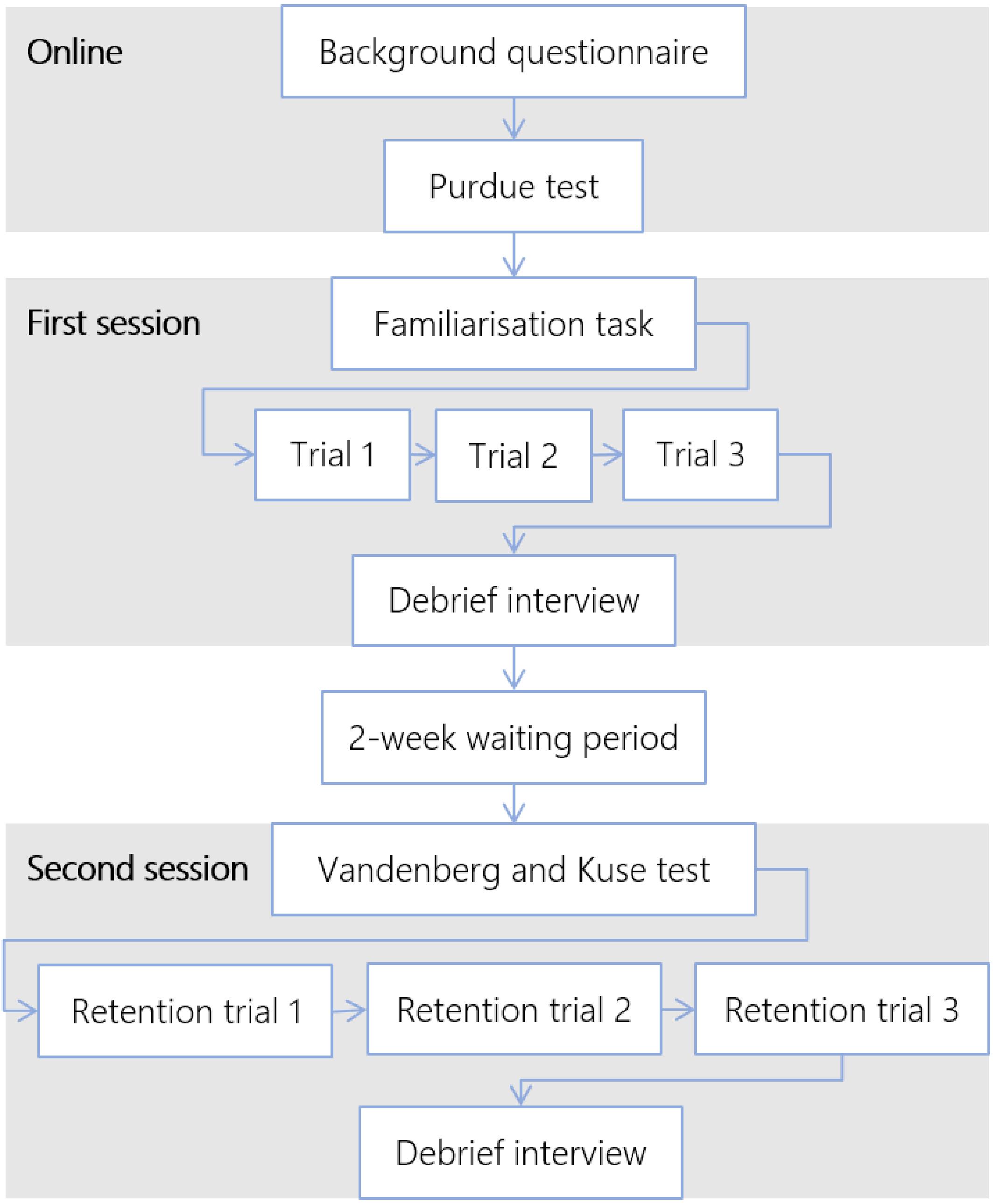


Figure 6.5: Overview of the experimental procedure.

A Kruskal-Wallis H test showed that there was an overall statistically significant difference in training times for the first puzzle between the different experimental conditions, $\chi2(5) = 25.648, p < 0.001$, with a mean rank score of 15.35 for P, 38.85 for PB, 13.85 for $\mathrm{PV_I}$, 40.15 for $\mathrm{PV_IB}$, 36.25 for $\mathrm{V_E}$ and 38.55 for $\mathrm{V_EA}$.

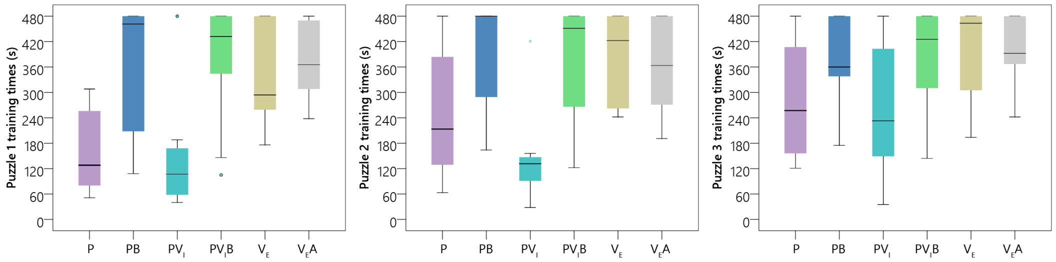


Figure 6.6: Boxplot containing training times for each of the puzzles. Medians are shown as dark horizontal lines. Boxes represent the IQR. Whiskers represent either the extreme data points or extend to 1.5 x IQR. Outliers (data points outside the whiskers) are shown by circles. A value, X, is an outlier if X $<$ lower quartile $-$ 1.5 x interquartile range or if X $<$ upper quartile $+$ 1.5 x IQR. See Table 6.4 for pairwise interactions.

Table 6.4: Test statistics using Dunn's procedure [1] for training times between the different experimental conditions. Significance values have been adjusted by the Bonferroni correction for multiple tests. Interaction is not significant unless it is explicitly indicated as specified in the legend.

		PB	**PV$_{\mathrm{I}}$**	**PV$_{\mathrm{I}}$B**	**VE**	**V$_{\mathrm{E}}$A**
Puzzle 1	P	-23.50[a]	1.50	-24.80[a]	-20.90	-23.20[a]
	PB		25.00[a]	-1.30	2.60	0.30
	PV$_{\mathrm{I}}$			-26.30[a]	-22.40	-24.70[a]
	PV$_{\mathrm{I}}$B				-3.90	-1.60
	VE					-2.30
Puzzle 2	P	-17.50	11.90	-14.20	-14.95	-13.25
	PB		29.40[a]	3.30	2.55	4.25
	PV$_{\mathrm{I}}$			-26.10[a]	-26.85[a]	-25.15[a]
	PV$_{\mathrm{I}}$B				0.75	-0.95
	VE					1.70

[a] interaction is significant at the 0.05 level (two-tailed)

A Kruskal-Wallis H test showed that there was an overall statistically significant difference in training times for the second puzzle between the different experimental conditions, $\chi 2(5) = 22.764, p < 0.001$, with a mean rank score of 22.50 for P, 40.00 for PB, 10.60 for PV$_{\mathrm{I}}$, 36.70 for PV$_{\mathrm{I}}$B, 37.45 for V$_{\mathrm{E}}$ and 35.75 for V$_{\mathrm{E}}$A.

A Kruskal-Wallis H test showed that there was no overall statistically significant difference in training times for the third puzzle between the different experimental conditions, $\chi 2(5) = 10.701, p = 0.058$, with a mean rank score of 21.95 for P, 33.50 for PB, 18.90 for $\mathrm{PV_I}$, 35.70 for $\mathrm{PV_IB}$, 37.15 for $\mathrm{V_E}$ and 35.80 for $\mathrm{V_EA}$.

Pairwise comparisons were performed using Dunn's procedure [1] with a Bonferroni correction for multiple comparisons with adjusted p-values. These are displayed in Table 6.4. Note that pairwise comparisons for puzzles in which the Kruskal-Wallis H test showed no overall statistically significant difference have not been included.

The post hoc analysis revealed statistically significant differences in training times for the first puzzle. There was a statistically significant difference between P (mean rank = 15.35) and PB (mean rank = 38.85) ($p = 0.036$), $\mathrm{PV_IB}$ (mean rank = 40.15) ($p = 0.020$) and $\mathrm{V_EA}$ (mean rank = 38.55) ($p = 0.041$). There was also a statistically significant difference between $\mathrm{PV_I}$ (mean rank = 13.85) and PB (mean rank = 38.85) ($p = 0.018$), $\mathrm{PV_IB}$ (mean rank = 40.15) ($p = 0.010$) and $\mathrm{V_EA}$ (mean rank = 38.55) ($p = 0.021$).

The post hoc analysis revealed statistically significant differences in training times for the second puzzle. There was a statistically significant difference between $\mathrm{PV_I}$ (mean rank = 10.60) and PB (mean rank = 40.00) ($p = 0.002$), $\mathrm{PV_IB}$ (mean rank = 36.70) ($p = 0.010$), $\mathrm{V_E}$ (mean rank = 37.45) ($p = 0.007$) and $\mathrm{V_EA}$ (mean rank = 35.75) ($p = 0.015$).

Separate statistical analysis showed no significant effect of gender or age on training times for all puzzles.

6.3.2.2 Immediate testing success rates

A binomial logistic regression was performed to ascertain the effects of experimental condition on the likelihood that participants succeed at assembling each puzzle during the immediate testing phase. Figure 6.7 shows the number of successful and unsuccessful completions of each puzzle for all experimental conditions. $\mathrm{PV_IB}$ was chosen as the reference category as this was the condition that produced the highest number of successful puzzle completions, overall.

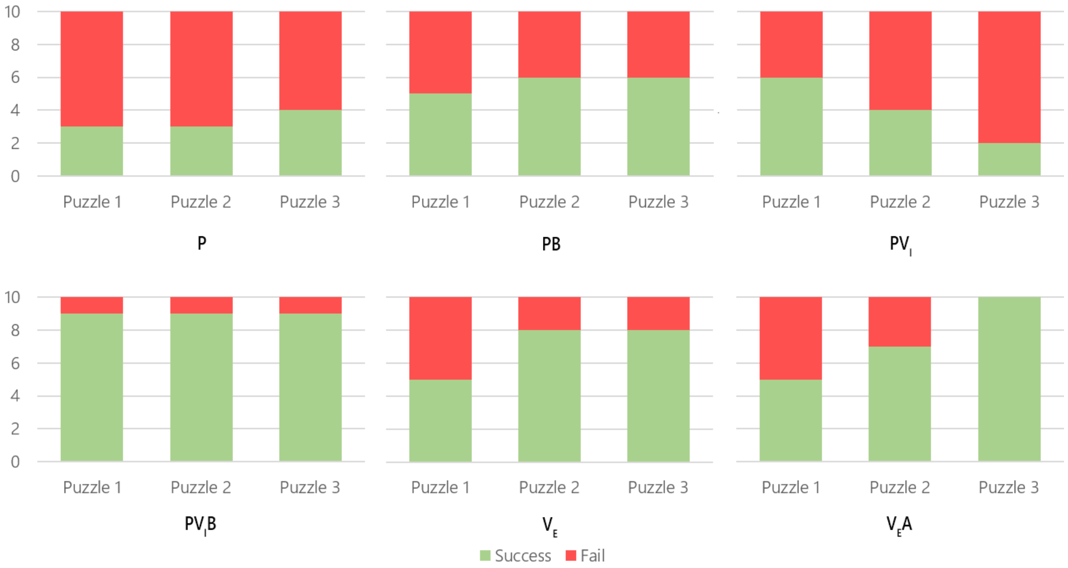


Figure 6.7: Number of successful (green) and failed (red) attempts at solving the three puzzles in the immediate testing phase for each of the experimental conditions.

The binomial logistic regression model was not statistically significant, $\chi2(5) = 8.809, p = 0.117$ for the first puzzle. The model explained 18.3% (Nagelkerke R^2) of the variance in success rate and correctly classified 61.7% of cases. The Wald criterion demonstrated that only condition P made a significant contribution to prediction (p = 0.016). The model suggested that participants in this condition were 0.05 times as likely to successfully assemble the first puzzle than participants in the reference category ($\mathrm{PV_IB}$).

The binomial logistic regression model was statistically significant, $\chi2(5) = 12.016, p = 0.035$ for the second puzzle. The model explained 24.7% (Nagelkerke R^2) of the variance in success rate and correctly classified 71.7% of cases. The Wald criterion demonstrated that P and $\mathrm{PV_I}$ made a significant contribution to prediction ($p = 0.016$ and $p = 0.035$, respectively). The model suggested that participants in the P experimental condition were 0.048 times as likely to successfully assemble the second puzzle than participants in the reference category ($\mathrm{PV_IB}$). The model suggested that participants in the $\mathrm{PV_I}$ experimental condition were 0.074 times as likely to successfully assemble the second puzzle than participants in the reference category ($\mathrm{PV_IB}$).

The binomial logistic regression model was statistically significant, $\chi 2(5) = 24.255, p < 0.001$ for the third puzzle. The model explained 45.8% (Nagelkerke R^2) of the variance in success rate and correctly classified 78.3% of cases. The Wald criterion demonstrated that P and PV_I made a significant contribution to prediction (p = 0.035 and p = 0.007, respectively). The model suggested that participants in the P experimental condition were 0.074 times as likely to successfully assemble the third puzzle than participants in the reference category (PV_IB). The model suggested that participants in the PV_I experimental condition were 0.028 times as likely to successfully assemble the third puzzle than participants in the reference category (PV_IB). Condition V_EA did not contribute to this model (Wald = .000). However, it is important to note that all participants in this condition successfully completed the third puzzle.

A binomial logistic regression was then performed to ascertain the effects of successful completion of the first puzzle on the likelihood that participants succeed at assembling the second puzzle during the immediate testing phase. The logistic regression model was statistically significant, $\chi 2(1) = 12.993, p < 0.001$. The model explained 26.5% (Nagelkerke R^2) of the variance in success rate and correctly classified 73.3% of cases. The model suggested that participants who succeeded at correctly assembling the first puzzle were 7.65 times as likely to successfully assemble the second puzzle than participants in the reference category (PV_IB).

A binomial logistic regression was also performed to ascertain the effects of successful completion of the second puzzle on the likelihood that participants succeed at assembling the third puzzle during the immediate testing phase. The logistic regression model was statistically significant, $\chi 2(1) = 15.174, p < 0.001$. The model explained 30.8% (Nagelkerke R^2) of the variance in success rate and correctly classified 76.7% of cases. The model suggested that participants who succeeded at correctly assembling the second puzzle were 9.687 times as likely to successfully assemble the third puzzle than participants in the reference category (PV_IB).

As a result, a binomial logistic regression was performed to ascertain the effects of successful completion of the first puzzle and experimental condition on the likelihood that participants succeed at assembling the second puzzle during the immediate testing phase. The logistic regression model was statistically significant, $\chi 2(1) = 22.265, p = 0.001$. The model explained 42.1% (Nagelkerke R^2) of the variance in success rate and correctly classified 75% of cases. The Wald criterion demonstrated that none of the experimental conditions made a significant contribution to prediction. The Wald criterion also showed that successful completion of the previous puzzle did contribute significantly to prediction ($p = 0.003$). The model suggested that participants who succeeded at correctly assembling the first puzzle were 8.273 times as likely to successfully assemble the second puzzle than participants in the reference category (PV_IB). This model presented with the highest percentage of completely classified observations for the second puzzle.

A binomial logistic regression was also performed to ascertain the effects of successful completion of the second puzzle and experimental condition on the likelihood that participants succeed at assembling the third puzzle during the immediate testing phase. The logistic regression model was statistically significant, $\chi 2(1) = 32.441, p < 0.001$. The model explained 57.5% (Nagelkerke R^2) of the variance in success rate and correctly classified 83.3% of cases. The Wald criterion demonstrated that condition PV_I and successful completion of the previous puzzle made a significant contribution to prediction ($p = 0.030$ and $p = 0.007$, respectively). The model suggested that participants in the PV_I condition were 0.048 times as likely to successfully assemble the third puzzle than participants in the reference category (PV_IB). Participants who successfully completed the second puzzle were 8.475 times as likely to successfully assemble the third puzzle than participants in the reference category (PV_IB). Note that condition V_EA did not contribute to this model (Wald = .000). However, it is important to note that all participants in this condition successfully completed the third puzzle. This model presented with the highest percentage of completely classified observations for the

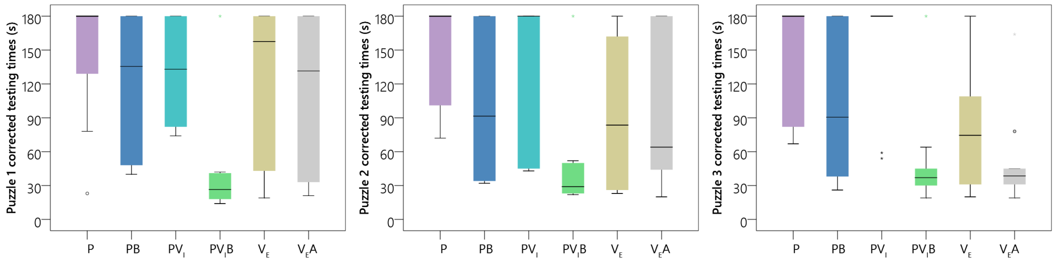


Figure 6.8: Boxplot containing corrected immediate testing times for each of the puzzles. Medians are shown as dark horizontal lines. Boxes represent the IQR. Whiskers represent either the extreme data points or extend to 1.5 x IQR. Outliers (data points outside the whiskers) are shown by circles. A value, X, is an outlier if X < lower quartile − 1.5 x interquartile range or if X < upper quartile + 1.5 x IQR. See Table 6.5 for pairwise interactions.

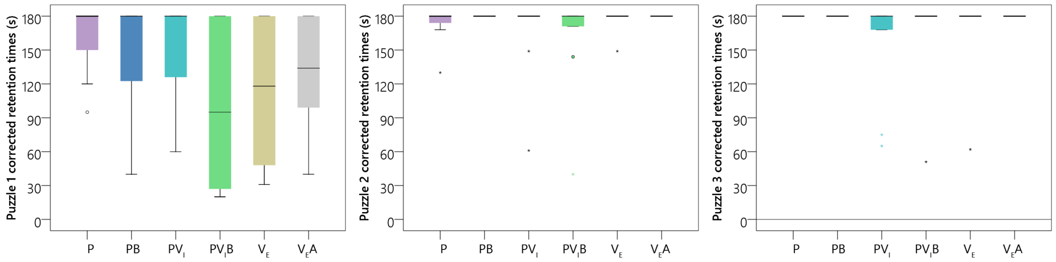


Figure 6.9: Boxplot containing corrected retention testing times for each of the puzzles. Medians are shown as dark horizontal lines. Boxes represent the IQR. Whiskers represent either the extreme data points or extend to 1.5 x IQR. Outliers (data points outside the whiskers) are shown by circles. A value, X, is an outlier if X < lower quartile − 1.5 x interquartile range or if X < upper quartile + 1.5 x IQR.

third puzzle.

To summarise, the binomial logistic regression model for the first puzzle was not statistically significant, with only condition P significantly contributing to the model. For the second puzzle, the binomial logistic regression model with the highest percentage of correctly classified observations was the one that ascertained the effect of successful completion of the previous puzzle during immediate testing. For the third puzzle, the binomial logistic regression model with the highest percentage of correctly classified observations was the one that ascertained the effect of both experimental condition and successful completion of the previous puzzle. These results show some support for H1 and H3.

Table 6.5: Test statistics using Dunn's procedure [1] for immediate testing times between the different experimental conditions. Significance values have been adjusted by the Bonferroni correction for multiple tests. Interaction is not significant unless it is explicitly indicated as specified in the legend.

		PB	**PV_I**	**$\text{PV}_\text{I}\text{B}$**	**V_E**	**$\text{V}_\text{E}\text{A}$**
Puzzle 1	P	5.00	4.25	27.35[a]	6.80	8.80
	PB		-0.75	22.35[a]	1.80	3.80
	PV_I			23.10[a]	2.55	4.55
	$\text{PV}_\text{I}\text{B}$				20.55	18.55
	VE					2.00
Puzzle 2	P	11.30	4.60	29.05[a]	16.90	14.65
	PB		-6.70	17.75	5.60	3.35
	PV_I			24.45[a]	12.30	10.05
	$\text{PV}_\text{I}\text{B}$				12.15	14.40
	VE					-2.25
Puzzle 3	P	11.50	-2.15	25.45[a]	15.70	25.40[a]
	PB		-13.65	13.95	4.20	13.90
	PV_I			27.60[a]	17.85	27.55[a]
	$\text{PV}_\text{I}\text{B}$				9.75	0.50
	VE					9.70

[a] interaction is significant at the 0.05 level (two-tailed)

6.3.2.3 Immediate testing completion times

We compared puzzle completion times between the different experimental conditions during the immediate testing phase. Completion time values were corrected by assigning the upper time limit (180s) to all unsuccessful attempts (see Section 6.3.1). All the corrected data satisfied the assumption of homogeneity.

Boxplots with immediate testing times for each of the puzzles are shown in Figure 6.8. Non-parametric statistical analysis was performed for immediate testing times because our data was not normally distributed as shown by a Shapiro-Wilk test.

A Kruskal-Wallis H test showed that there was an overall statistically significant difference in time taken to assemble the first puzzle in the testing phase between the different experimental conditions, $\chi^2(5) = 16.618$, p = 0.005, with a mean rank score of 39.20 for P, 34.20 for PB, 34.95 for PV_I, 11.85 for PV_IB, 32.40 for VE and 30.40 for V_EA.

A Kruskal-Wallis H test showed that there was an overall statistically significant difference in time taken to assemble the second puzzle in the testing phase between the different experimental conditions, $\chi^2(5) = 17.986, p = 0.003$, with a mean rank score of 43.25 for P, 31.95 for PB, 38.65 for PV_I, 14.20 for PV_IB, 26.35 for V_E and 28.60 for V_EA.

A Kruskal-Wallis H test showed that there was an overall statistically significant difference in time taken to assemble the third puzzle in the testing phase between the different experimental conditions, $\chi^2(5) = 24.536, p < 0.001$, with a mean rank score of 43.15 for P, 31.65 for PB, 45.30 for PV_I, 17.70 for PV_IB, 27.45 for V_E and 17.75 for V_EA.

Pairwise comparisons were performed using Dunn's procedure [1] with a Bonferroni correction for multiple comparisons with adjusted p-values. These are displayed in Table 6.5. Note that pairwise comparisons for puzzles in which the Kruskal-Wallis H test showed no overall statistically significant difference have not been included.

The post hoc analysis revealed statistically significant differences in immediate testing times for the first puzzle. There was a statistically significant difference between PV_IB (mean rank = 11.85) and P (mean rank = 39.20) ($p = 0.004$), PB (mean rank = 34.20) ($p = 0.003$) and PV_I (mean rank = 34.95) ($p = 0.002$).

The post hoc analysis revealed statistically significant differences in immediate testing times for the second puzzle. There was a statistically significant difference between PV_IB (mean rank = 14.20) and P (mean rank = 43.25) ($p = 0.002$) and PV_I (mean rank = 38.65) ($p = 0.019$).

The post hoc analysis revealed statistically significant differences in immediate testing times for the third puzzle. There was a statistically significant difference between PV_IB (mean rank = 17.70) and P (mean rank = 43.15) ($p = 0.013$) and PV_I (mean rank = 45.30) ($p = 0.005$). There was a statistically significant difference between V_EA (mean rank = 17.75) and P (mean rank = 43.15) ($p = 0.013$) and PV_I

(mean rank = 45.30) ($p = 0.005$).

The analysis of immediate testing completion times shows some support for H2 and H3. Separate statistical analysis showed no significant effect of gender or age on training times for all puzzles.

6.3.2.4 Subjective questionnaire ratings

There was no statistically significant difference in rated difficulty, ease of use and seriousness between groups as determined by one-way ANOVA for the first puzzle.

There was a statistically significant difference in ease of use of the training environment (F(5,54) = 5.006, p = 0.001) between groups as determined by one-way ANOVA for the second puzzle. A Tukey post hoc test revealed that participants in the P condition (M = 2.5, SD = 1.08) rated the ease of use of the training environment as significantly more difficult than participants in the V_E (M = 4.4, SD = 0.70, p = 0.001) and V_EA (M = 4.1, SD = 1.1, p = 0.007) conditions. No other significant interactions were found for the second puzzle.

There was a statistically significant difference in task difficulty (F(5,54) =4.613, p = 0.001) between groups as determined by one-way ANOVA for the third puzzle. A Tukey post hoc test revealed that participants in the P condition (M = 1.9, SD = 1.00) rated the difficulty of the task as significantly more difficult than participants in the V_EA (M = 4.1, SD = 0.88, p = 0.002) condition. Participants in the PB condition (M = 2.7, SD = 1.34) also rated the difficulty of the task as significantly more difficult than participants in the VEA condition (M = 4.1, SD = 1.34, p = 0.003). No other significant interactions were found for the third puzzle.

There was a statistically significant difference in ease of use of the training environment (F(5,54) = 3.044, p = 0.017) between groups as determined by one-way ANOVA for the third puzzle. A Tukey post hoc test revealed that participants in the P condition (M = 2.4, SD = 1.35) rated the ease of use of the training environment as significantly more difficult than participants in the V_EA (M = 4.4, SD = 0.70, p = 0.007) condition. No other significant interactions were found for the third puzzle.

Table 6.6: Number of successful attempts, failed attempts and participants solving the three puzzles in the retention testing phase for each of the experimental conditions.

		P	**PB**	**PV$_I$I**	**PV$_I$B**	**V$_E$**	**V$_E$A**
Puzzle 1	Success	2	3	3	6	7	5
	Fail	6	6	7	4	3	4
Puzzle 2	Success	2	0	2	4	1	0
	Fail	6	9	8	6	9	9
Puzzle 3	Success	0	0	3	1	1	0
	Fail	8	9	7	9	9	9
N		8	9	10	10	10	9

6.3.3 Retention Testing

6.3.3.1 Participants

A total of 56 participants who completed the first part session returned to complete the second session two weeks later (average number of days between training session and retention session: 14.16, SD = 0.918). Overall, retention testing performance was lower than expected for all conditions both in terms of success rates and completion times. We believe this is due to the high complexity of the 3D puzzles.

6.3.3.2 Retention testing success rates

A binomial logistic regression was performed to ascertain the effects of experimental condition on the likelihood that participants succeed at assembling each puzzle during the immediate testing phase. PV$_I$B was chosen as the reference category (the condition with most successful puzzle completions, overall).

The binomial logistic regression model was not statistically significant, $\chi^2(5) = 6.240, p = 0.284$ for the first puzzle. The model explained 14.3% (Nagelkerke R^2) of the variance in success rate and correctly classified 65.5% of cases. The Wald criterion demonstrated that none of the conditions made a significant contribution to prediction.

The binomial logistic regression model was not statistically significant, $\chi^2(5) = 10.054, p = 0.074$ for the second puzzle. The model explained 28.3%

(Nagelkerke R^2) of the variance in success rate and correctly classified 83.6% of cases. The Wald criterion demonstrated that none of the conditions made a significant contribution to prediction.

The binomial logistic regression model was not statistically significant, $\chi^2(5) = 8.289, p = 0.141$ for the third puzzle. The model explained 30.7% (Nagelkerke R^2) of the variance in success rate and correctly classified 90.9% of cases. The Wald criterion demonstrated that none of the conditions made a significant contribution to prediction.

6.3.3.3 Retention testing completion times

We compared puzzle retention testing times between the different experimental conditions. Completion time values were corrected by assigning the upper time limit (180s) to all unsuccessful attempts. All the corrected data satisfied the assumption of homogeneity.

Boxplots with training times for each of the puzzles are shown in Figure 6.9. Non-parametric statistical analysis was performed for retention testing times because our data was not normally distributed as shown by a Shapiro-Wilk test.

A Kruskal-Wallis H test showed that there was no overall statistically significant difference in time taken to assemble the first puzzle in the retention testing phase between the different experimental conditions, $\chi^2(5) = 8.101$, p = 0.151, with a mean rank score of 34.69 for P, 32.88 for PB, 33.45 for PV_I, 20.90 for PV_IB, 21.95 for V_E and 26.28 for V_EA.

A Kruskal-Wallis H test showed that there was no overall statistically significant difference in time taken to assemble the second puzzle in the retention testing phase between the different experimental conditions, $\chi^2(5) = 5.832$, p = 0.323, with a mean rank score of 25.25 for P, 32.00 for PB, 26.35 for PV_I, 23.70 for PV_IB, 29.35 for V_E and 32.00 for V_EA.

A Kruskal-Wallis H test showed that there was no overall statistically significant difference in time taken to assemble the third puzzle in the retention testing phase between the different experimental conditions, $\chi^2(5) = 7.151$, p = 0.210, with a mean rank score of 30.50 for P, 30.50 for PB, 22.55 for PV_I, 27.55 for

PV_IB, 27.65 for V_E and 30.50 for V_EA.

6.3.3.4 Subjective questionnaire ratings

There was no statistically significant difference in rated difficulty and seriousness between groups as determined by one-way ANOVA for any of the three puzzles. Tukey post hoc tests showed no significant interactions.

6.4 Discussion

In terms of training times, post hoc analysis revealed a significant difference between the physical conditions where no blocks were available during training (P and PV_I) and the rest of the physical conditions where blocks were available during training (PB and PV_IB), amongst other significant interactions. For the second puzzle, we observe a significant difference in training times between PV_I and all other conditions except P, amongst other effects. For the third puzzle we found no significant interactions. We believe it is important to note the lack of significant differences in terms of training times between the virtual conditions and condition PV_IB, the overall best performing condition. We also believe that the lower training times for conditions P and PB could be due to the lack of blocks to practice with during training, which meant participants did not have any activities to perform during training and therefore decided to move on to the next stage of the experimental task. This could be related to a high number of unsuccessful puzzle completions in these conditions. An increase in training times for these conditions in puzzles 2 and 3 could be due to participants understanding the complexity of the tasks after the immediate testing for the first puzzle and deciding to spend more time inspecting the paper instructions and video (when available).

Regarding success rates for immediate testing, we observed that condition PV_IB yielded the highest number of successful completions of the three puzzles (see Figure 6.7). We also observed that condition P yielded the lowest number of successful completions of the puzzles during immediate testing. Condition PB showed a ceiling effect in the second and third puzzle. Successful puzzle completions in condition PV_I decreased with each puzzle. Immediate testing success rates for the virtual conditions, V_E and V_EA, increased with each puzzle. Our analysis showed that the binomial logistic regression model for the first puzzle was not statistically significant, with only condition P significantly contributing to the model. For the second puzzle, the binomial logistic regression model with the highest percentage of correctly classified observations was the one that ascertained the effect of the successful completion of the previous puzzle during immediate testing. For the third puzzle, the binomial logistic regression model with the highest percentage of correctly classified observations was ones that ascertained the effect of experimental condition as well as the successful completion of the previous puzzle during immediate testing.

In terms of immediate testing completion times, we observed how (parallel to an increase in success rate) the immediate testing times for condition V_EA decreased with each puzzle. A statistically significant difference was not found between this condition and condition PV_IB (the condition with overall lowest testing times). This result could indicate that the availability of static and animated instructions in the virtual training environment contributed to effective training.

Anecdotal evidence from the training videos as well as participant feedback during debrief interviews shows that virtually trained participants initially struggled to assemble the pieces during the immediate testing phase. We believe this is due to the lack of experience in handling and joining the physical blocks during training. However, after the first and second tasks, participants refined their strategy during the training stage to include physically plausible movements of the puzzles pieces. This is, participants replicated the movement they would then perform with the physical blocks in the virtual training environment and avoided allowing the pieces to go through each other, as no physics restrictions were assigned to the virtual blocks in the virtual training environment.

Subjective questionnaire ratings answered by participants during the first session showed no statistically significant difference in rated difficulty, ease of use and seriousness between groups as determined by one-way ANOVA for the first puzzle. For the second puzzle, results indicated that participants in the P condition rated the training environment as significantly more difficult to use than participants in the V_E and V_EA conditions. When asked about task difficulty in the third puzzle, participants in the P and PB conditions rated the difficulty of the task as significantly more difficult than participants in the V_EA condition. In terms of ease of use, participants in the P condition rated the ease of use of the training environment as significantly more difficult than participants in the V_EA condition.

One of the limitations in our design was the high complexity of the puzzles. Overall, retention testing resulted in lower performance than we had expected and we believe this is due to the difficulty associated with remembering the process to solve the three puzzles two weeks after the training. This was further validated by verbal feedback from our participants during the second session. Our previous piloting of the task had not shown this effect. Future studies should further evaluate the suitability of the task for retention. This evaluation should aim to balance the amount of training and complexity of the task to avoid floor and ceiling effects in subsequent retention sessions.

Chapter 7

Conclusions

This chapter summarises the work presented in the thesis. We reiterate the main results from each of the four studies presented. We discuss the limitations of our designs as well as the implications for virtual training. We then synthesise and discuss overall conclusions across all studies. Finally, we introduce directions for future work in this area.

7.1 Conclusions on Distorting Real and Virtual Space

In Chapter 3 and Chapter 4 we introduced two studies on human spatial memory. Participants were asked to collect an object in a room, exit the room, re-enter the room and then place the object back where they had found it. The room was geometrically modified between learning and placement of the object. We then compared participant object placement with different models derived from previous spatial cognition experiments. Chapter 3 reports the physical, real-world version of the study. Chapter 4 reports the virtual counterpart version of the study.

Overall, results highlight strong differences in participant behaviour, with contrasting models describing different response object locations. No single model can account for all participant behaviour accross systems (in both the real world and VR versions).

These studies have not replicated findings from Hartley et al., who reported that human spatial representations are likely determined by proximities to environment boundaries, similar to rodent place cell firing. This study was run in a desktop VR system. Our results show that self-motion contributes significantly to spatial representations in combination with geometric information. It is not clear if distance information is favoured over angle information when remembering object locations. These findings indicate that grid cells, which have a role in the brains coordinate system for navigation, may also have a substantial role in human spatial memory, although further work is needed to confirm this. Whilst this study cannot refute findings from Hartley et al., results here certainly show that the findings are not sufficient to reconcile a behaviourally based neural account of human spatial memory.

Our results highlight the need to test more suitable models that merge memory for geometry and self-motion, or combinations of the ones presented here. It also emphasises the importance of spatial layout for training VEs as well as the user's starting location and facing direction, which have a strong effect on participant behaviour. In addition, it further motivates the design of experimental setups that will maximise the differences in locations predicted by the models and make it more clear if participant behaviour is actually following a specific model. Future work includes comparing our results with a desktop VR version of the study.

We believe that this study could inform the design of followup studies on spatial cognition in immersive VR and assist experts in the design of training simulations where users are required to remember object locations. Followup experiments should further examine the effectiveness of spatial training in immersive VR, specifically in situations in which the work environment, where the acquired skills will be used, is unknown or difficult to replicate as a virtual model. Under this circumstances, trainees would undergo the training simulation in a virtual space that is different to the work environment. Design of training VEs would strongly benefit from more accurate behavioural models.

7.2 Conclusions on the Effect of Environmental Features, Self-Avatar and Level of Immersion on Object Location Memory in Virtual Environments

In Chapter 5 we ran a study on object location memory. The experimental task involves several judgements, including distance estimation, and it is not clear exactly what strategies participants use to learn object locations [5]. Previous work has shown that distance estimation is impaired within immersive VR, although including a self-avatar and increasing confidence in fidelity can reduce this impairment [25, 102, 103, 104, 38, 26, 32, 34, 37, 105]. Our results suggest that level of immersion is extremely important for accurate object location learning and recall, and that higher environmental fidelity may reinforce learning transfer from VEs to the real world. However, most importantly, they indicate that providing users with a virtual body can interfere with successful completion of the task. This motivates studies of more complex self-representations.

We believe that the main outcomes of this study could be generalised to other spatial learning scenarios and assist experts in the design of training simulations related to spatial memory, where trainees are required to remember component or tool locations as part of the task. Overall, our results denote that HMD training resembles real world training more than desktop learning, related to higher object location memory accuracy. However, desktop training applications can be suitable and offer acceptable results when precise location learning accuracy is not required. Regarding self-avatars, our results suggest that a low fidelity avatar representation can degrate object location memory. In our experimental task, this observation is particularly important when the training transfer takes place from a low fidelity VE, where only basic geometric cues are available, to the real world equivalent.

7.3 Conclusions on A Comparison of Virtual and Physical Training Transfer of Bimanual Assembly Tasks

In Chapter 6 we ran a study that compares the effectiveness of virtual and traditional paper- and video-based training transfer of a bimanual assembly task motivated by previous research [59, 2]. In a between-subjects experimental design, participants were trained to solve three six-piece burr puzzles in a virtual training environment or a physical training environment. The conditions were designed to account for situations in which the physical puzzle blocks are available or not during training. The conditions were also devised to include static instructions (paper) or combinations of static and animated instructions (video or 3D animations).

Following training, participants were asked to solve physical versions of the puzzles. Participants then completed a retention session two weeks after the training. During the course of the study participants answered mental rotations tests and questionnaires measuring several aspects of the experience.

We hypothesised that the experimental conditions where the physical blocks were available during training (PB and PV_IB) would result in better performance than the other conditions as measured by success rates and puzzle completion times. Overall, we expected that those conditions were video or 3D animations were available (PV_IB and V_EA) would result in lower assembly times during immediate and retention testing. Out of those, we predicted that condition PV_IB, with animated instructions (video), would yield the highest performance. Although there were conditions we expected to deliver worse or better performance, we had no hypothesis on the full order so all the analysis presented in this manuscript is two-tailed.

Our results highlight the effectiveness of the virtual training environment. Success rates and completion times indicate that the performance of virtually trained participants (in conditions V_E and V_EA) increased with each puzzle, reaching the level of the best performing physical condition (PV_IB).

Following a between-subjects experimental design, participants were trained to assemble three versions of a 3D burr puzzle in one of six experimental conditions (see Table 6.1 for definitions). All participants completed an immediate testing phase and a retention test two weeks after the training, both with physical versions of the puzzles. Participants were trained and tested in solving the puzzles in increasing order of complexity in that they had to remember a higher number of steps in the assembly process for each of the puzzles.

We analysed performance in terms of success rates as well as immediate testing times and retention testing times. Our results show that the performance of virtually trained participants was promising. Condition $\mathrm{V_EA}$ yielded success rates and immediate testing times similar to the best performing physical condition ($\mathrm{PV_IB}$, in which physical blocks and animated instructions were available during training) for the last and most complex puzzle. We believe these results are of great importance given that virtually trained participants did not have the chance to interact with the physical blocks at any point during training. We also observed that participants were more likely to successfully assemble a puzzle during immediate testing if they had successfully assembled the previous one. Retention testing performance was unexpectedly low due to the high complexity of the task. We believe that the results of this study further validate the effectiveness of virtual training for bimanual assembly tasks.

7.4 Overall Conclusions

We discussed the requirement for more sophisticated models of spatial cognition that merge memory for geometry and self-motion. A more granular analysis of participant behaviour, supported by in-depth debrief interviews to understand their strategies, could help refine the current models. This could include head direction or eye-gaze data collection to analyse the cues that participants were using to place the object in Chapter 3 and Chapter 4.

We observed the need to explore orientation cues other than lighting to aid participants in reorienting to a previous facing direction as it is not clear whether the amber lighting at the UCL PAMELA facility was sufficient for this purpose in in Chapter 3 and Chapter 4. Results from this study also highlighted the importance of room layout, starting position and facing direction, as these largely influenced participant behaviour.

We learnt about the importance of featural, non-geometric cues (by design, such as plugs and wall decorations, or accidental, such as wear and tear) for accurate object location memory (in Chapter 3, Chapter 4 and Chapter 5). The degree of environmental fidelity of a VE will be dependent on the training task and the level of accuracy needed in performing the task after training.

We observed the superiority of HMD-based VEs over desktop-based VEs for accurate object location learning as well as performance degradation with single-point tracked avatars in Chapter 5.

We noted the success of transparent hand representation and ambidextrous interaction for training transfer of a bimanual assembly task in Chapter 6. Observations from this study also raised the question around haptics: whether the adaptation period for virtually trained participants during which they adjust their strategy to replicate only physically plausible movements can be shortened if a haptic device is available. This study also raised the question of task complexity and the difficulty in designing experimental tasks that avoid floor and ceiling effects in subsequent retention testing.

7.5 Directions for Future Work

The question on whether VR consumer technology can deliver effective training for any type of task remains unanswered. Recent advances in the field have provided evidence that virtual training is possible and promising. However, more research is needed to ascertain the full effectiveness of this technology for training by testing across a richer spectrum of tasks. We suggest including real world training as a baseline condition in future research endeavours in this area, as well as testing in the real world, to effectively measure the training transfer. We also recommend the evaluation of other parameters that mediate the transfer from virtual to real enviroments such as spatial sound, locomotion, avatar self-representation and haptics in order to refine and optimise the design of VEs for training.

Equally, there is a need to further explore the validity of research studies that use VEs as proxy environments to replicate real-world scenarios. Despite the numerous benefits of using this technology, including the high level of experimental control and the ease of data capture, it is unclear whether this technology induces other nuances that can affect the way participants understand, behave in or solve a task. We encourage further exploration of the long-term effects of VR system usage with longitudinal trials that span longer periods of time to better understand learning curves and familiarisation.

Moreover, the spatial memory models explored in this thesis failed to capture the behaviour of our participants, not only within each of the experimental setups, but across levels of immersion. We recommend augmenting the modalities of data capture to include head direction or eye gaze in order to better understand which cues determine spatial memory. We also suggest extending the work presented in this thesis to explore other spatial layouts and boundary distortions in the process of defining new models. This raises the need for further exploration in the fields of behavioural neuroscience and experimental psychology as well as the requirement for these disciplines to continue to collaborate with researchers in the field of computer science, and, specifically, VEs.

Appendix A

Publications

The following publications, appearing in or submitted to peer-reviewed conferences and journals, have derived from this project. These are presented in chronological order according to date of publication.

Murcia-López, M., Zisch, F.E., Steptoe, W., Spiers, H.J., & Steed, A. (2015). Preliminary Findings on the Effect of Scale Transformations on Spatial Memory in Immersive and Desktop Virtual Environments. *Cognitive Processing*, 16, S75.

Murcia-López, M., & Steed, A. (2016). The Effect of Environmental Features, Self-Avatar, and Immersion on Object Location Memory in Virtual Environments. *Frontiers in ICT*, 3, 24.

Xu, M., **Murcia-López, M.**, & Steed, A. (2017, March). Object Location Memory Error in Virtual and Real Environments. In Virtual Reality (VR), 2017 IEEE (pp. 315-316).

Murcia-López, M., & Steed, A. (2018). A Comparison of Virtual and Physical Training Transfer of Bimanual Assembly Tasks. *IEEE Transactions on Visualization and Computer Graphics*, 24(4), 1574-1583.

Zisch, F.E.[1], **Murcia-López, M.**[1], Coutrout, A.[1], Motala, A., Greaves, J., De Cothi, W., Newton, C., Steed, A., Gage, S.A., & Spiers, H.J. (2018). Distorting Space: The Impact of Environmental Boundary Changes on Human Spatial Memory in Real and Virtual Spaces. *Manuscript in preparation.*

---

[1] These authors contributed equally to the work.

Appendix B

List of Acronyms

2D Two Dimensional

3D Three Dimensional

AD Absolute Distance

ANOVA Analysis of Variance

BP Boundary Proximity

CA Corner Angle

DC Distance Criteria

FDA Fixed Distance Allocentric

FDE Fixed Distance Egocentric

FOV Field of View

FOR Field of Regard

FRA Fixed Ratio Allocentric

FRE Fixed Ratio Egocentric

HMD Head-Mounted Display

HSSMI High Speed Sustainable Manufacturing Institute

IQR Interquartile Ranges

P Paper (Experimental condition in Chapter 6)

PAMELA Pedestrian Accessibility Movement Environment Laboratory

PB Paper and Blocks (Experimental condition in Chapter 6)

PI Path Integration

PV$_{\mathrm{I}}$ Paper and Video (Experimental condition in Chapter 6)

PV$_{\mathrm{I}}$B Paper, Video and Blocks (Experimental condition in Chapter 6)

PV Path Vector

PVR Path Vector Ratio

QC Quadrant Criteria

UCL University College London

UK United Kingdom

VE Virtual Environment

V$_{\mathrm{E}}$ Virtual Environment (Experimental condition in Chapter 6)

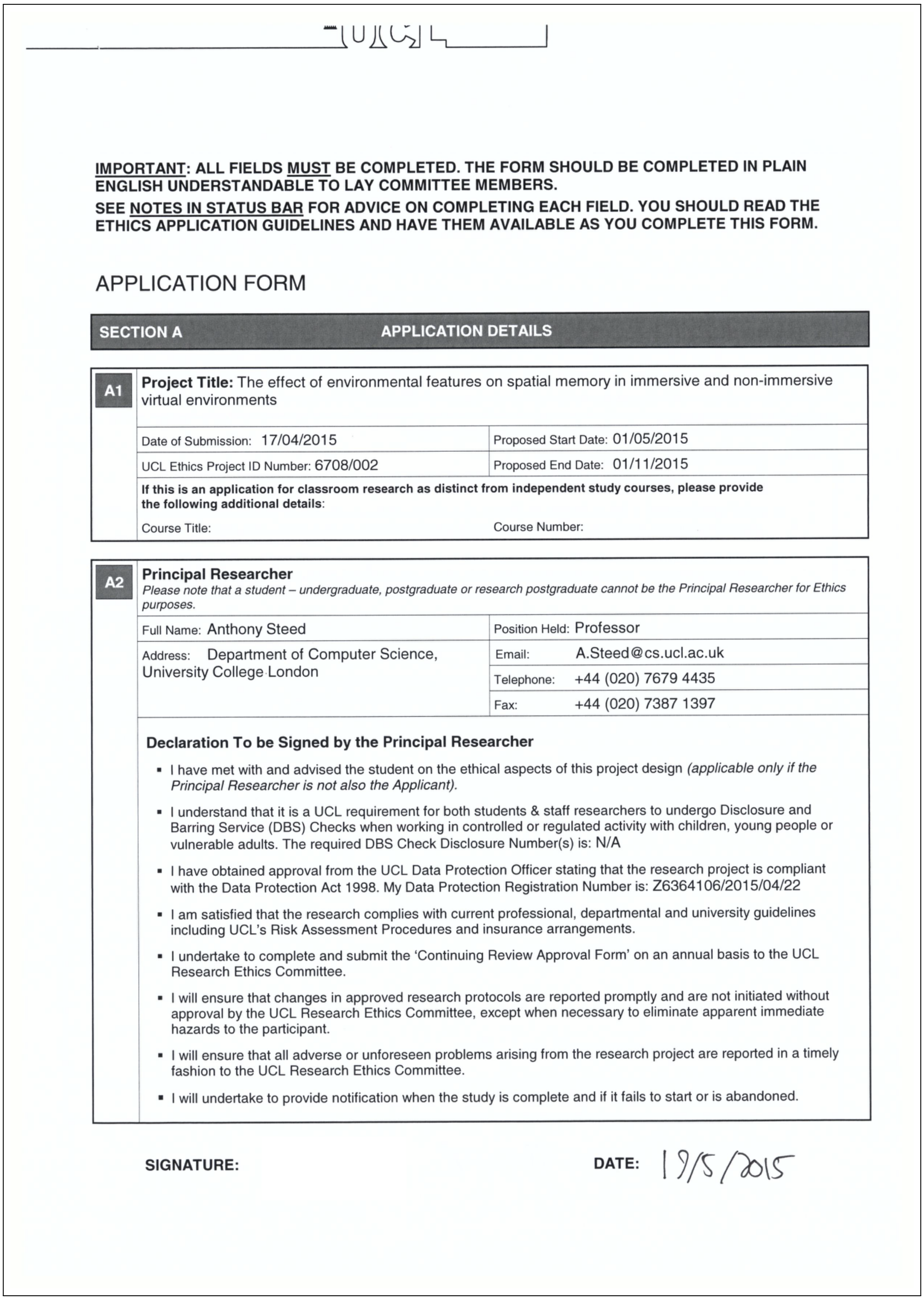
V$_{\mathrm{E}}$A Virtual Environment with Animations (Experimental condition in Chapter 6)

VR Virtual Reality

Appendix C

Ethics Application for The Effect of Environmental Features, Self-Avatar and Level of Immersion on Object Location Memory in Virtual Environments

UCL Research Ethics Committee Application (Project ID: 6708/002) for the study presented in Chapter 5. This application document contains the information sheet as well as the informed consent form that participants had to sign to take part in the study.



IMPORTANT: ALL FIELDS **MUST** BE COMPLETED. THE FORM SHOULD BE COMPLETED IN PLAIN ENGLISH UNDERSTANDABLE TO LAY COMMITTEE MEMBERS.

SEE NOTES IN STATUS BAR FOR ADVICE ON COMPLETING EACH FIELD. YOU SHOULD READ THE ETHICS APPLICATION GUIDELINES AND HAVE THEM AVAILABLE AS YOU COMPLETE THIS FORM.

APPLICATION FORM

SECTION A APPLICATION DETAILS

A1

Project Title: The effect of environmental features on spatial memory in immersive and non-immersive virtual environments

Date of Submission: 17/04/2015	Proposed Start Date: 01/05/2015
UCL Ethics Project ID Number: 6708/002	Proposed End Date: 01/11/2015

If this is an application for classroom research as distinct from independent study courses, please provide the following additional details:

Course Title:	Course Number:

A2

Principal Researcher

Please note that a student – undergraduate, postgraduate or research postgraduate cannot be the Principal Researcher for Ethics purposes.

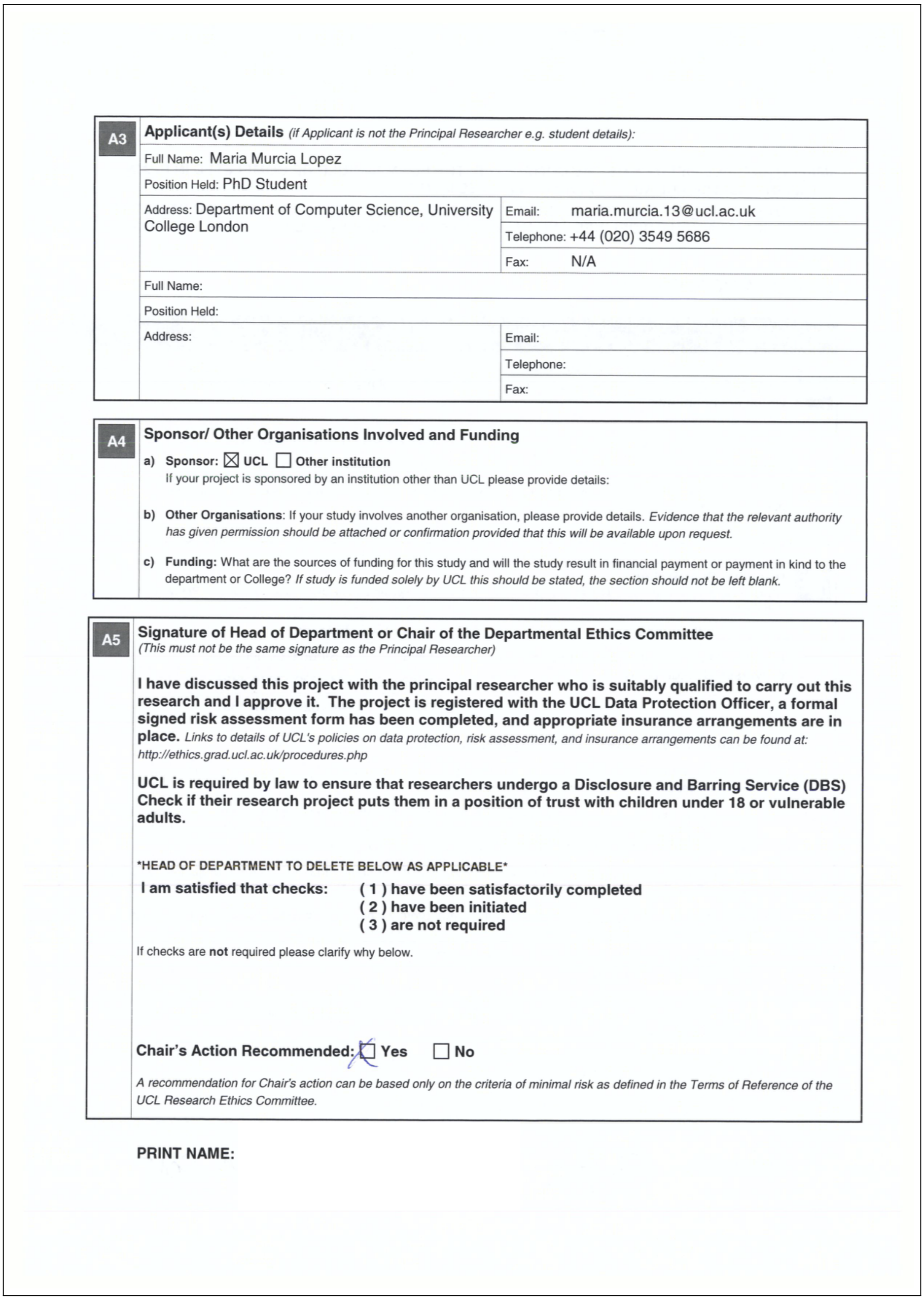
Full Name: Anthony Steed	Position Held:	Professor
Address: Department of Computer Science, University College London	Email:	A.Steed@cs.ucl.ac.uk
	Telephone:	+44 (020) 7679 4435
	Fax:	+44 (020) 7387 1397

Declaration To be Signed by the Principal Researcher

- I have met with and advised the student on the ethical aspects of this project design *(applicable only if the Principal Researcher is not also the Applicant).*
- I understand that it is a UCL requirement for both students & staff researchers to undergo Disclosure and Barring Service (DBS) Checks when working in controlled or regulated activity with children, young people or vulnerable adults. The required DBS Check Disclosure Number(s) is: N/A
- I have obtained approval from the UCL Data Protection Officer stating that the research project is compliant with the Data Protection Act 1998. My Data Protection Registration Number is: Z6364106/2015/04/22
- I am satisfied that the research complies with current professional, departmental and university guidelines including UCL's Risk Assessment Procedures and insurance arrangements.
- I undertake to complete and submit the 'Continuing Review Approval Form' on an annual basis to the UCL Research Ethics Committee.
- I will ensure that changes in approved research protocols are reported promptly and are not initiated without approval by the UCL Research Ethics Committee, except when necessary to eliminate apparent immediate hazards to the participant.
- I will ensure that all adverse or unforeseen problems arising from the research project are reported in a timely fashion to the UCL Research Ethics Committee.
- I will undertake to provide notification when the study is complete and if it fails to start or is abandoned.

SIGNATURE:

DATE: 19/5/2015



A3 **Applicant(s) Details** *(if Applicant is not the Principal Researcher e.g. student details):*

Full Name: Maria Murcia Lopez

Position Held: PhD Student

Address: Department of Computer Science, University College London

Email: maria.murcia.13@ucl.ac.uk

Telephone: +44 (020) 3549 5686

Fax: N/A

Full Name:

Position Held:

Address:

Email:

Telephone:

Fax:

A4 **Sponsor/ Other Organisations Involved and Funding**

a) **Sponsor:** ☒ **UCL** ☐ **Other institution**
If your project is sponsored by an institution other than UCL please provide details:

b) **Other Organisations**: If your study involves another organisation, please provide details. *Evidence that the relevant authority has given permission should be attached or confirmation provided that this will be available upon request.*

c) **Funding:** What are the sources of funding for this study and will the study result in financial payment or payment in kind to the department or College? *If study is funded solely by UCL this should be stated, the section should not be left blank.*

A5 **Signature of Head of Department or Chair of the Departmental Ethics Committee**
(This must not be the same signature as the Principal Researcher)

I have discussed this project with the principal researcher who is suitably qualified to carry out this research and I approve it. The project is registered with the UCL Data Protection Officer, a formal signed risk assessment form has been completed, and appropriate insurance arrangements are in place. *Links to details of UCL's policies on data protection, risk assessment, and insurance arrangements can be found at: http://ethics.grad.ucl.ac.uk/procedures.php*

UCL is required by law to ensure that researchers undergo a Disclosure and Barring Service (DBS) Check if their research project puts them in a position of trust with children under 18 or vulnerable adults.

HEAD OF DEPARTMENT TO DELETE BELOW AS APPLICABLE

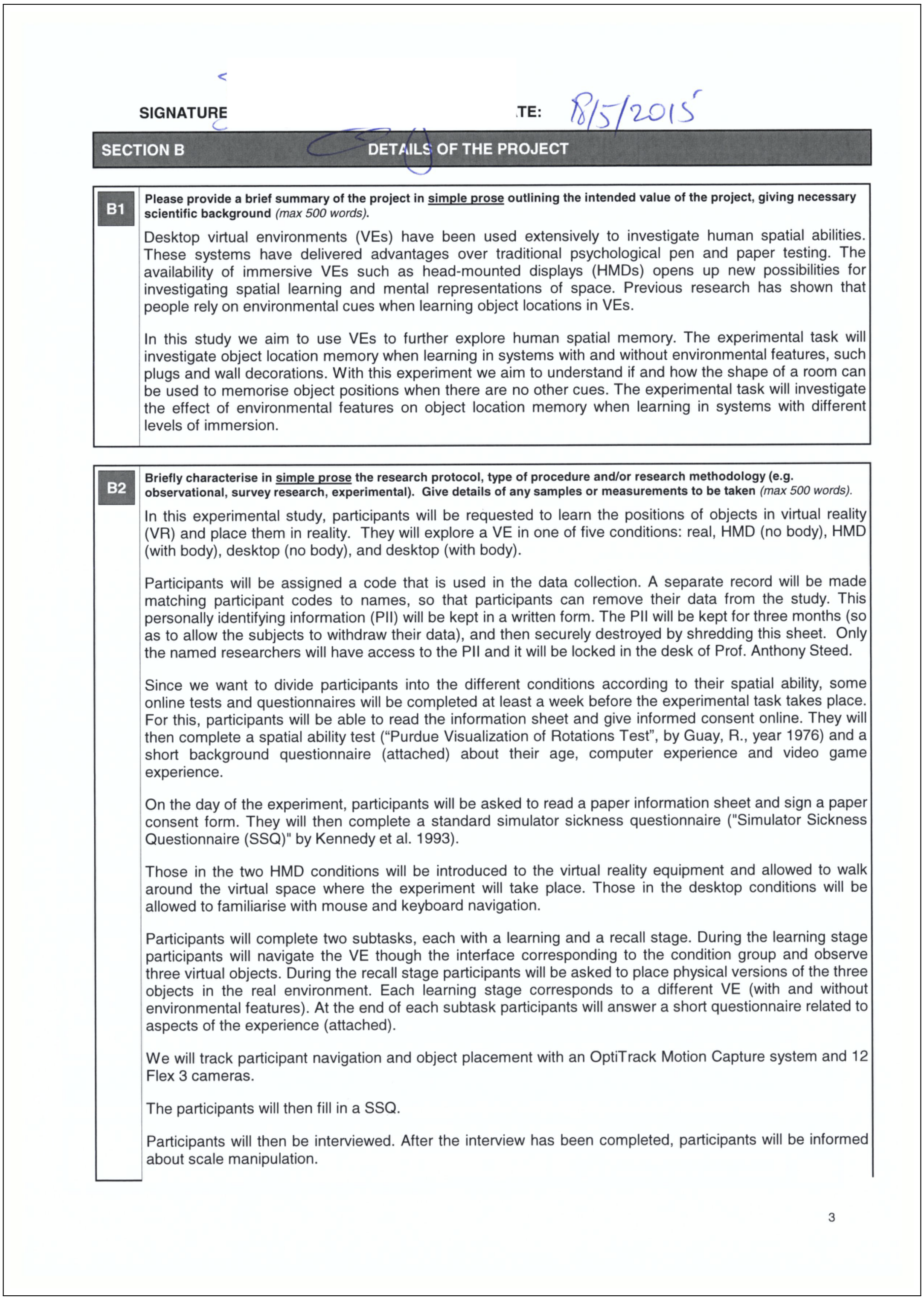
I am satisfied that checks: **(1) have been satisfactorily completed**
(2) have been initiated
(3) are not required

If checks are **not** required please clarify why below.

Chair's Action Recommended: ☒ **Yes** ☐ **No**

A recommendation for Chair's action can be based only on the criteria of minimal risk as defined in the Terms of Reference of the UCL Research Ethics Committee.

PRINT NAME:



SIGNATURE .TE: 8/5/2015

SECTION B DETAILS OF THE PROJECT

B1 **Please provide a brief summary of the project in <u>simple prose</u> outlining the intended value of the project, giving necessary scientific background** *(max 500 words).*

Desktop virtual environments (VEs) have been used extensively to investigate human spatial abilities. These systems have delivered advantages over traditional psychological pen and paper testing. The availability of immersive VEs such as head-mounted displays (HMDs) opens up new possibilities for investigating spatial learning and mental representations of space. Previous research has shown that people rely on environmental cues when learning object locations in VEs.

In this study we aim to use VEs to further explore human spatial memory. The experimental task will investigate object location memory when learning in systems with and without environmental features, such plugs and wall decorations. With this experiment we aim to understand if and how the shape of a room can be used to memorise object positions when there are no other cues. The experimental task will investigate the effect of environmental features on object location memory when learning in systems with different levels of immersion.

B2 **Briefly characterise in <u>simple prose</u> the research protocol, type of procedure and/or research methodology (e.g. observational, survey research, experimental). Give details of any samples or measurements to be taken** *(max 500 words).*

In this experimental study, participants will be requested to learn the positions of objects in virtual reality (VR) and place them in reality. They will explore a VE in one of five conditions: real, HMD (no body), HMD (with body), desktop (no body), and desktop (with body).

Participants will be assigned a code that is used in the data collection. A separate record will be made matching participant codes to names, so that participants can remove their data from the study. This personally identifying information (PII) will be kept in a written form. The PII will be kept for three months (so as to allow the subjects to withdraw their data), and then securely destroyed by shredding this sheet. Only the named researchers will have access to the PII and it will be locked in the desk of Prof. Anthony Steed.

Since we want to divide participants into the different conditions according to their spatial ability, some online tests and questionnaires will be completed at least a week before the experimental task takes place. For this, participants will be able to read the information sheet and give informed consent online. They will then complete a spatial ability test (“Purdue Visualization of Rotations Test”, by Guay, R., year 1976) and a short background questionnaire (attached) about their age, computer experience and video game experience.

On the day of the experiment, participants will be asked to read a paper information sheet and sign a paper consent form. They will then complete a standard simulator sickness questionnaire ("Simulator Sickness Questionnaire (SSQ)" by Kennedy et al. 1993).

Those in the two HMD conditions will be introduced to the virtual reality equipment and allowed to walk around the virtual space where the experiment will take place. Those in the desktop conditions will be allowed to familiarise with mouse and keyboard navigation.

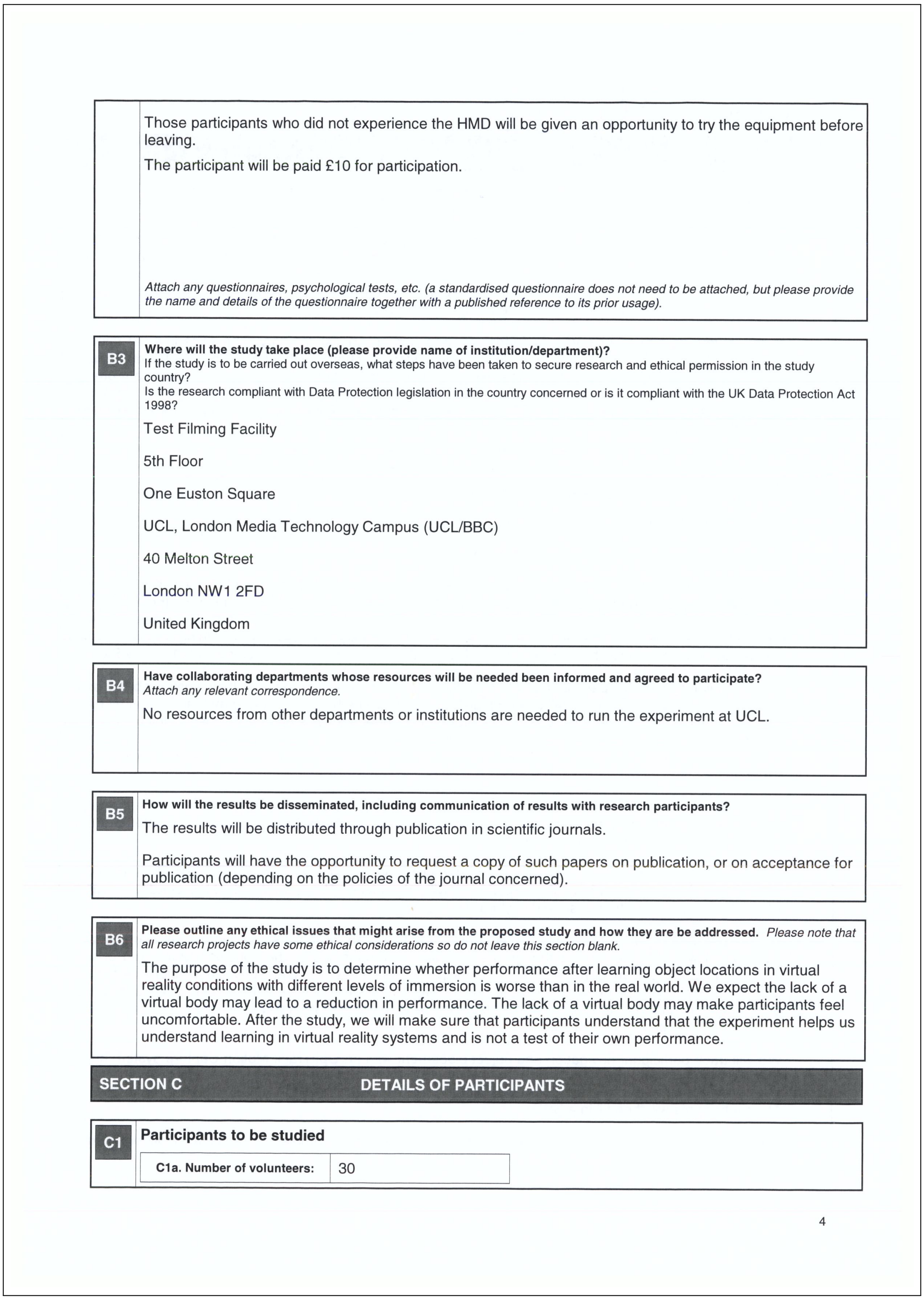
Participants will complete two subtasks, each with a learning and a recall stage. During the learning stage participants will navigate the VE though the interface corresponding to the condition group and observe three virtual objects. During the recall stage participants will be asked to place physical versions of the three objects in the real environment. Each learning stage corresponds to a different VE (with and without environmental features). At the end of each subtask participants will answer a short questionnaire related to aspects of the experience (attached).

We will track participant navigation and object placement with an OptiTrack Motion Capture system and 12 Flex 3 cameras.

The participants will then fill in a SSQ.

Participants will then be interviewed. After the interview has been completed, participants will be informed about scale manipulation.

3



Those participants who did not experience the HMD will be given an opportunity to try the equipment before leaving.

The participant will be paid £10 for participation.

Attach any questionnaires, psychological tests, etc. (a standardised questionnaire does not need to be attached, but please provide the name and details of the questionnaire together with a published reference to its prior usage).

B3 **Where will the study take place (please provide name of institution/department)?**
If the study is to be carried out overseas, what steps have been taken to secure research and ethical permission in the study country?
Is the research compliant with Data Protection legislation in the country concerned or is it compliant with the UK Data Protection Act 1998?

Test Filming Facility

5th Floor

One Euston Square

UCL, London Media Technology Campus (UCL/BBC)

40 Melton Street

London NW1 2FD

United Kingdom

B4 **Have collaborating departments whose resources will be needed been informed and agreed to participate?**
Attach any relevant correspondence.

No resources from other departments or institutions are needed to run the experiment at UCL.

B5 **How will the results be disseminated, including communication of results with research participants?**

The results will be distributed through publication in scientific journals.

Participants will have the opportunity to request a copy of such papers on publication, or on acceptance for publication (depending on the policies of the journal concerned).

B6 **Please outline any ethical issues that might arise from the proposed study and how they are be addressed.** *Please note that all research projects have some ethical considerations so do not leave this section blank.*

The purpose of the study is to determine whether performance after learning object locations in virtual reality conditions with different levels of immersion is worse than in the real world. We expect the lack of a virtual body may lead to a reduction in performance. The lack of a virtual body may make participants feel uncomfortable. After the study, we will make sure that participants understand that the experiment helps us understand learning in virtual reality systems and is not a test of their own performance.

SECTION C DETAILS OF PARTICIPANTS

C1 **Participants to be studied**

C1a. Number of volunteers:	30

Upper age limit:	N/A
Lower age limit:	18

C1b. Please justify the age range and sample size:
Adult participants are required to reduce subject bias. There are two VEs and 2 orders in which the subtasks can be ordered. We want to test all orders an equal number of times (3 times in each condition). We therefore need 30 participants.

C2 **If you are using data or information held by a third party, please explain how you will obtain this. You should confirm that the information has been obtained in accordance with the UK Data Protection Act 1998.**

N/A

C3 **Will the research include children or vulnerable adults such as individuals with a learning disability or cognitive impairment or individuals in a dependent or unequal relationship?** ☐ **Yes** ☒ **No**

How will you ensure that participants in these groups are competent to give consent to take part in this study? *If you have relevant correspondence, please attach it.*

C4 **Will payment or any other incentive, such as gift service or free services, be made to any research participant?**

☒ **Yes** ☐ **No**

If yes, please specify the level of payment to be made and/or the source of the funds/gift/free service to be used.

Participants will be paid £10.

Please justify the payment/other incentive you intend to offer.

The payment is to cover travel expenses.

C5 **Recruitment**

(i) Describe how potential participants will be identified:

Any person of 18 years of age or older.

(ii) Describe how potential participants will be approached:

Advertisements will be placed around the UCL campus and email circulars will be sent out, as we expect some potential recruits to be students or staff at UCL.

(iii) Describe how participants will be recruited:

As above

Attach recruitment emails/adverts/webpages. A data protection disclaimer should be included in the text of such literature.

C6 **Will the participants participate on a fully voluntary basis?** ☒ Yes ☐ No

Will UCL students be involved as participants in the research project? ☒ Yes ☐ No

If yes, care must be taken to ensure that they are recruited in such a way that they do not feel any obligation to a teacher or member of staff to participate.

Please state how you will bring to the attention of the participants their right to withdraw from the study without penalty?

This is stated to them verbally and also written on the consent form and information sheet.

C7 **CONSENT**

Please describe the process you will use when seeking and obtaining consent.

Participants will be asked to read an information sheet and sign a consent form online before completing online questionnaires prior to the appointment.

During the day of the experiment, participants will be given a copy of the experimental information sheet and the consent sheet and asked to read them carefully. They will then be asked if they have understood the information and if they have any questions.

The participants will then be reminded that they can leave the experiment at any time without giving a reason, and that they will still receive £10 for travel expenses if they do so. If they agree they will then be asked to sign the consent form before the experimental task begins.

A copy of the participant information sheet and consent form must be attached to this application. For your convenience proformas are provided in C10 below. These should be filled in and modified as necessary.

In cases where it is not proposed to obtain the participants informed consent, please explain why below.

C8 **Will any form of deception be used that raises ethical issues? If so, please explain.**

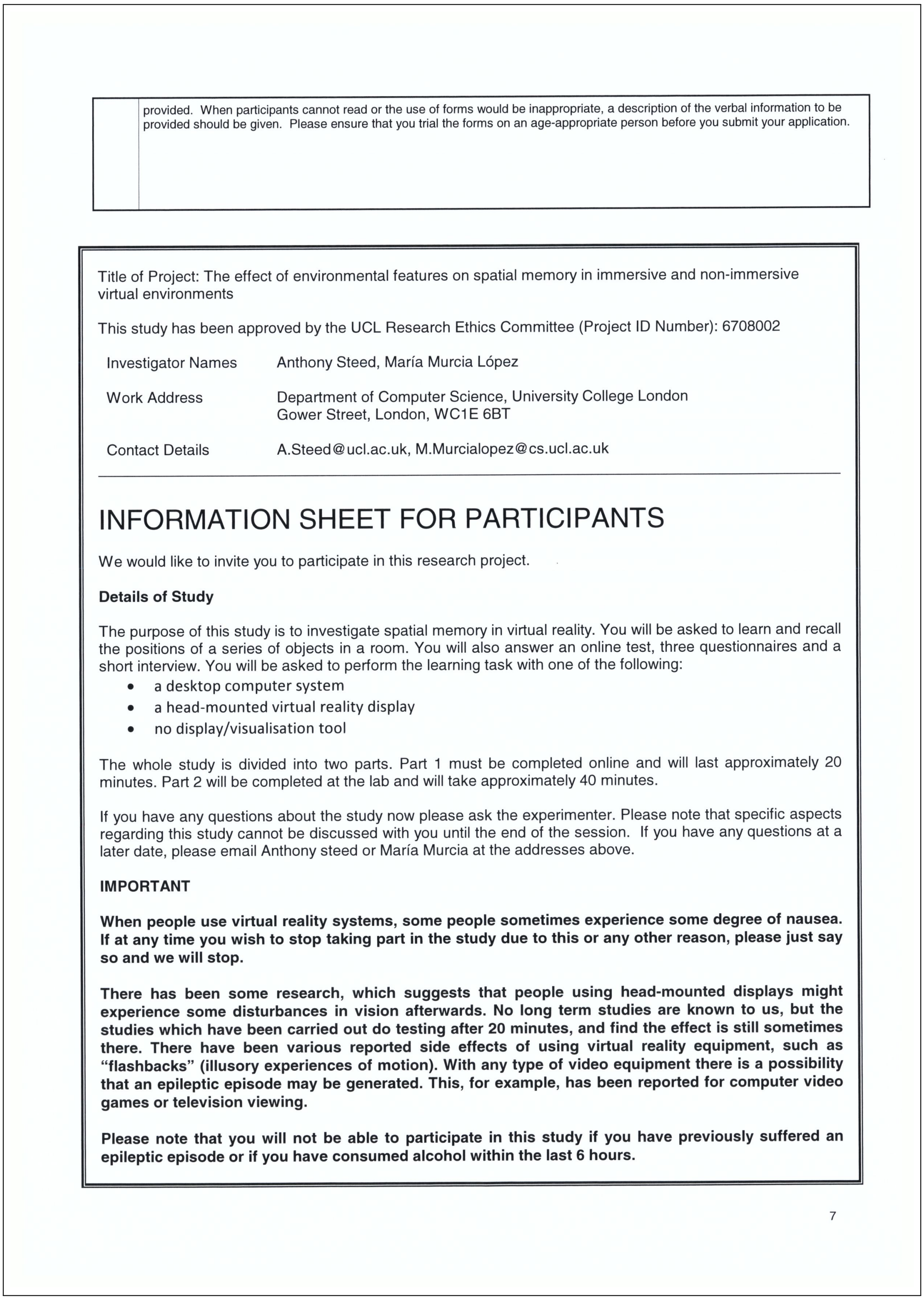
No deception is used that raises ethical issues.

C9 **Will you provide a full debriefing at the end of the data collection phase?** ☒ Yes ☐ No

If 'No', please explain why below.

C10 **Information Sheets And Consent Forms**

A poorly written Information Sheet(s) and Consent Form(s) that lack clarity and simplicity frequently delay ethics approval of research projects. The wording and content of the Information Sheet and Consent Form must be appropriate to the age and educational level of the research participants and clearly state in simple non-technical language what the participant is agreeing to. Use the active voice e.g. "we will book" rather than "bookings will be made". Refer to participants as "you" and yourself as "I" or "we". An appropriate translation of the Forms should be provided where the first language of the participants is not English. If you have different participant groups you should provide Information Sheets and Consent Forms as appropriate (e.g. one for children and one for parents/guardians) using the templates below. Where children are of a reading age, a written Information Sheet should be



provided. When participants cannot read or the use of forms would be inappropriate, a description of the verbal information to be provided should be given. Please ensure that you trial the forms on an age-appropriate person before you submit your application.

Title of Project: The effect of environmental features on spatial memory in immersive and non-immersive virtual environments

This study has been approved by the UCL Research Ethics Committee (Project ID Number): 6708002

Investigator Names	Anthony Steed, María Murcia López
Work Address	Department of Computer Science, University College London Gower Street, London, WC1E 6BT
Contact Details	A.Steed@ucl.ac.uk, M.Murcialopez@cs.ucl.ac.uk

INFORMATION SHEET FOR PARTICIPANTS

We would like to invite you to participate in this research project.

Details of Study

The purpose of this study is to investigate spatial memory in virtual reality. You will be asked to learn and recall the positions of a series of objects in a room. You will also answer an online test, three questionnaires and a short interview. You will be asked to perform the learning task with one of the following:

- a desktop computer system
- a head-mounted virtual reality display
- no display/visualisation tool

The whole study is divided into two parts. Part 1 must be completed online and will last approximately 20 minutes. Part 2 will be completed at the lab and will take approximately 40 minutes.

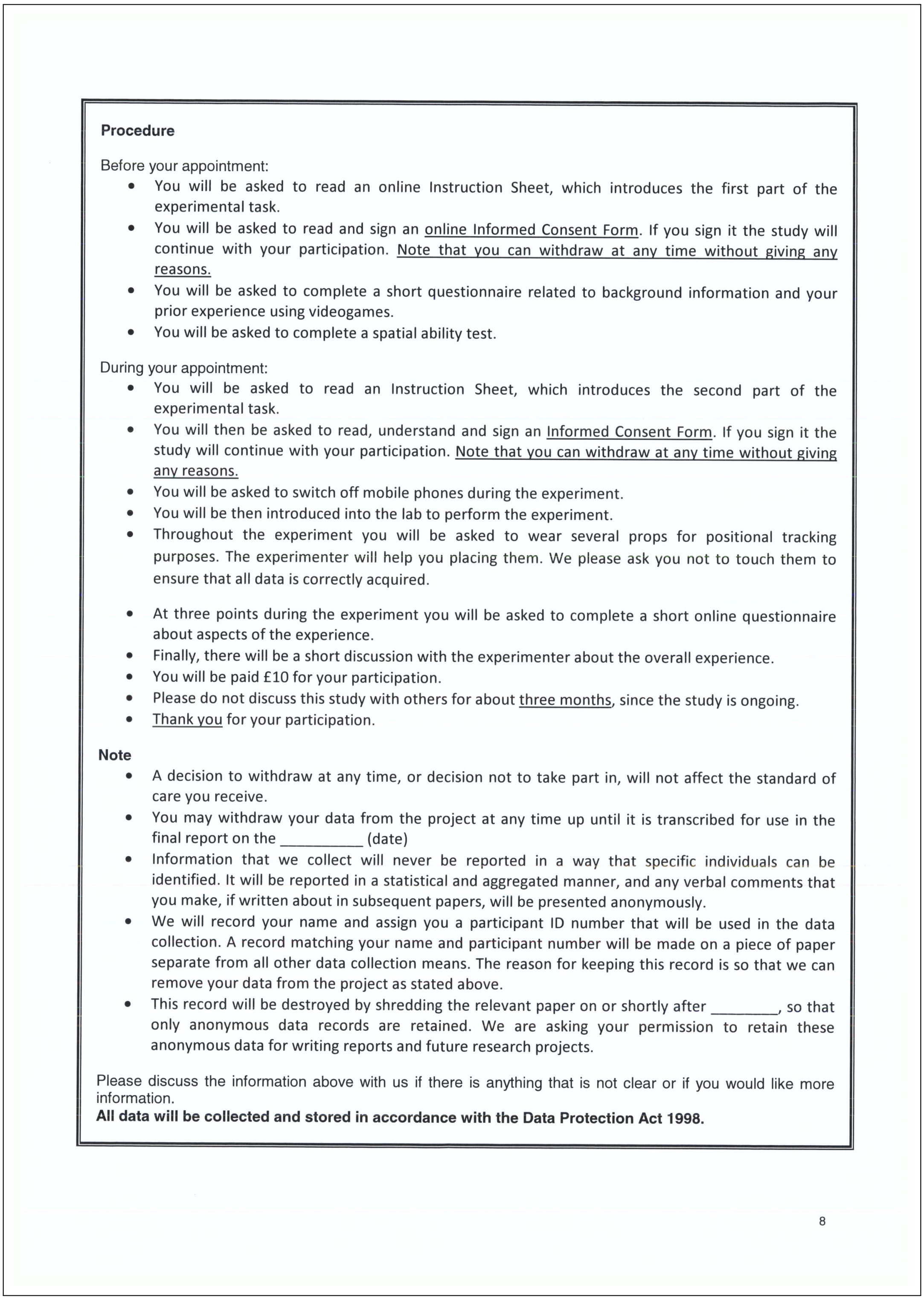
If you have any questions about the study now please ask the experimenter. Please note that specific aspects regarding this study cannot be discussed with you until the end of the session. If you have any questions at a later date, please email Anthony steed or María Murcia at the addresses above.

IMPORTANT

When people use virtual reality systems, some people sometimes experience some degree of nausea. If at any time you wish to stop taking part in the study due to this or any other reason, please just say so and we will stop.

There has been some research, which suggests that people using head-mounted displays might experience some disturbances in vision afterwards. No long term studies are known to us, but the studies which have been carried out do testing after 20 minutes, and find the effect is still sometimes there. There have been various reported side effects of using virtual reality equipment, such as "flashbacks" (illusory experiences of motion). With any type of video equipment there is a possibility that an epileptic episode may be generated. This, for example, has been reported for computer video games or television viewing.

Please note that you will not be able to participate in this study if you have previously suffered an epileptic episode or if you have consumed alcohol within the last 6 hours.



Procedure

Before your appointment:

- You will be asked to read an online Instruction Sheet, which introduces the first part of the experimental task.
- You will be asked to read and sign an online Informed Consent Form. If you sign it the study will continue with your participation. Note that you can withdraw at any time without giving any reasons.
- You will be asked to complete a short questionnaire related to background information and your prior experience using videogames.
- You will be asked to complete a spatial ability test.

During your appointment:

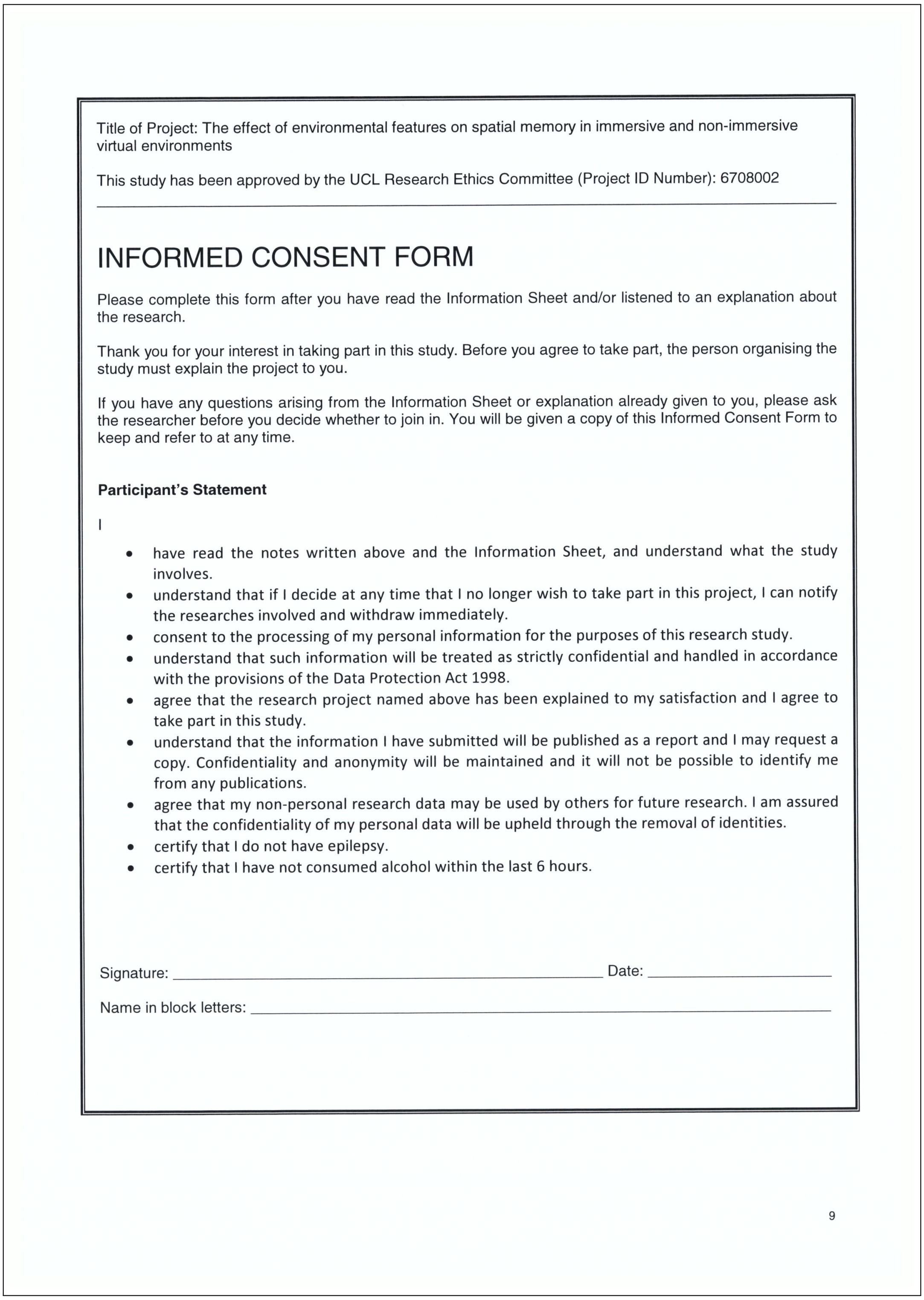
- You will be asked to read an Instruction Sheet, which introduces the second part of the experimental task.
- You will then be asked to read, understand and sign an Informed Consent Form. If you sign it the study will continue with your participation. Note that you can withdraw at any time without giving any reasons.
- You will be asked to switch off mobile phones during the experiment.
- You will be then introduced into the lab to perform the experiment.
- Throughout the experiment you will be asked to wear several props for positional tracking purposes. The experimenter will help you placing them. We please ask you not to touch them to ensure that all data is correctly acquired.
- At three points during the experiment you will be asked to complete a short online questionnaire about aspects of the experience.
- Finally, there will be a short discussion with the experimenter about the overall experience.
- You will be paid £10 for your participation.
- Please do not discuss this study with others for about three months, since the study is ongoing.
- Thank you for your participation.

Note

- A decision to withdraw at any time, or decision not to take part in, will not affect the standard of care you receive.
- You may withdraw your data from the project at any time up until it is transcribed for use in the final report on the __________ (date)
- Information that we collect will never be reported in a way that specific individuals can be identified. It will be reported in a statistical and aggregated manner, and any verbal comments that you make, if written about in subsequent papers, will be presented anonymously.
- We will record your name and assign you a participant ID number that will be used in the data collection. A record matching your name and participant number will be made on a piece of paper separate from all other data collection means. The reason for keeping this record is so that we can remove your data from the project as stated above.
- This record will be destroyed by shredding the relevant paper on or shortly after ________, so that only anonymous data records are retained. We are asking your permission to retain these anonymous data for writing reports and future research projects.

Please discuss the information above with us if there is anything that is not clear or if you would like more information.

All data will be collected and stored in accordance with the Data Protection Act 1998.



Title of Project: The effect of environmental features on spatial memory in immersive and non-immersive virtual environments

This study has been approved by the UCL Research Ethics Committee (Project ID Number): 6708002

INFORMED CONSENT FORM

Please complete this form after you have read the Information Sheet and/or listened to an explanation about the research.

Thank you for your interest in taking part in this study. Before you agree to take part, the person organising the study must explain the project to you.

If you have any questions arising from the Information Sheet or explanation already given to you, please ask the researcher before you decide whether to join in. You will be given a copy of this Informed Consent Form to keep and refer to at any time.

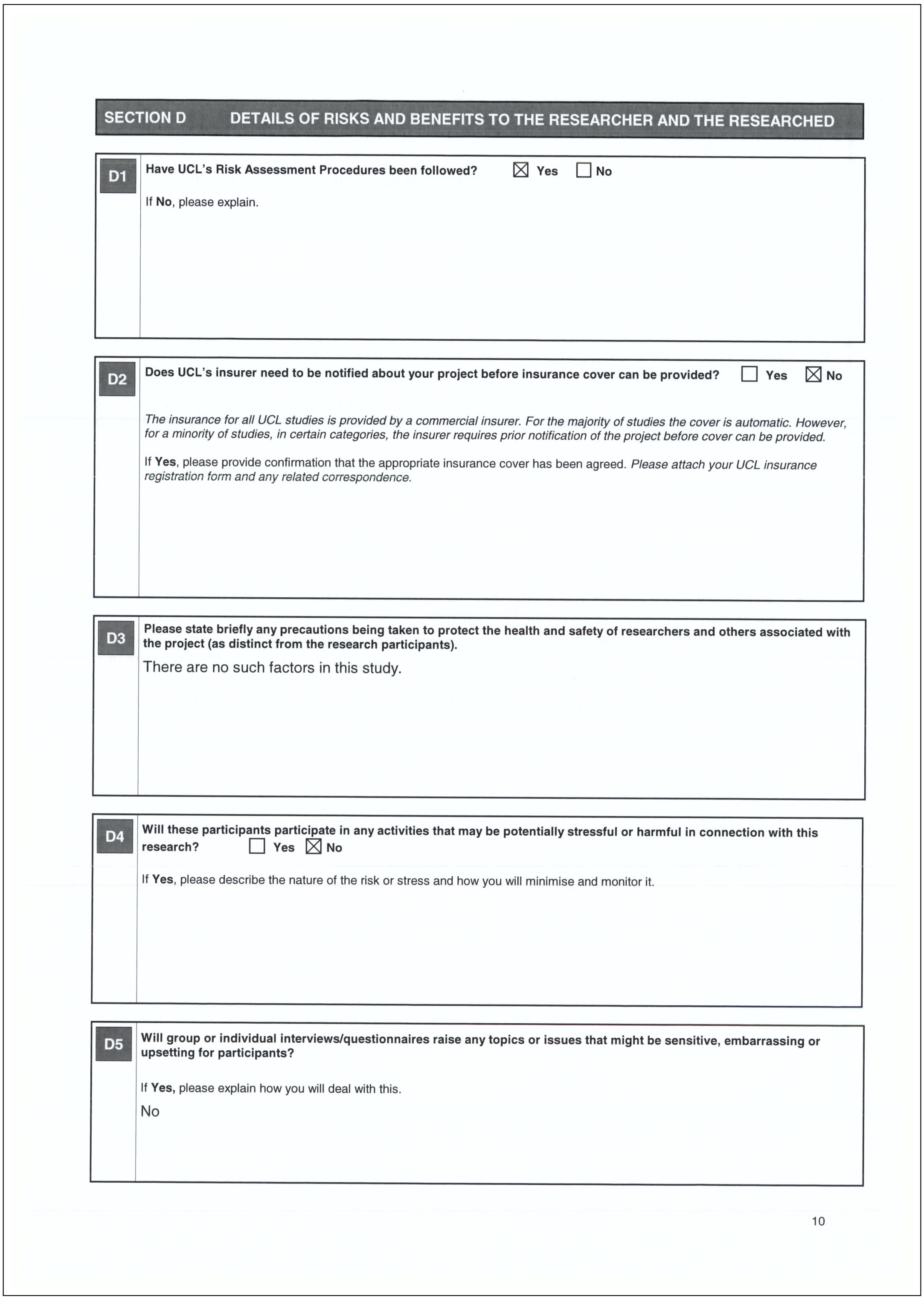
Participant's Statement

I

- have read the notes written above and the Information Sheet, and understand what the study involves.
- understand that if I decide at any time that I no longer wish to take part in this project, I can notify the researches involved and withdraw immediately.
- consent to the processing of my personal information for the purposes of this research study.
- understand that such information will be treated as strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.
- agree that the research project named above has been explained to my satisfaction and I agree to take part in this study.
- understand that the information I have submitted will be published as a report and I may request a copy. Confidentiality and anonymity will be maintained and it will not be possible to identify me from any publications.
- agree that my non-personal research data may be used by others for future research. I am assured that the confidentiality of my personal data will be upheld through the removal of identities.
- certify that I do not have epilepsy.
- certify that I have not consumed alcohol within the last 6 hours.

Signature: ______________________________ Date: ________________

Name in block letters: ______________________________



SECTION D DETAILS OF RISKS AND BENEFITS TO THE RESEARCHER AND THE RESEARCHED

D1 **Have UCL's Risk Assessment Procedures been followed?** ☒ **Yes** ☐ **No**

If **No**, please explain.

D2 **Does UCL's insurer need to be notified about your project before insurance cover can be provided?** ☐ **Yes** ☒ **No**

The insurance for all UCL studies is provided by a commercial insurer. For the majority of studies the cover is automatic. However, for a minority of studies, in certain categories, the insurer requires prior notification of the project before cover can be provided.

If **Yes**, please provide confirmation that the appropriate insurance cover has been agreed. *Please attach your UCL insurance registration form and any related correspondence.*

D3 **Please state briefly any precautions being taken to protect the health and safety of researchers and others associated with the project (as distinct from the research participants).**

There are no such factors in this study.

D4 **Will these participants participate in any activities that may be potentially stressful or harmful in connection with this research?** ☐ **Yes** ☒ **No**

If **Yes**, please describe the nature of the risk or stress and how you will minimise and monitor it.

D5 **Will group or individual interviews/questionnaires raise any topics or issues that might be sensitive, embarrassing or upsetting for participants?**

If **Yes,** please explain how you will deal with this.

No

D6 **Please describe any expected benefits to the participant.**

The participant will experience the use of novel technologies.

D7 **Specify whether the following procedures are involved:**

Any invasive procedure(s) ☐ **Yes** ☒ **No**

Physical contact ☐ **Yes** ☒ **No**

Any procedure(s) that may cause mental distress ☐ **Yes** ☒ **No**

Please state briefly any precautions being taken to protect the health and safety of the research participants.

D8 **Does the research involve the use of drugs?** ☐ **Yes** ☒ **No**

If **Yes**, please name the drug/product and its intended use in the research and then complete Appendix I

Does the project involve the use of genetically modified materials? ☐ **Yes** ☒ **No**

If **Yes**, has approval from the Genetic Modification Safety Committee been obtained for work? ☐ Yes ☐ No

If **Yes**, please quote the Genetic Modification Reference Number:

D9 **Will any non-ionising radiation be used on the research participant(s)?** ☐ **Yes** ☒ **No**

If **Yes**, please complete Appendix II.

D10 **Are you using a medical device in the UK that is CE-marked and is being used within its product indication?** ☐**Yes** ☒ **No**

If **Yes**, please complete Appendix III.

CHECKLIST

Please submit ether 12 copies (1 original + 11 double sided photocopies) of your completed application form for full committee review or 3 copies (1 original + 2 double sided copies) for chair's action, together with the appropriate supporting documentation from the list below to the UCL Research Ethics Committee Administrator. You should also submit your application form electronically to the Administrator at: ethics@ucl.ac.uk

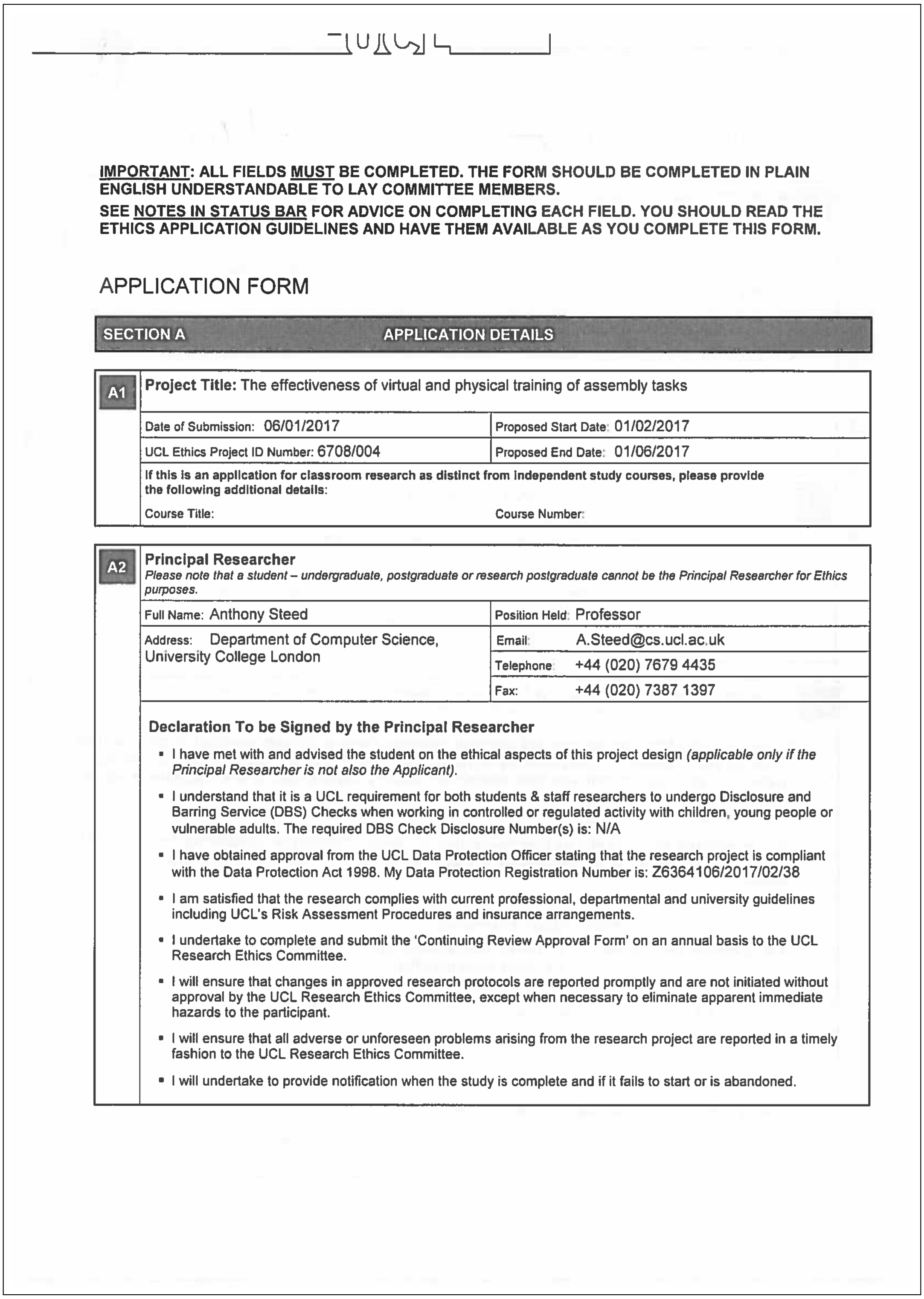
Documents to be Attached to Application Form (if applicable)	**Ticked if attached**	**Tick if not relevant**
Section B: Details of the Project		
• Questionnaire(s) / Psychological Tests	☒	☐
• Relevant correspondence relating to involvement of collaborating department/s and agreed participation in the research.	☐	☒
Section C: Details of Participants		
• Parental/guardian consent form for research involving participants under 18	☐	☒
• Participant/s information sheet	☒	☐
• Participant/s consent form/s	☒	☐
• Advertisement	☒	☐
Section D: Details of Risks and Benefits to the Researcher and the Researched		
• Insurance registration form and related correspondence	☐	☒
Appendix I: Research Involving the Use of Drugs		
• Relevant correspondence relating to agreed arrangements for dispensing with the pharmacy	☐	☒
• Written confirmation from the manufacturer that the drug/substance has has been manufactured to GMP	☐	☒
• Proposed volunteer contract	☐	☒
• Full declaration of financial or direct interest	☐	☒
• Copies of certificates: CTA etc…	☐	☒
Appendix II: Use of Non-Ionising Radiation		
Appendix III: Use Medical Devices		

Please note that correspondence regarding the application will normally be sent to the Principal Researcher and copied to other named individuals.

Appendix D

Ethics Application for A Comparison of Virtual and Physical Training Transfer of Bimanual Assembly Tasks

UCL Research Ethics Committee Application (Project ID: 6708/004) for the study presented in Chapter 6. This application document contains the information sheet as well as the informed consent form that participants had to sign to take part in the study.



IMPORTANT: ALL FIELDS MUST BE COMPLETED. THE FORM SHOULD BE COMPLETED IN PLAIN ENGLISH UNDERSTANDABLE TO LAY COMMITTEE MEMBERS.

SEE NOTES IN STATUS BAR FOR ADVICE ON COMPLETING EACH FIELD. YOU SHOULD READ THE ETHICS APPLICATION GUIDELINES AND HAVE THEM AVAILABLE AS YOU COMPLETE THIS FORM.

APPLICATION FORM

SECTION A APPLICATION DETAILS

A1

Project Title: The effectiveness of virtual and physical training of assembly tasks

Date of Submission: 06/01/2017	Proposed Start Date: 01/02/2017
UCL Ethics Project ID Number: 6708/004	Proposed End Date: 01/06/2017

If this is an application for classroom research as distinct from independent study courses, please provide the following additional details:

Course Title:	Course Number:

A2

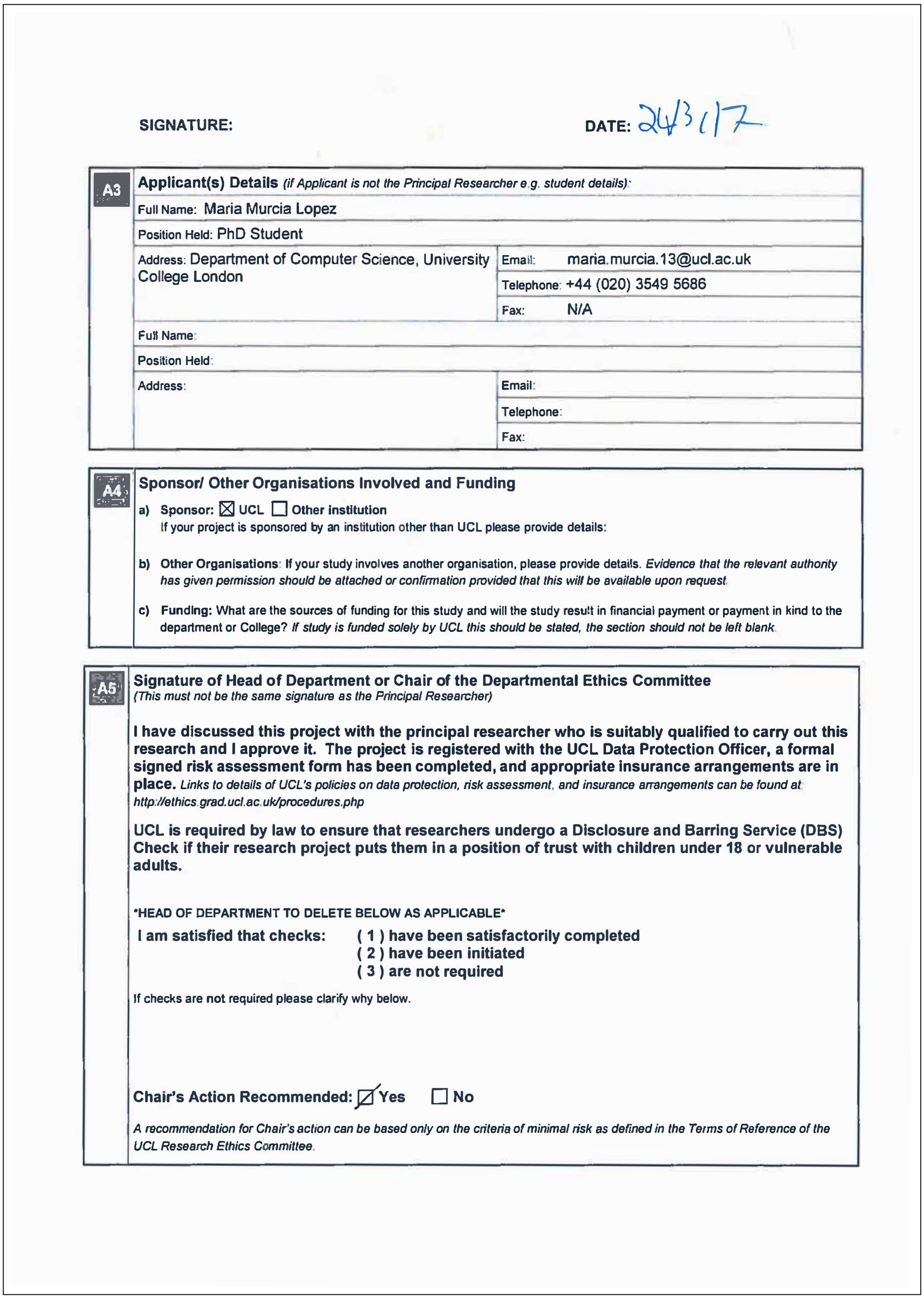
Principal Researcher

Please note that a student – undergraduate, postgraduate or research postgraduate cannot be the Principal Researcher for Ethics purposes.

Full Name: Anthony Steed	Position Held: Professor
Address: Department of Computer Science, University College London	Email: A.Steed@cs.ucl.ac.uk
	Telephone: +44 (020) 7679 4435
	Fax: +44 (020) 7387 1397

Declaration To be Signed by the Principal Researcher

- I have met with and advised the student on the ethical aspects of this project design *(applicable only if the Principal Researcher is not also the Applicant).*
- I understand that it is a UCL requirement for both students & staff researchers to undergo Disclosure and Barring Service (DBS) Checks when working in controlled or regulated activity with children, young people or vulnerable adults. The required DBS Check Disclosure Number(s) is: N/A
- I have obtained approval from the UCL Data Protection Officer stating that the research project is compliant with the Data Protection Act 1998. My Data Protection Registration Number is: Z6364106/2017/02/38
- I am satisfied that the research complies with current professional, departmental and university guidelines including UCL's Risk Assessment Procedures and insurance arrangements.
- I undertake to complete and submit the 'Continuing Review Approval Form' on an annual basis to the UCL Research Ethics Committee.
- I will ensure that changes in approved research protocols are reported promptly and are not initiated without approval by the UCL Research Ethics Committee, except when necessary to eliminate apparent immediate hazards to the participant.
- I will ensure that all adverse or unforeseen problems arising from the research project are reported in a timely fashion to the UCL Research Ethics Committee.
- I will undertake to provide notification when the study is complete and if it fails to start or is abandoned.



SIGNATURE: **DATE:** 24/3/17

A3 **Applicant(s) Details** *(if Applicant is not the Principal Researcher e.g. student details):*

Full Name: Maria Murcia Lopez

Position Held: PhD Student

Address: Department of Computer Science, University College London

Email: maria.murcia.13@ucl.ac.uk

Telephone: +44 (020) 3549 5686

Fax: N/A

Full Name:

Position Held:

Address:

Email:

Telephone:

Fax:

A4 **Sponsor/ Other Organisations Involved and Funding**

a) Sponsor: ☒ **UCL** ☐ **Other institution**
If your project is sponsored by an institution other than UCL please provide details:

b) Other Organisations: If your study involves another organisation, please provide details. *Evidence that the relevant authority has given permission should be attached or confirmation provided that this will be available upon request.*

c) Funding: What are the sources of funding for this study and will the study result in financial payment or payment in kind to the department or College? *If study is funded solely by UCL this should be stated, the section should not be left blank.*

A5 **Signature of Head of Department or Chair of the Departmental Ethics Committee**
(This must not be the same signature as the Principal Researcher)

I have discussed this project with the principal researcher who is suitably qualified to carry out this research and I approve it. The project is registered with the UCL Data Protection Officer, a formal signed risk assessment form has been completed, and appropriate insurance arrangements are in place. *Links to details of UCL's policies on data protection, risk assessment, and insurance arrangements can be found at: http://ethics.grad.ucl.ac.uk/procedures.php*

UCL is required by law to ensure that researchers undergo a Disclosure and Barring Service (DBS) Check if their research project puts them in a position of trust with children under 18 or vulnerable adults.

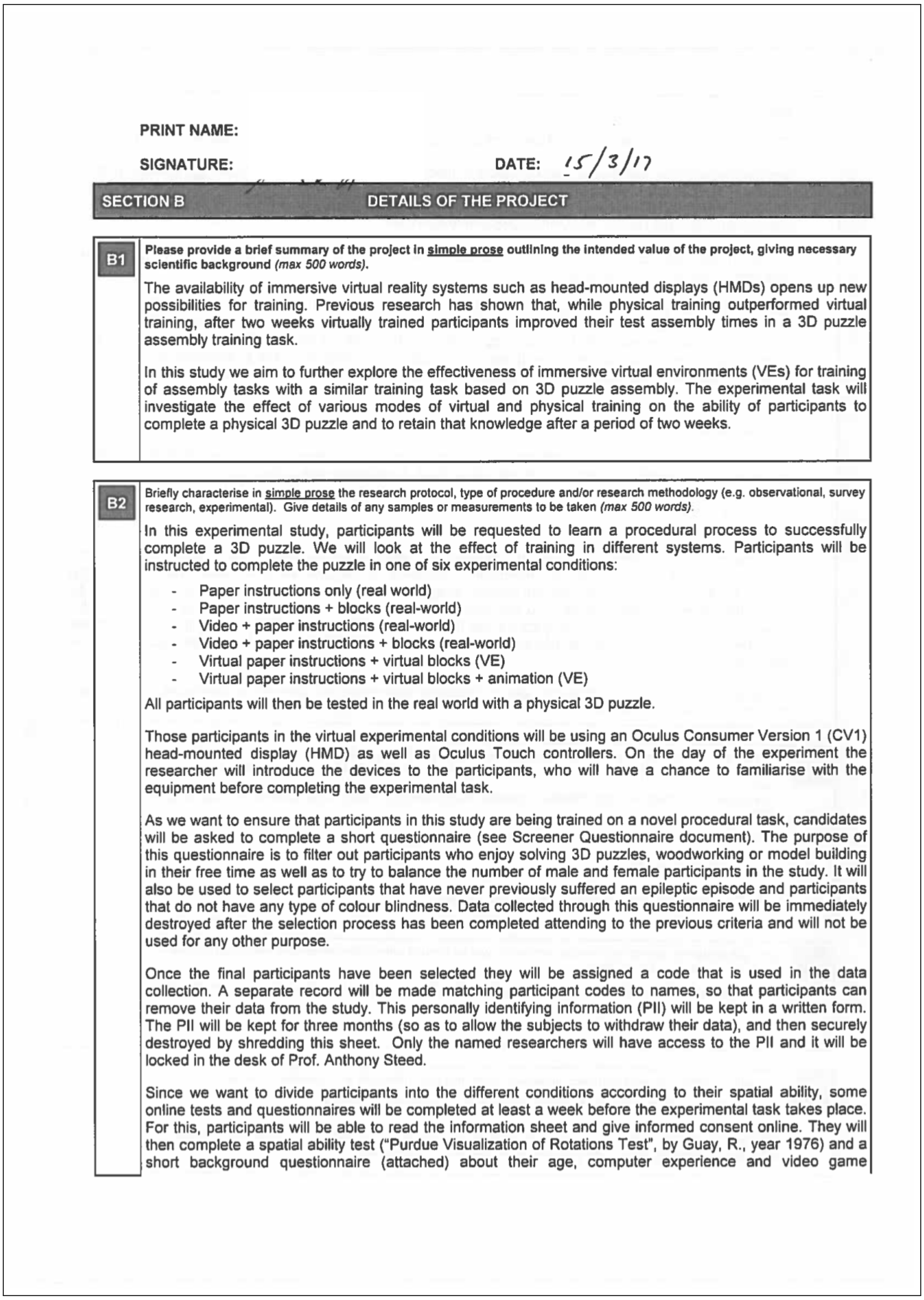
HEAD OF DEPARTMENT TO DELETE BELOW AS APPLICABLE

I am satisfied that checks:
(1) have been satisfactorily completed
(2) have been initiated
(3) are not required

If checks are not required please clarify why below.

Chair's Action Recommended: ☑ **Yes** ☐ **No**

A recommendation for Chair's action can be based only on the criteria of minimal risk as defined in the Terms of Reference of the UCL Research Ethics Committee.



PRINT NAME:

SIGNATURE: **DATE:** 15/3/17

SECTION B **DETAILS OF THE PROJECT**

B1 **Please provide a brief summary of the project in simple prose outlining the intended value of the project, giving necessary scientific background** *(max 500 words).*

The availability of immersive virtual reality systems such as head-mounted displays (HMDs) opens up new possibilities for training. Previous research has shown that, while physical training outperformed virtual training, after two weeks virtually trained participants improved their test assembly times in a 3D puzzle assembly training task.

In this study we aim to further explore the effectiveness of immersive virtual environments (VEs) for training of assembly tasks with a similar training task based on 3D puzzle assembly. The experimental task will investigate the effect of various modes of virtual and physical training on the ability of participants to complete a physical 3D puzzle and to retain that knowledge after a period of two weeks.

B2 **Briefly characterise in simple prose the research protocol, type of procedure and/or research methodology (e.g. observational, survey research, experimental). Give details of any samples or measurements to be taken** *(max 500 words).*

In this experimental study, participants will be requested to learn a procedural process to successfully complete a 3D puzzle. We will look at the effect of training in different systems. Participants will be instructed to complete the puzzle in one of six experimental conditions:

- Paper instructions only (real world)
- Paper instructions + blocks (real-world)
- Video + paper instructions (real-world)
- Video + paper instructions + blocks (real-world)
- Virtual paper instructions + virtual blocks (VE)
- Virtual paper instructions + virtual blocks + animation (VE)

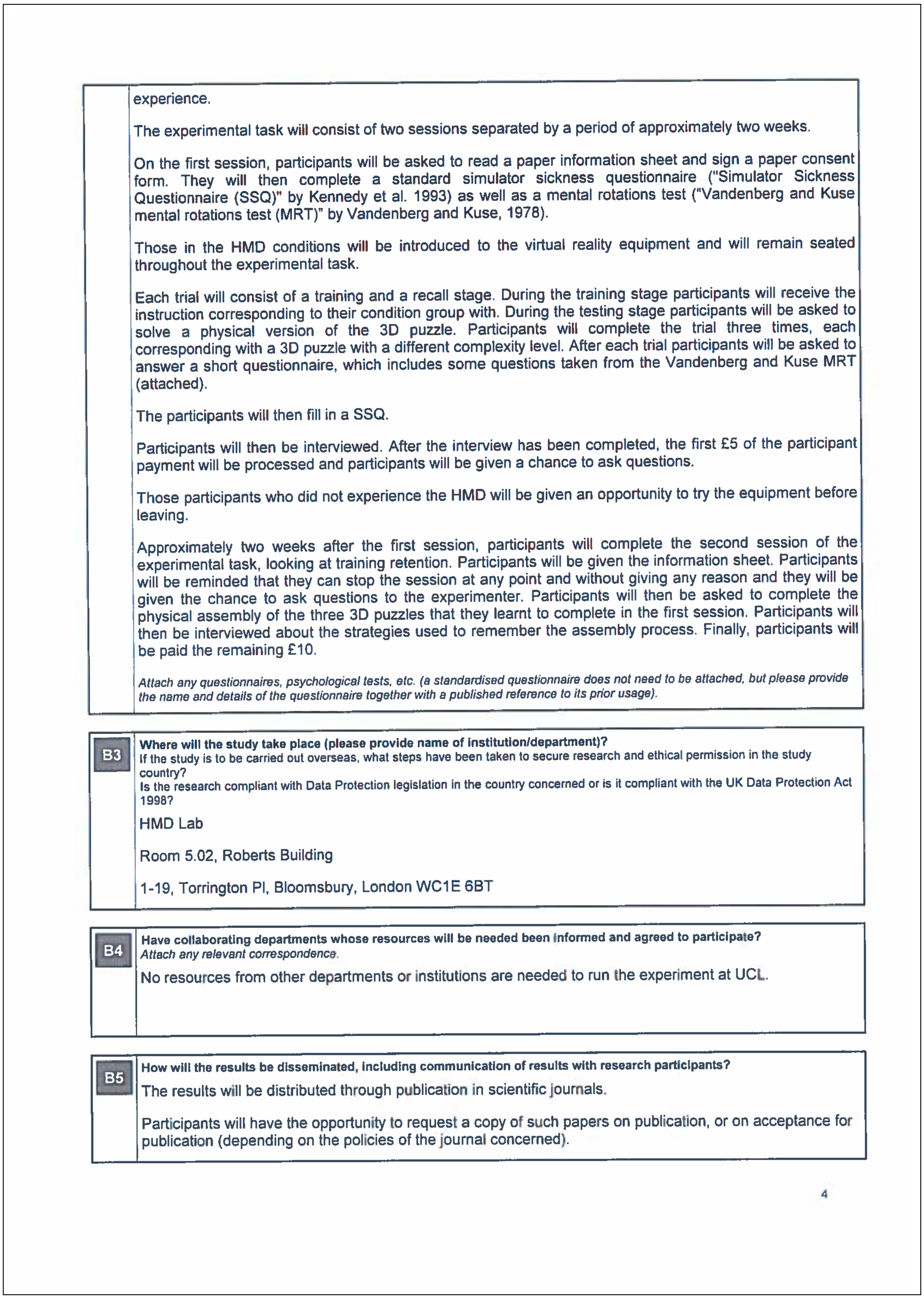
All participants will then be tested in the real world with a physical 3D puzzle.

Those participants in the virtual experimental conditions will be using an Oculus Consumer Version 1 (CV1) head-mounted display (HMD) as well as Oculus Touch controllers. On the day of the experiment the researcher will introduce the devices to the participants, who will have a chance to familiarise with the equipment before completing the experimental task.

As we want to ensure that participants in this study are being trained on a novel procedural task, candidates will be asked to complete a short questionnaire (see Screener Questionnaire document). The purpose of this questionnaire is to filter out participants who enjoy solving 3D puzzles, woodworking or model building in their free time as well as to try to balance the number of male and female participants in the study. It will also be used to select participants that have never previously suffered an epileptic episode and participants that do not have any type of colour blindness. Data collected through this questionnaire will be immediately destroyed after the selection process has been completed attending to the previous criteria and will not be used for any other purpose.

Once the final participants have been selected they will be assigned a code that is used in the data collection. A separate record will be made matching participant codes to names, so that participants can remove their data from the study. This personally identifying information (PII) will be kept in a written form. The PII will be kept for three months (so as to allow the subjects to withdraw their data), and then securely destroyed by shredding this sheet. Only the named researchers will have access to the PII and it will be locked in the desk of Prof. Anthony Steed.

Since we want to divide participants into the different conditions according to their spatial ability, some online tests and questionnaires will be completed at least a week before the experimental task takes place. For this, participants will be able to read the information sheet and give informed consent online. They will then complete a spatial ability test ("Purdue Visualization of Rotations Test", by Guay, R., year 1976) and a short background questionnaire (attached) about their age, computer experience and video game



experience.

The experimental task will consist of two sessions separated by a period of approximately two weeks.

On the first session, participants will be asked to read a paper information sheet and sign a paper consent form. They will then complete a standard simulator sickness questionnaire ("Simulator Sickness Questionnaire (SSQ)" by Kennedy et al. 1993) as well as a mental rotations test ("Vandenberg and Kuse mental rotations test (MRT)" by Vandenberg and Kuse, 1978).

Those in the HMD conditions will be introduced to the virtual reality equipment and will remain seated throughout the experimental task.

Each trial will consist of a training and a recall stage. During the training stage participants will receive the instruction corresponding to their condition group with. During the testing stage participants will be asked to solve a physical version of the 3D puzzle. Participants will complete the trial three times, each corresponding with a 3D puzzle with a different complexity level. After each trial participants will be asked to answer a short questionnaire, which includes some questions taken from the Vandenberg and Kuse MRT (attached).

The participants will then fill in a SSQ.

Participants will then be interviewed. After the interview has been completed, the first £5 of the participant payment will be processed and participants will be given a chance to ask questions.

Those participants who did not experience the HMD will be given an opportunity to try the equipment before leaving.

Approximately two weeks after the first session, participants will complete the second session of the experimental task, looking at training retention. Participants will be given the information sheet. Participants will be reminded that they can stop the session at any point and without giving any reason and they will be given the chance to ask questions to the experimenter. Participants will then be asked to complete the physical assembly of the three 3D puzzles that they learnt to complete in the first session. Participants will then be interviewed about the strategies used to remember the assembly process. Finally, participants will be paid the remaining £10.

Attach any questionnaires, psychological tests, etc. (a standardised questionnaire does not need to be attached, but please provide the name and details of the questionnaire together with a published reference to its prior usage).

B3 **Where will the study take place (please provide name of institution/department)?**
If the study is to be carried out overseas, what steps have been taken to secure research and ethical permission in the study country?
Is the research compliant with Data Protection legislation in the country concerned or is it compliant with the UK Data Protection Act 1998?

HMD Lab

Room 5.02, Roberts Building

1-19, Torrington Pl, Bloomsbury, London WC1E 6BT

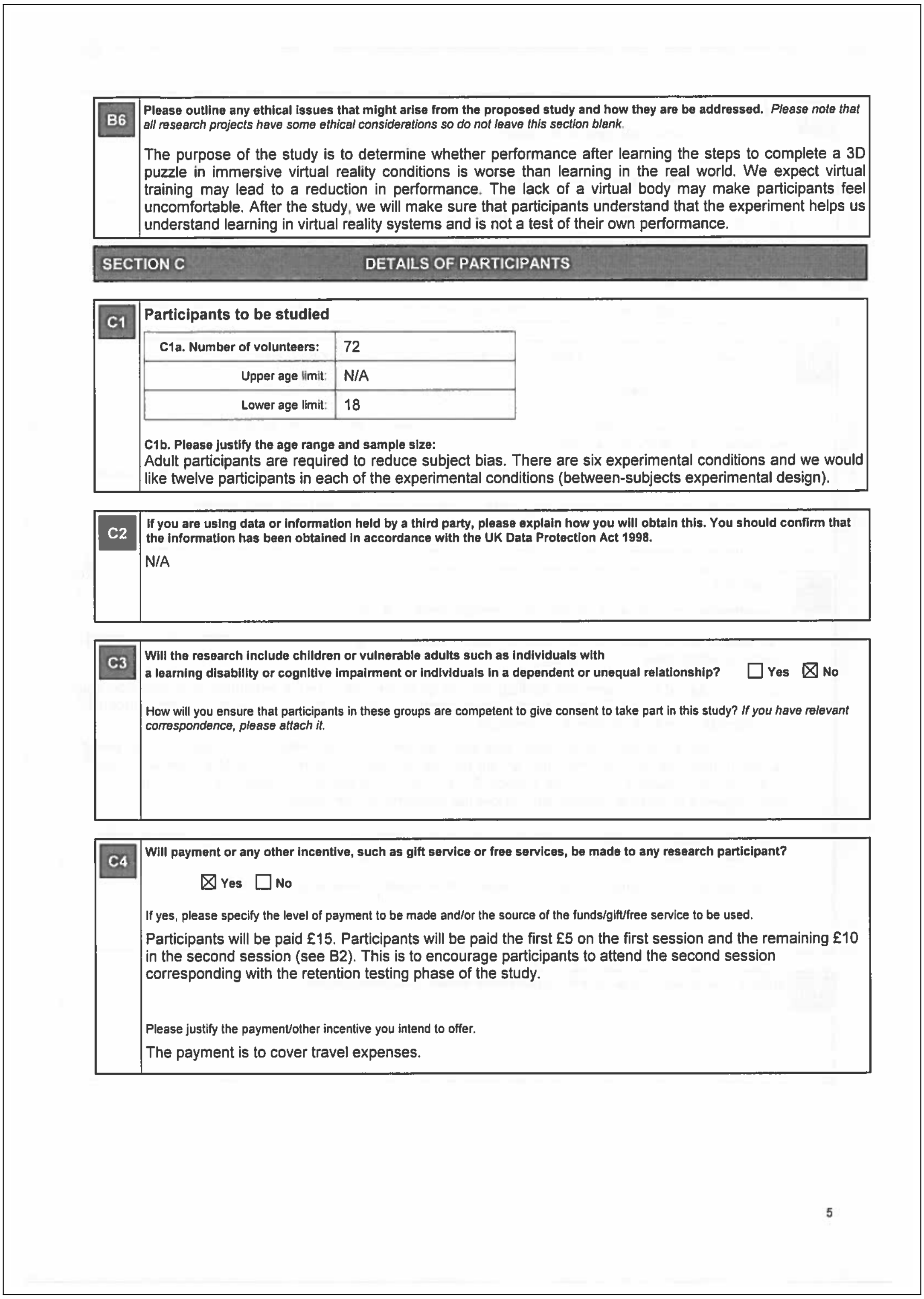
B4 **Have collaborating departments whose resources will be needed been informed and agreed to participate?**
Attach any relevant correspondence.

No resources from other departments or institutions are needed to run the experiment at UCL.

B5 **How will the results be disseminated, including communication of results with research participants?**

The results will be distributed through publication in scientific journals.

Participants will have the opportunity to request a copy of such papers on publication, or on acceptance for publication (depending on the policies of the journal concerned).



B6 **Please outline any ethical issues that might arise from the proposed study and how they are be addressed.** *Please note that all research projects have some ethical considerations so do not leave this section blank.*

The purpose of the study is to determine whether performance after learning the steps to complete a 3D puzzle in immersive virtual reality conditions is worse than learning in the real world. We expect virtual training may lead to a reduction in performance. The lack of a virtual body may make participants feel uncomfortable. After the study, we will make sure that participants understand that the experiment helps us understand learning in virtual reality systems and is not a test of their own performance.

SECTION C DETAILS OF PARTICIPANTS

C1 **Participants to be studied**

C1a. Number of volunteers:	72
Upper age limit:	N/A
Lower age limit:	18

C1b. Please justify the age range and sample size:
Adult participants are required to reduce subject bias. There are six experimental conditions and we would like twelve participants in each of the experimental conditions (between-subjects experimental design).

C2 **If you are using data or information held by a third party, please explain how you will obtain this. You should confirm that the information has been obtained in accordance with the UK Data Protection Act 1998.**

N/A

C3 **Will the research include children or vulnerable adults such as individuals with a learning disability or cognitive impairment or individuals in a dependent or unequal relationship?** ☐ **Yes** ☒ **No**

How will you ensure that participants in these groups are competent to give consent to take part in this study? *If you have relevant correspondence, please attach it.*

C4 **Will payment or any other incentive, such as gift service or free services, be made to any research participant?**

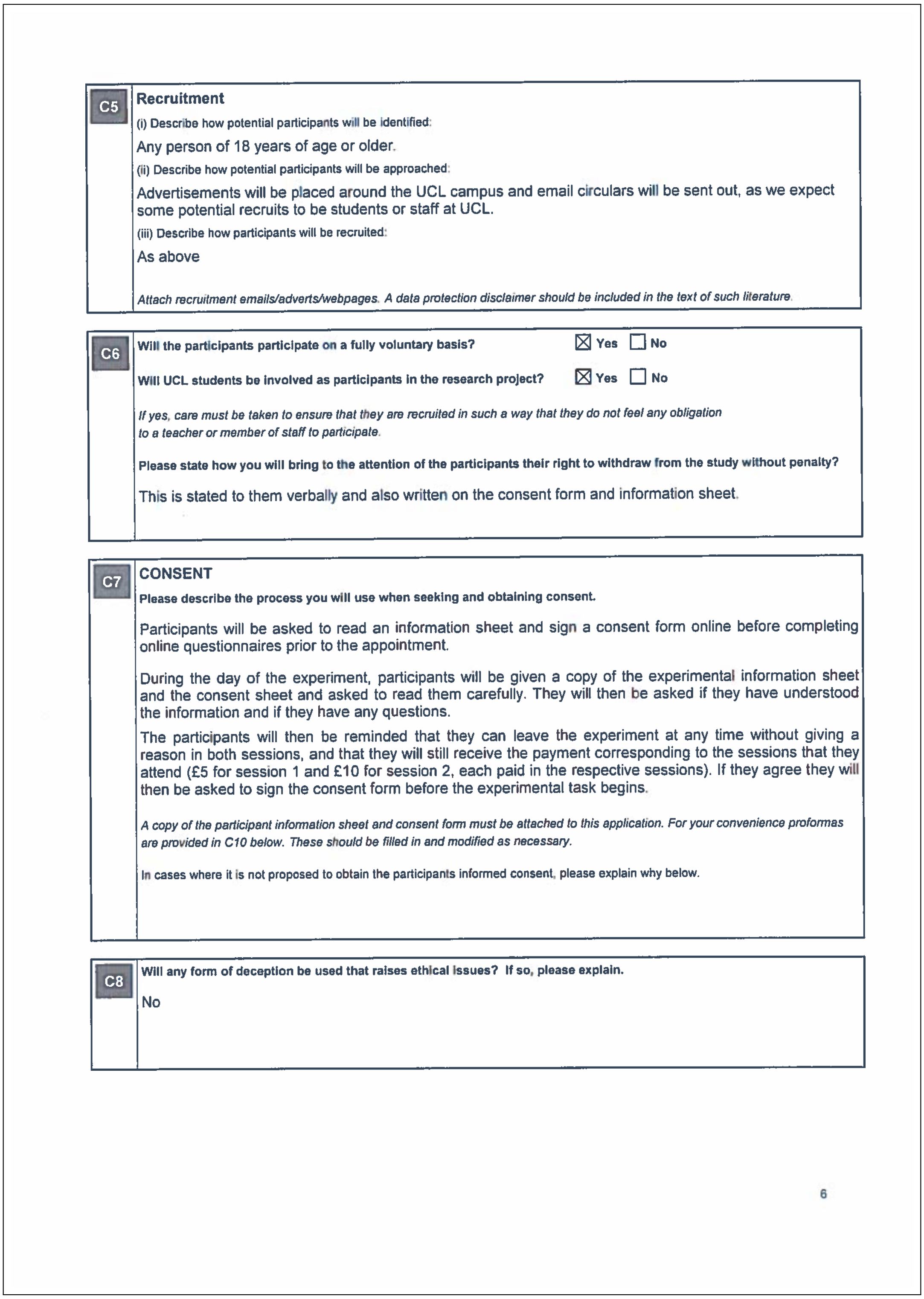
☒ **Yes** ☐ **No**

If yes, please specify the level of payment to be made and/or the source of the funds/gift/free service to be used.

Participants will be paid £15. Participants will be paid the first £5 on the first session and the remaining £10 in the second session (see B2). This is to encourage participants to attend the second session corresponding with the retention testing phase of the study.

Please justify the payment/other incentive you intend to offer.

The payment is to cover travel expenses.



C5 **Recruitment**

(i) Describe how potential participants will be identified:

Any person of 18 years of age or older.

(ii) Describe how potential participants will be approached:

Advertisements will be placed around the UCL campus and email circulars will be sent out, as we expect some potential recruits to be students or staff at UCL.

(iii) Describe how participants will be recruited:

As above

Attach recruitment emails/adverts/webpages. A data protection disclaimer should be included in the text of such literature.

C6 **Will the participants participate on a fully voluntary basis?** ☒ **Yes** ☐ **No**

Will UCL students be involved as participants in the research project? ☒ **Yes** ☐ **No**

If yes, care must be taken to ensure that they are recruited in such a way that they do not feel any obligation to a teacher or member of staff to participate.

Please state how you will bring to the attention of the participants their right to withdraw from the study without penalty?

This is stated to them verbally and also written on the consent form and information sheet.

C7 **CONSENT**

Please describe the process you will use when seeking and obtaining consent.

Participants will be asked to read an information sheet and sign a consent form online before completing online questionnaires prior to the appointment.

During the day of the experiment, participants will be given a copy of the experimental information sheet and the consent sheet and asked to read them carefully. They will then be asked if they have understood the information and if they have any questions.

The participants will then be reminded that they can leave the experiment at any time without giving a reason in both sessions, and that they will still receive the payment corresponding to the sessions that they attend (£5 for session 1 and £10 for session 2, each paid in the respective sessions). If they agree they will then be asked to sign the consent form before the experimental task begins.

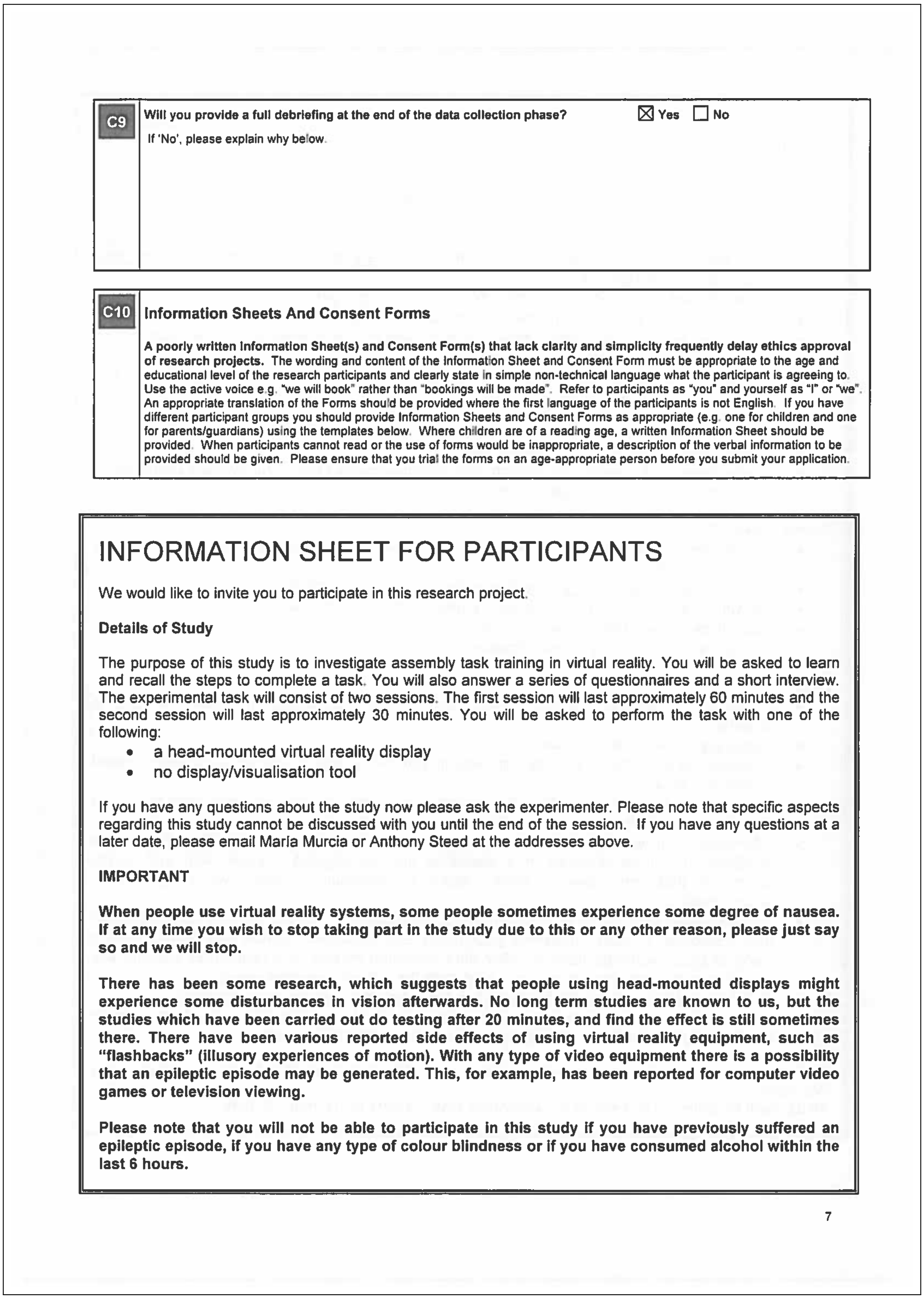
A copy of the participant information sheet and consent form must be attached to this application. For your convenience proformas are provided in C10 below. These should be filled in and modified as necessary.

In cases where it is not proposed to obtain the participants informed consent, please explain why below.

C8 **Will any form of deception be used that raises ethical issues? If so, please explain.**

No

6



C9 **Will you provide a full debriefing at the end of the data collection phase?** ☒ Yes ☐ No

If 'No', please explain why below.

C10 **Information Sheets And Consent Forms**

A poorly written Information Sheet(s) and Consent Form(s) that lack clarity and simplicity frequently delay ethics approval of research projects. The wording and content of the Information Sheet and Consent Form must be appropriate to the age and educational level of the research participants and clearly state in simple non-technical language what the participant is agreeing to. Use the active voice e.g. "we will book" rather than "bookings will be made". Refer to participants as "you" and yourself as "I" or "we". An appropriate translation of the Forms should be provided where the first language of the participants is not English. If you have different participant groups you should provide Information Sheets and Consent Forms as appropriate (e.g. one for children and one for parents/guardians) using the templates below. Where children are of a reading age, a written Information Sheet should be provided. When participants cannot read or the use of forms would be inappropriate, a description of the verbal information to be provided should be given. Please ensure that you trial the forms on an age-appropriate person before you submit your application.

INFORMATION SHEET FOR PARTICIPANTS

We would like to invite you to participate in this research project.

Details of Study

The purpose of this study is to investigate assembly task training in virtual reality. You will be asked to learn and recall the steps to complete a task. You will also answer a series of questionnaires and a short interview. The experimental task will consist of two sessions. The first session will last approximately 60 minutes and the second session will last approximately 30 minutes. You will be asked to perform the task with one of the following:

- a head-mounted virtual reality display
- no display/visualisation tool

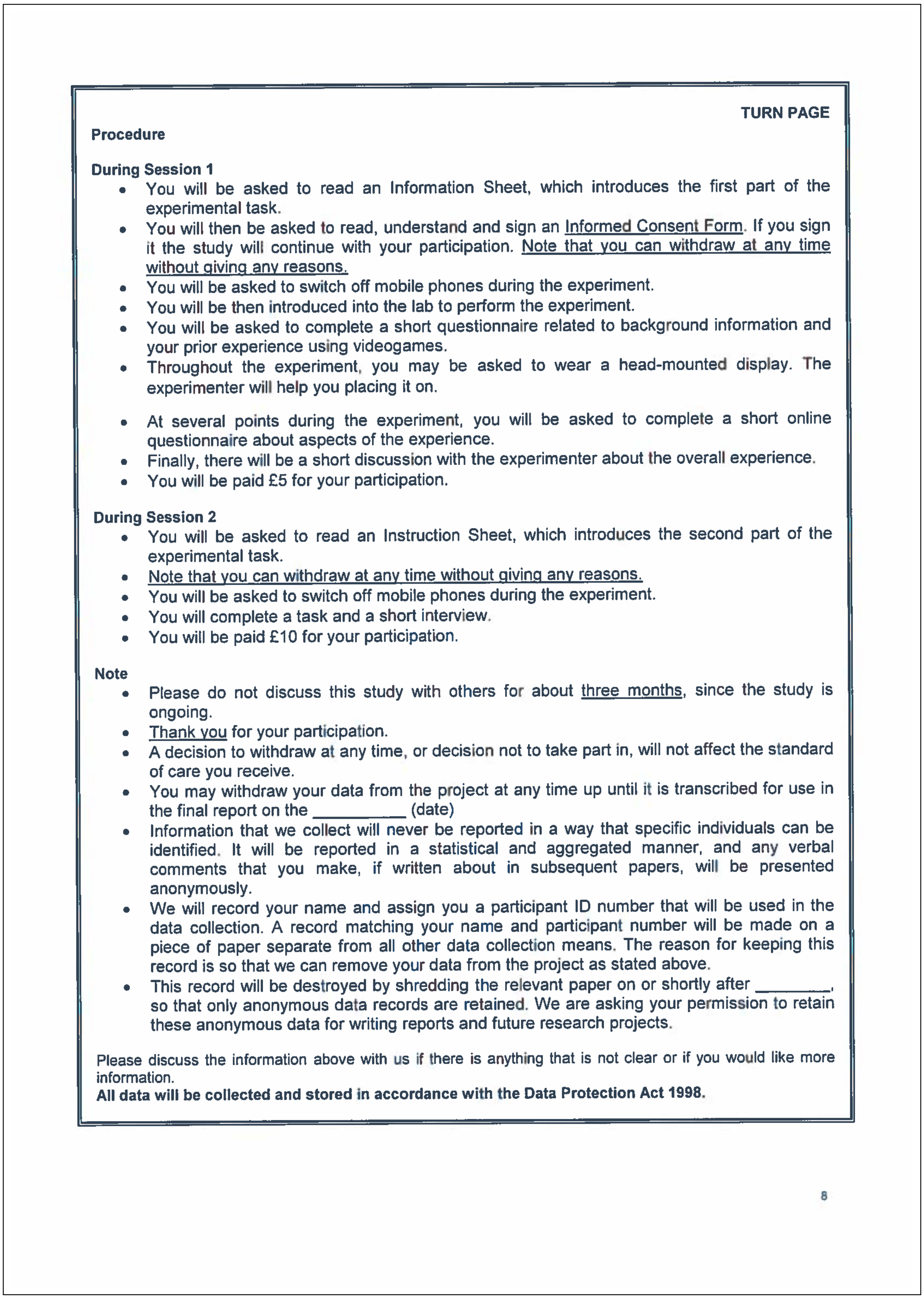
If you have any questions about the study now please ask the experimenter. Please note that specific aspects regarding this study cannot be discussed with you until the end of the session. If you have any questions at a later date, please email Maria Murcia or Anthony Steed at the addresses above.

IMPORTANT

When people use virtual reality systems, some people sometimes experience some degree of nausea. If at any time you wish to stop taking part in the study due to this or any other reason, please just say so and we will stop.

There has been some research, which suggests that people using head-mounted displays might experience some disturbances in vision afterwards. No long term studies are known to us, but the studies which have been carried out do testing after 20 minutes, and find the effect is still sometimes there. There have been various reported side effects of using virtual reality equipment, such as "flashbacks" (illusory experiences of motion). With any type of video equipment there is a possibility that an epileptic episode may be generated. This, for example, has been reported for computer video games or television viewing.

Please note that you will not be able to participate in this study if you have previously suffered an epileptic episode, if you have any type of colour blindness or if you have consumed alcohol within the last 6 hours.



TURN PAGE

Procedure

During Session 1

- You will be asked to read an Information Sheet, which introduces the first part of the experimental task.
- You will then be asked to read, understand and sign an Informed Consent Form. If you sign it the study will continue with your participation. Note that you can withdraw at any time without giving any reasons.
- You will be asked to switch off mobile phones during the experiment.
- You will be then introduced into the lab to perform the experiment.
- You will be asked to complete a short questionnaire related to background information and your prior experience using videogames.
- Throughout the experiment, you may be asked to wear a head-mounted display. The experimenter will help you placing it on.
- At several points during the experiment, you will be asked to complete a short online questionnaire about aspects of the experience.
- Finally, there will be a short discussion with the experimenter about the overall experience.
- You will be paid £5 for your participation.

During Session 2

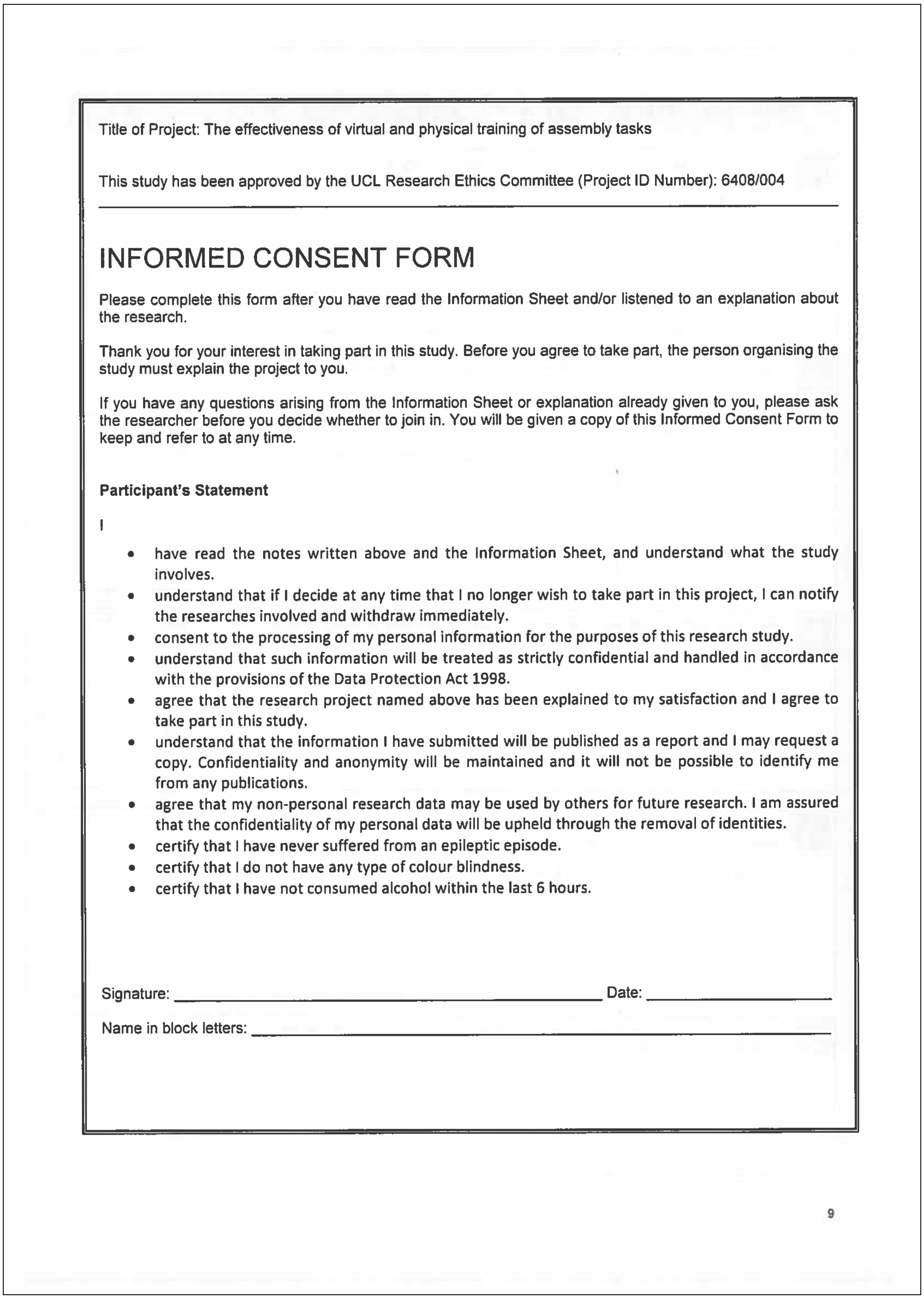
- You will be asked to read an Instruction Sheet, which introduces the second part of the experimental task.
- Note that you can withdraw at any time without giving any reasons.
- You will be asked to switch off mobile phones during the experiment.
- You will complete a task and a short interview.
- You will be paid £10 for your participation.

Note

- Please do not discuss this study with others for about three months, since the study is ongoing.
- Thank you for your participation.
- A decision to withdraw at any time, or decision not to take part in, will not affect the standard of care you receive.
- You may withdraw your data from the project at any time up until it is transcribed for use in the final report on the __________ (date)
- Information that we collect will never be reported in a way that specific individuals can be identified. It will be reported in a statistical and aggregated manner, and any verbal comments that you make, if written about in subsequent papers, will be presented anonymously.
- We will record your name and assign you a participant ID number that will be used in the data collection. A record matching your name and participant number will be made on a piece of paper separate from all other data collection means. The reason for keeping this record is so that we can remove your data from the project as stated above.
- This record will be destroyed by shredding the relevant paper on or shortly after ________, so that only anonymous data records are retained. We are asking your permission to retain these anonymous data for writing reports and future research projects.

Please discuss the information above with us if there is anything that is not clear or if you would like more information.

All data will be collected and stored in accordance with the Data Protection Act 1998.



Title of Project: The effectiveness of virtual and physical training of assembly tasks

This study has been approved by the UCL Research Ethics Committee (Project ID Number): 6408/004

INFORMED CONSENT FORM

Please complete this form after you have read the Information Sheet and/or listened to an explanation about the research.

Thank you for your interest in taking part in this study. Before you agree to take part, the person organising the study must explain the project to you.

If you have any questions arising from the Information Sheet or explanation already given to you, please ask the researcher before you decide whether to join in. You will be given a copy of this Informed Consent Form to keep and refer to at any time.

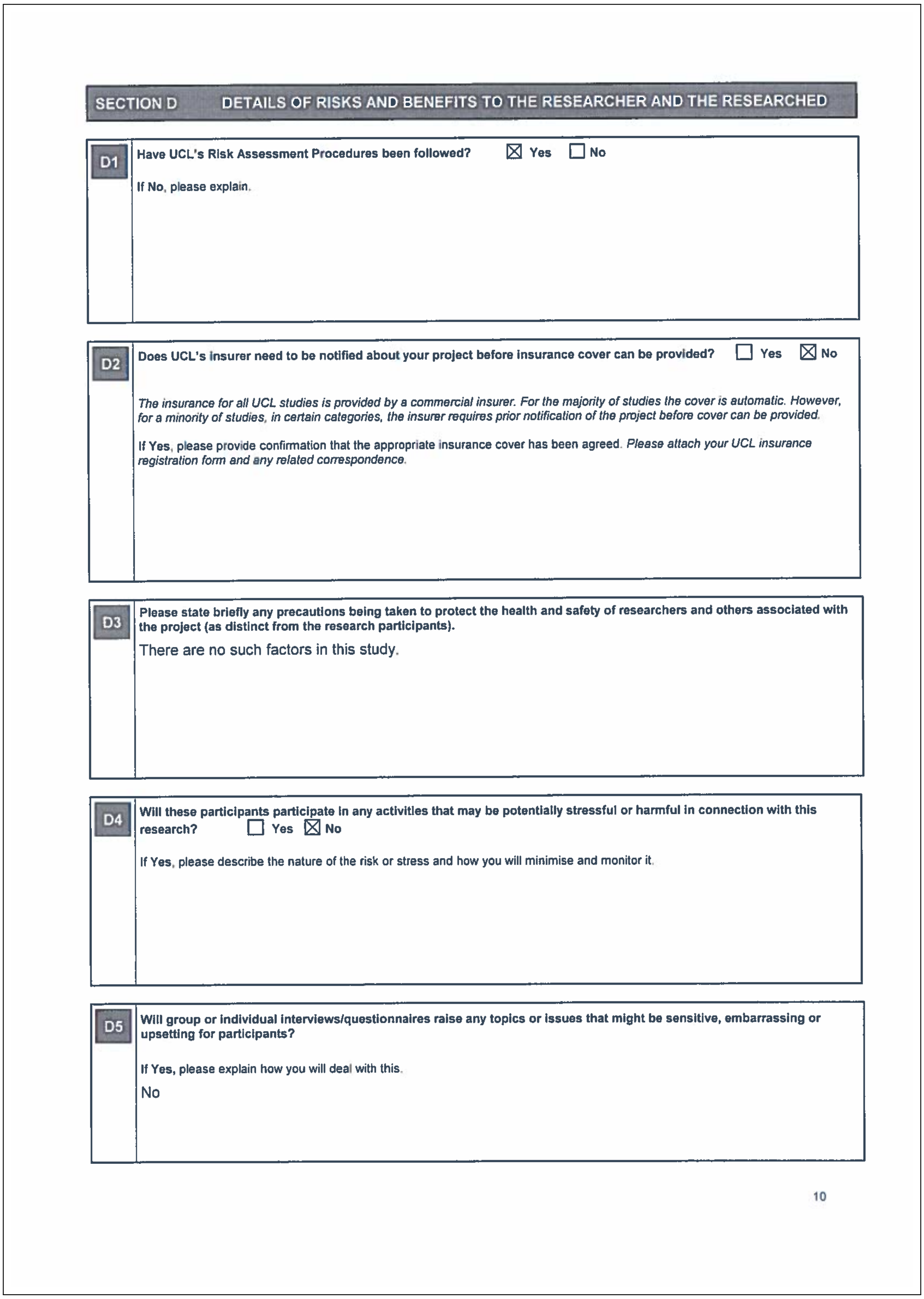
Participant's Statement

I

- have read the notes written above and the Information Sheet, and understand what the study involves.
- understand that if I decide at any time that I no longer wish to take part in this project, I can notify the researches involved and withdraw immediately.
- consent to the processing of my personal information for the purposes of this research study.
- understand that such information will be treated as strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.
- agree that the research project named above has been explained to my satisfaction and I agree to take part in this study.
- understand that the information I have submitted will be published as a report and I may request a copy. Confidentiality and anonymity will be maintained and it will not be possible to identify me from any publications.
- agree that my non-personal research data may be used by others for future research. I am assured that the confidentiality of my personal data will be upheld through the removal of identities.
- certify that I have never suffered from an epileptic episode.
- certify that I do not have any type of colour blindness.
- certify that I have not consumed alcohol within the last 6 hours.

Signature: ______________________________ Date: ______________

Name in block letters: ______________________________



SECTION D DETAILS OF RISKS AND BENEFITS TO THE RESEARCHER AND THE RESEARCHED

D1 **Have UCL's Risk Assessment Procedures been followed?** ☒ **Yes** ☐ **No**

If **No**, please explain.

D2 **Does UCL's insurer need to be notified about your project before insurance cover can be provided?** ☐ **Yes** ☒ **No**

The insurance for all UCL studies is provided by a commercial insurer. For the majority of studies the cover is automatic. However, for a minority of studies, in certain categories, the insurer requires prior notification of the project before cover can be provided.

If **Yes**, please provide confirmation that the appropriate insurance cover has been agreed. *Please attach your UCL insurance registration form and any related correspondence.*

D3 **Please state briefly any precautions being taken to protect the health and safety of researchers and others associated with the project (as distinct from the research participants).**

There are no such factors in this study.

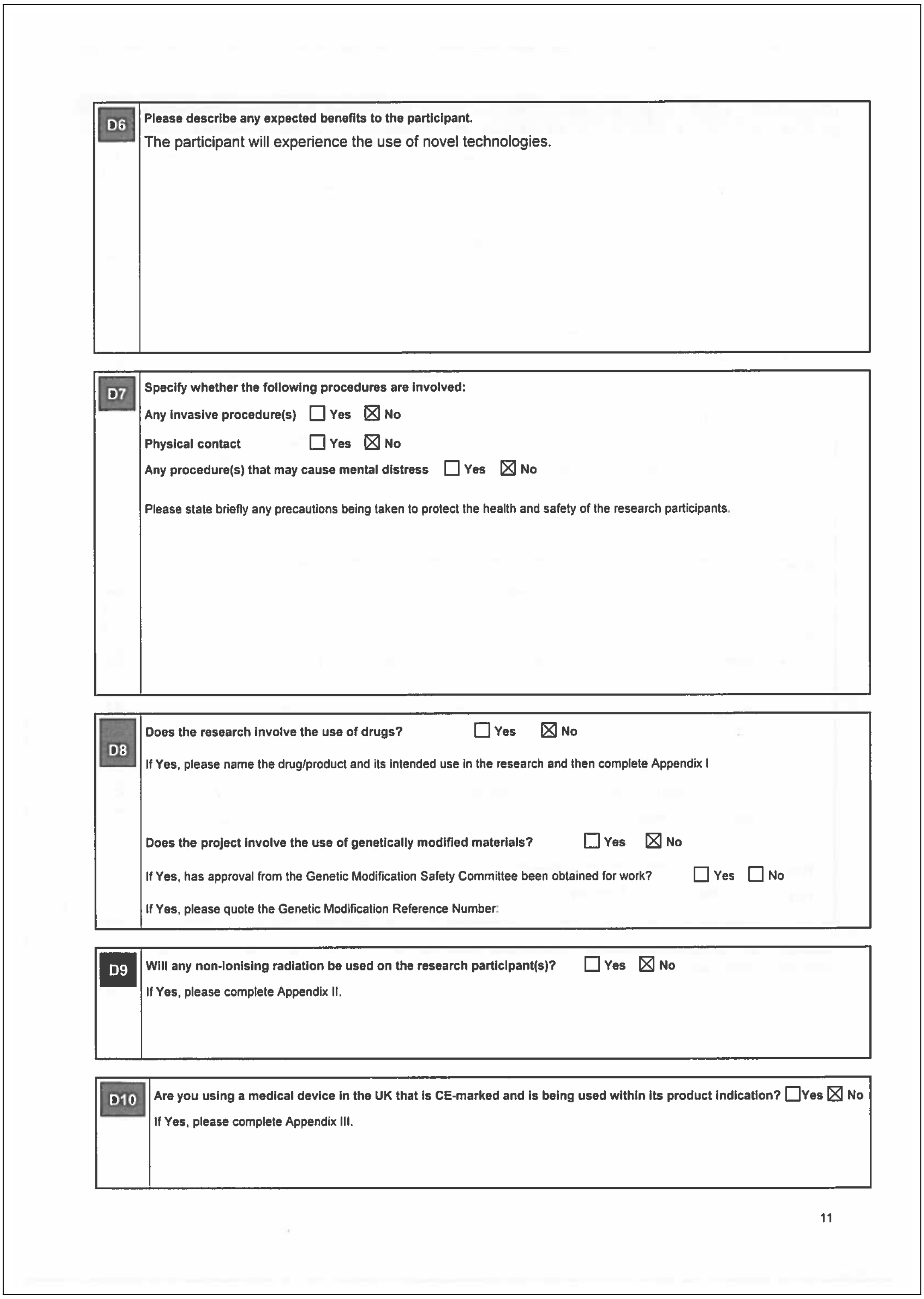
D4 **Will these participants participate in any activities that may be potentially stressful or harmful in connection with this research?** ☐ **Yes** ☒ **No**

If **Yes**, please describe the nature of the risk or stress and how you will minimise and monitor it.

D5 **Will group or individual interviews/questionnaires raise any topics or issues that might be sensitive, embarrassing or upsetting for participants?**

If **Yes**, please explain how you will deal with this.

No



D6 **Please describe any expected benefits to the participant.**

The participant will experience the use of novel technologies.

D7 **Specify whether the following procedures are involved:**

Any invasive procedure(s) ☐ **Yes** ☒ **No**

Physical contact ☐ **Yes** ☒ **No**

Any procedure(s) that may cause mental distress ☐ **Yes** ☒ **No**

Please state briefly any precautions being taken to protect the health and safety of the research participants.

D8 **Does the research involve the use of drugs?** ☐ **Yes** ☒ **No**

If Yes, please name the drug/product and its intended use in the research and then complete Appendix I

Does the project involve the use of genetically modified materials? ☐ **Yes** ☒ **No**

If Yes, has approval from the Genetic Modification Safety Committee been obtained for work? ☐ Yes ☐ No

If Yes, please quote the Genetic Modification Reference Number:

D9 **Will any non-ionising radiation be used on the research participant(s)?** ☐ **Yes** ☒ **No**

If Yes, please complete Appendix II.

D10 **Are you using a medical device in the UK that is CE-marked and is being used within its product indication?** ☐**Yes** ☒ **No**

If Yes, please complete Appendix III.

CHECKLIST

Please submit ether 12 copies (1 original + 11 double sided photocopies) of your completed application form for full committee review or 3 copies (1 original + 2 double sided copies) for chair's action, together with the appropriate supporting documentation from the list below to the UCL Research Ethics Committee Administrator. You should also submit your application form electronically to the Administrator at: ethics@ucl.ac.uk

Documents to be Attached to Application Form (if applicable)	**Ticked if attached**	**Tick if not relevant**
Section B: Details of the Project		
• Questionnaire(s) / Psychological Tests	☒	☐
• Relevant correspondence relating to involvement of collaborating department/s and agreed participation in the research.	☐	☒
Section C: Details of Participants		
• Parental/guardian consent form for research involving participants under 18	☐	☒
• Participant/s information sheet	☒	☐
• Participant/s consent form/s	☒	☐
• Advertisement	☒	☐
Section D: Details of Risks and Benefits to the Researcher and the Researched		
• Insurance registration form and related correspondence	☐	☒
Appendix I: Research Involving the Use of Drugs		
• Relevant correspondence relating to agreed arrangements for dispensing with the pharmacy	☐	☒
• Written confirmation from the manufacturer that the drug/substance has has been manufactured to GMP	☐	☒
• Proposed volunteer contract	☐	☒
• Full declaration of financial or direct interest	☐	☒
• Copies of certificates: CTA etc...	☐	☒
Appendix II: Use of Non-Ionising Radiation		
Appendix III: Use Medical Devices		

Please note that correspondence regarding the application will normally be sent to the Principal Researcher and copied to other named individuals.

Appendix E

Colophon

This document was set using LaTeXand BibTeXwith the UCL Thesis document class, composed with ShareLaTeX and the following tools:

Matlab. IBM SPSS. Autodesk. AutoCAD. Evernote.

Bibliography

[1] Olive Jean Dunn. Multiple Comparisons Using Rank Sums. *Technometrics*, 6(3):241–252, 1964.

[2] Patrick Carlson, Anicia Peters, Stephen B Gilbert, Judy M Vance, and Andy Luse. Virtual Training: Learning Transfer of Assembly Tasks. *IEEE Transactions on Visualization and Computer Graphics*, 21(6):770–782, Jan. 2015.

[3] Neil Burgess, Eleanor A Maguire, and John O'Keefe. The Human Hippocampus and Spatial and Episodic Memory. *Neuron*, 35(4):625–641, 2002.

[4] Michael J Tarr and William H Warren. Virtual Reality in Behavioral Neuroscience and Beyond. *nature neuroscience*, 5:1089–1092, 2002.

[5] Tom Hartley, Iris Trinkler, and Neil Burgess. Geometric Determinants of Human Spatial Memory. *Cognition*, 94(1):39–75, 2004.

[6] Evangelos Vassos, Esben Agerbo, Ole Mors, and Carsten Bøcker Pedersen. Urban-rural Differences in Incidence Rates of Psychiatric Disorders in Denmark. *The British Journal of Psychiatry*, 208(5):435–440, 2016.

[7] John P Eberhard. Applying Neuroscience to Architecture. *Neuron*, 62(6):753–756, 2009.

[8] Samantha L Allison, Anne M Fagan, John C Morris, and Denise Head. Spatial Navigation in Preclinical Alzheimers Disease. *Journal of Alzheimer's Disease*, 52(1):77–90, 2016.

[9] James P Bliss, Philip D Tidwell, and Michael A Guest. The Effectiveness of Virtual Reality for Administering Spatial Navigation Training to Firefighters. *Presence: Teleoperators and Virtual Environments*, 6(1):73–86, 1997.

[10] Mel Slater, Martin Usoh, and Anthony Steed. Taking Steps: the Influence of a Walking Technique on Presence in Virtual Reality. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 2(3):201–219, 1995.

[11] Mel Slater. A Note on Presence Terminology. *Presence connect*, 3(3):1–5, 2003.

[12] Mel Slater. 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, 2009.

[13] Eric D. Ragan. The Effects of Higher Levels of Immersion on Procedure Memorization Performance and Implications for Educational Virtual Environments, 2010.

[14] Doug A Bowman and Ryan P McMahan. Virtual Reality: How Much Immersion is Enough? *Computer*, 40(7), 2007.

[15] Jack M Loomis, James J Blascovich, and Andrew C Beall. Immersive Virtual Environment Technology as a Basic Research Tool in Psychology. *Behavior Research Methods, Instruments, & Computers*, 31(4):557–564, 1999.

[16] Eric D Ragan, Doug A Bowman, Regis Kopper, Cheryl Stinson, Siroberto Scerbo, and Ryan P McMahan. Effects of Field of View and Visual Complexity on Virtual Reality Training Effectiveness for a Visual Scanning Task. *Visualization and Computer Graphics, IEEE Transactions on*, 21(7):794–807, 2015.

[17] Martin Usoh, Kevin Arthur, Mary C. Whitton, Rui Bastos, Anthony Steed, Mel Slater, and Frederick P. Brooks, Jr. Walking, Walking-in-place, Flying, in Virtual Environments. In *Proceedings of the 26th Annual*

Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '99, pages 359–364, New York, NY, USA, 1999. ACM Press/Addison-Wesley Publishing Co.

[18] Ryan P McMahan, Doug A Bowman, David J Zielinski, and Rachael B Brady. Evaluating Display Fidelity and Interaction Fidelity in a Virtual Reality Game. *Visualization and Computer Graphics, IEEE Transactions on*, 18(4):626–633, 2012.

[19] Doug A. Bowman, Ajith Sowndararajan, Eric D. Ragan, and Regis Kopper. Higher Levels of Immersion Improve Procedure Memorization Performance. *Proceedings of the 15th Joint virtual reality Eurographics conference on Virtual Environments*, 2009.

[20] David Waller, Earl Hunt, and David Knapp. The Transfer of Spatial Knowledge in Virtual Environment Training. *Presence: Teleoperators and Virtual Environments*, 7(2):129–143, March 1998.

[21] Charles R Gallistel. *The Organization of Learning*. The MIT Press, 1990.

[22] Ken Cheng. A Purely Geometric Module in the Rat's Spatial Representation. *Cognition*, 23(2):149–178, 1986.

[23] Jonathan W Kelly, Timothy P McNamara, Bobby Bodenheimer, Thomas H Carr, and John J Rieser. Individual Differences in Using Geometric and Featural Cues to Maintain Spatial Orientation: Cue Quantity and Cue Ambiguity are More Important than Cue Type. *Psychonomic bulletin & review*, 16(1):176–181, 2009.

[24] Maria Murcia Lopez, Fiona Zisch, William Steptoe, Hugo Spiers, and Anthony Steed. Preliminary Findings on the Effect of Scale Transformations on Spatial Memory in Immersive and Desktop Virtual Environments. *Cognitive Processing*, 16:S75–S75, 2015.

[25] Victoria Interrante, Brian Ries, and Lee Anderson. Distance Perception in Immersive Virtual Environments, Revisited. In *Virtual Reality Conference, 2006*, pages 3–10. IEEE, 2006.

[26] Lane Phillips, Brian Ries, Victoria Interrante, Michael Kaeding, and Lee Anderson. Distance Perception in NPR Immersive Virtual Environments, Revisited. In *Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization*, pages 11–14. ACM, 2009.

[27] Masahiro Mori. The Uncanny Valley. *Energy*, 7(4):33–35, 1970.

[28] Mel Slater and Martin Usoh. Body Centred Interaction in Immersive Virtual Environments. *Artificial life and virtual reality*, 1:125–148, 1994.

[29] Erin A McManus, Bobby Bodenheimer, Stephan Streuber, Stephan De La Rosa, Heinrich H Bülthoff, and Betty J Mohler. The Influence of Avatar (Self and Character) Animations on Distance Estimation, Object Interaction and Locomotion in Immersive Virtual Environments. In *Proceedings of the ACM SIGGRAPH Symposium on applied perception in graphics and visualization*, pages 37–44. ACM, 2011.

[30] Anthony Steed, Ye Pan, Fiona Zisch, and William Steptoe. The Impact of a Self-avatar on Cognitive Load in Immersive Virtual Reality. In *2016 IEEE Virtual Reality Conference (Vr)*, volume 23, pages 67–76. IEEE, 2016.

[31] Mark H Draper, Maxwell J Wells, Valerie J Gawron, and Tom A Furness III. Exploring the Influence of a Virtual Body on Spatial Awareness. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 40, pages 1146–1150. SAGE Publications Sage CA: Los Angeles, CA, 1996.

[32] Lane Phillips, Brian Ries, Michael Kaeding, and Victoria Interrante. Avatar Self-Embodiment Enhances Distance Perception Accuracy in Non-Photorealistic Immersive Virtual Environments. In *Virtual Reality Conference (VR), 2010 IEEE*, pages 115–1148. IEEE, 2010.

[33] Eric D Ragan, Curtis Wilkes, Yong Cao, and Doug A Bowman. The Effects of Virtual Character Animation on Spatial Judgments. 2012.

[34] Brian Ries, Victoria Interrante, Michael Kaeding, and Lee Anderson. The Effect of Self-Embodiment on Distance Perception in Immersive Virtual Environments. In *Proceedings of the 2008 ACM symposium on Virtual reality software and technology*, pages 167–170. ACM, 2008.

[35] Elham Ebrahimi, Leah S Hartman, Andrew Robb, Christopher C Pagano, and Sabarish V Babu. Investigating the Effects of Anthropomorphic Fidelity of Self-Avatars on Near Field Depth Perception in Immersive Virtual Environments. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 1–8. IEEE, 2018.

[36] A Full-body [a.

[37] Brian Ries, Victoria Interrante, Michael Kaeding, and Lane Phillips. Analyzing the Effect of a Virtual Avatar's Geometric and Motion Fidelity on Ego-Centric Spatial Perception in Immersive Virtual Environments. In *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology*, pages 59–66. ACM, 2009.

[38] Betty J Mohler, Sarah H Creem-Regehr, William B Thompson, and Heinrich H Bulthoff. The Effect of Viewing a Self-Avatar on Distance Judgments in an HMD-based Virtual Environment. *Presence: Teleoperators and Virtual Environments*, 19(3):230–242, 2010.

[39] María Murcia-López and Anthony Steed. The Effect of Environmental Features, Self-Avatar, and Immersion on Object Location Memory in Virtual Environments. *Frontiers in ICT*, 3:24, 2016.

[40] Corey J Bohil, Bradly Alicea, and Frank A Biocca. Virtual Reality in Neuroscience Research and Therapy. *Nature reviews neuroscience*, 12(12):752–762, 2011.

[41] Francesco Vincelli. From Imagination to Virtual Reality: the Future of Clinical Psychology. *CyberPsychology and Behavior*, 2(3):241–248, 1999.

[42] Jim Blascovich, Jack Loomis, Andrew C Beall, Kimberly R Swinth, Crystal L Hoyt, and Jeremy N Bailenson. Immersive Virtual Environment Technology as a Methodological Tool for Social Psychology. *Psychological Inquiry*, 13(2):103–124, 2002.

[43] Giuseppe Riva. Virtual Reality: an Experiential Tool for Clinical Psychology. *British Journal of Guidance & Counselling*, 37(3):337–345, 2009.

[44] Kay M Stanney, Ronald R Mourant, and Robert S Kennedy. Human Factors Issues in Virtual Environments: A Review of The Literature. *Presence: Teleoperators and Virtual Environments*, 7(4):327–351, March 1998.

[45] Dylan Schmorrow, Joseph Cohn, and Denise Nicholson. *The PSI Handbook of Virtual Environments for Training and Education [Three Volumes]: Developments for the Military and Beyond.* Praeger Publishers, 2008.

[46] María Murcia-López and Anthony Steed. The Effect of Environmental Features, Self-avatar and Immersion on Object Location Memory in Virtual Environments. *Frontiers in ICT*, 3:24, Nov. 2016.

[47] Danae Stanton, Nigel Foreman, P Wilson, H Duffy, and Richard Parnell. Use of Virtual Environments to Acquire Spatial Understanding of Real-World Multi-Level Environments. *ICDVRAT 2002*, 2002.

[48] N Foreman and P Wilson. Prospects for VR Computing in Paediatric Medicine. *Ambulatory Child Health*, 1:176–182, 1995.

[49] Danaë Stanton, Paul Wilson, and Nigel Foreman. Using Virtual Reality Environments to Aid Spatial Awareness in Disabled Children. In *Proceedings of the 1st European conference on disability, virtual reality and associated technologies*, 1996.

[50] Paul N Wilson, Nigel Foreman, and Michael Tlauka. Transfer of Spatial Information from a Virtual to a Real Environment in Physically Disabled Children. *Disability and rehabilitation*, 18(12):633–637, 1996.

[51] Nigel Foreman, Paul Wilson, and Danae Stanton. VR and Spatial Awareness in Disabled Children. *Communications of the ACM*, 40(8):76–77, 1997.

[52] Danaë Stanton, Nigel Foreman, and Paul N Wilson. Uses of Virtual Reality in Clinical Training: Developing the Spatial Skills of Children with Mobility Impairments. *Studies in health technology and informatics*, 58:219–232, 1998.

[53] Craig R Hall and Carol D Horwitz. Virtual Reality for Training: Evaluating Retention of Procedural Knowledge. *International Journal of Virtual Reality*, 5(1):61–70, Nov. 2015.

[54] Nirit Gavish, Teresa Gutiérrez, Sabine Webel, Jorge Rodríguez, Matteo Peveri, Uli Bockholt, and Franco Tecchia. Evaluating Virtual Reality and Augmented Reality Training for Industrial Maintenance and Assembly Tasks. *Interactive Learning Environments*, 23(6):778–798, July 2015.

[55] Mar Gonzalez-Franco, Rodrigo Pizarro, Julio Cermeron, Katie Li, Jacob Thorn, Windo Hutabarat, Ashutosh Tiwari, and Pablo Bermell-Garcia. Immersive Mixed Reality for Manufacturing Training. *Frontiers in Robotics and AI*, 4(3):1, Feb. 2017.

[56] FD Rose, Elizabeth A Attree, BM Brooks, DM Parslow, and PR Penn. Training in Virtual Environments: Transfer to Real World Tasks and Equivalence to Real Task Training. *Ergonomics*, 43(4):494–511, Nov. 2010.

[57] Ajith Sowndararajan, Rongrong Wang, and Doug A Bowman. Quantifying the Benefits of Immersion for Procedural Training. In *Proceedings of the workshop on Immersive projection technologies/Emerging display technologiges*, page 2. ACM, Aug. 2008.

[58] Richard J Adams, Daniel Klowden, and Blake Hannaford. Virtual Training for a Manual Assembly Task. *Haptics-e, The Electronic Journal of Haptics Research*, 2(2), Oct. 2001.

[59] Mike Oren, Patrick Carlson, Stephen Gilbert, and Judy M Vance. Puzzle Assembly Training: Real World vs. Virtual Environment. In *Proceedings of the IEEE Virtual Reality Conference*, pages 27–30. IEEE, 2012.

[60] Dmitri Shuralyov and Wolfgang Stuerzlinger. A 3d Desktop Puzzle Assembly System. In *IEEE Symposium on 3DUI*, pages 139–140. IEEE, March 2011.

[61] Felix Ritter, Thomas Strothotte, Oliver Deussen, and Bernhard Preim. Virtual 3D Puzzles: A New Method for Exploring Geometric Models in VR. *IEEE Computer Graphics and Applications*, 21(5):11–13, 2001.

[62] Maria Murcia-Lopez and Anthony Steed. A Comparison of Virtual and Physical Training Transfer of Bimanual Assembly Tasks. *IEEE transactions on visualization and computer graphics*, 24(4):1574–1583, 2018.

[63] George M. Bodner and Roland B. Guay. The Purdue Visualization of Rotations Test. *The Chemical Educator*, 2(4):1–17, Oct. 1997.

[64] Albert A Rizzo and J Galen Buckwalter. Virtual Reality and Cognitive Assessment and Rehabilitation: The State of The Art. *Virtual Reality in Neuro-Psycho-Physiology: Cognitive, Clinical and Methodological Issues in Assessment and Rehabilitation*, 44:123, 1997.

[65] Steven G Vandenberg and Allan R Kuse. Mental Rotations, a Group Test of Three-Dimensional Spatial Visualization. *Perceptual and motor skills*, 47(2):599–604, Dec. 1978.

[66] Mores Prachyabrued and Christoph W Borst. Visual Feedback for Virtual Grasping. In *IEEE Symposium on 3DUI*, pages 19–26. IEEE, April 2014.

[67] Michael J Piller and Marc M Sebrechts. Spatial Learning in Transparent Virtual Environments. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 47(20):2133–2136, 2003.

[68] John O'Keefe and Jonathan Dostrovsky. The Hippocampus as a Spatial Map: Preliminary Evidence from Unit Activity in the Freely-Moving Rat. *Brain research*, 34(1):171–175, 1971.

[69] John O'keefe and Neil Burgess. Geometric Determinants of the Place Fields of Hippocampal Neurons. *Nature*, 381(6581):425, 1996.

[70] Nachum Ulanovsky and Cynthia F Moss. Hippocampal Cellular and Network Activity in Freely Moving Echolocating Bats. *Nature neuroscience*, 10(2):224, 2007.

[71] Arne D Ekstrom, Michael J Kahana, Jeremy B Caplan, Tony A Fields, Eve A Isham, Ehren L Newman, and Itzhak Fried. Cellular Networks Underlying Human Spatial Navigation. *Nature*, 425(6954):184, 2003.

[72] Edward C Tolman. Cognitive Maps in Rats and Men. *Psychological review*, 55(4):189, 1948.

[73] Stephen Kaplan. Cognitive Maps, Human Needs and the Designed Environment 5.4. *Environmental Design Research: Selected papers*, 1:275, 1973.

[74] Neil Burgess. Spatial Memory: How Egocentric and Allocentric Combine. *Trends in cognitive sciences*, 10(12):551–557, 2006.

[75] Beverly Roskos-Ewoldsen, Timothy P McNamara, Amy L Shelton, and Walter Carr. Mental Representations of Large and Small Spatial Layouts are Orientation Dependent. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(1):215, 1998.

[76] Melvyn A Goodale and A David Milner. Separate Visual Pathways for Perception and Action. *Trends in neurosciences*, 15(1):20–25, 1992.

[77] John O'Keefe and Lynn Nadel. *The Hippocampus as a Cognitive Map*. Oxford University Press, USA, 1978.

[78] Torkel Hafting, Marianne Fyhn, Sturla Molden, May-Britt Moser, and Edvard I Moser. Microstructure of a Spatial Map in the Entorhinal Cortex. *Nature*, 436(7052):801–806, 2005.

[79] Tom Hartley, Neil Burgess, C Lever, F Cacucci, and J O'Keefe. Modeling Place Fields in Terms of the Cortical Inputs to the Hippocampus. *Hippocampus*, 10(4):369–379, 2000.

[80] Colin Lever, Stephen Burton, Ali Jeewajee, John O'Keefe, and Neil Burgess. Boundary Vector Cells in the Subiculum of the Hippocampal Formation. *The journal of neuroscience*, 29(31):9771–9777, 2009.

[81] Trygve Solstad, Charlotte N Boccara, Emilio Kropff, May-Britt Moser, and Edvard I Moser. Representation of Geometric Borders in the Entorhinal Cortex. *Science*, 322(5909):1865–1868, 2008.

[82] Jeffrey S Taube, Robert U Muller, and James B Ranck. Head-Direction Cells Recorded from the Postsubiculum in Freely Moving Rats. I. Description and Quantitative Analysis. *The Journal of neuroscience*, 10(2):420–435, 1990.

[83] Kathryn J Jeffery. Self-Localization and the Entorhinal–Hippocampal System. *Current opinion in neurobiology*, 17(6):684–691, 2007.

[84] Eleanor A Maguire, Neil Burgess, James G Donnett, Richard SJ Frackowiak, Christopher D Frith, and John O'keefe. Knowing Where and Getting There: a Human Navigation Network. *Science*, 280(5365):921–924, 1998.

[85] Weimin Mou and Timothy P McNamara. Intrinsic Frames of Reference in Spatial Memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(1):162, 2002.

[86] Marcia L Spetch, Ken Cheng, and Suzanne E MacDonald. Learning the Configuration of a Landmark Array: I. Touch-Screen Studies with Pigeons and Humans. *Journal of Comparative Psychology*, 110(1):55, 1996.

[87] Marcia L Spetch, Ken Cheng, Suzanne E MacDonald, Brie A Linkenhoker, Debbie M Kelly, and Sharon R Doerkson. Use of Landmark Configuration in Pigeons and Humans: II. Generality Across Search Tasks. *Journal of Comparative Psychology*, 111(1):14, 1997.

[88] Ranxiao Frances Wang and Daniel J Simons. Active and Passive Scene Recognition Across Views. *Cognition*, 70(2):191–210, 1999.

[89] David Waller, Jack M Loomis, Reginald G Golledge, and Andrew C Beall. Place Learning in Humans: The Role of Distance and Direction Information. *Spatial Cognition and Computation*, 2(4):333–354, 2000.

[90] Ranxiao Frances Wang and Elizabeth S Spelke. Human Spatial Representation: Insights from Animals. *Trends in cognitive sciences*, 6(9):376–382, 2002.

[91] Weimin Mou, Timothy P McNamara, Christine M Valiquette, and Björn Rump. Allocentric and Egocentric Updating of Spatial Memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(1):142, 2004.

[92] Neil Burgess and John O'Keefe. Neuronal Computations Underlying the Firing of Place Cells and their Role in Navigation. *Hippocampus*, 6(6):749–762, 1996.

[93] M-L Mittelstaedt and H Mittelstaedt. Homing by Path Integration in a Mammal. *Naturwissenschaften*, 67(11):566–567, 1980.

[94] Jacob Greaves. Distorting the Structure of Space: Impacts on Spatial Cognition. Master's thesis, University College London, 2016. Unpublished.

[95] Anisa Motala. Distorting the Structure of Space: Impacts on Spatial Cognition. Master's thesis, University College London, 2016. Unpublished.

[96] Alexandra Kitson, Daniel Sproll, and Bernhard E Riecke. Influence of Ethnicity, Gender and Answering Mode on a Virtual Point-to-Origin Task. *Frontiers in behavioral neuroscience*, 10:22, 2016.

[97] Mary Hegarty, Anthony E Richardson, Daniel R Montello, Kristin Lovelace, and Ilavanil Subbiah. Development of a Self-Report Measure of Environmental Spatial Ability. *Intelligence*, 30(5):425–447, 2002.

[98] Isabel Briggs Myers, Mary H McCaulley, Naomi L Quenk, and Allen L Hammer. *MBTI Manual: A Guide to the Development and Use of the Myers-Briggs Type Indicator*, volume 3. Consulting Psychologists Press Palo Alto, CA, 1998.

[99] R.T. Hays and M.J. Singer. *Simulation Fidelity in Training System Design: Bridging the Gap Between Reality and Training*. Springer-Verlag, 1989.

[100] Havok. Rocketbox® libraries. `http://rocketbox.de/`, 2014.

[101] Stephen R Ellis and Brian M Menges. Localization of Virtual Objects in the Near Visual Field. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 40(3):415–431, 1998.

[102] Victoria Interrante, Brian Ries, Jason Lindquist, Michael Kaeding, and Lee Anderson. Elucidating Factors that can Facilitate Veridical Spatial Perception in Immersive Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 17(2):176–198, 2008.

[103] Qiufeng Lin, Xianshi Xie, Aysu Erdemir, Gayathri Narasimham, Timothy P. McNamara, John Rieser, and Bobby Bodenheimer. Egocentric Distance Perception in Real and HMD-based Virtual Environments: The Effect of Limited Scanning Method. In *Proceedings of the ACM SIGGRAPH*

Symposium on Applied Perception in Graphics and Visualization, APGV '11, pages 75–82, New York, NY, USA, 2011. ACM.

[104] Jack M Loomis and Joshua M Knapp. Visual Perception of Egocentric Distance in Real and Virtual Environments. *Virtual and adaptive environments*, 11:21–46, 2003.

[105] Wesley K Thompson, Peter Willemsen, Amy Gooch, Sarah H Creem-Regehr, Jack M Loomis, Andrew C Beall, et al. Does the Quality of the Computer Graphics Matter when Judging Distances in Visually Immersive Environments? *Presence*, 13(5):560–571, 2004.

[106] Roy A Ruddle, Stephen J Payne, and Dylan M Jones. Navigating Large-Scale Virtual Environments: What Differences Occur between Helmet-Mounted and Desk-Top Displays? *Presence: Teleoperators and Virtual Environments*, 8(2):157–168, 1999.

[107] Abhishek Seth, Judy M Vance, and James H Oliver. Combining Dynamic Modeling with Geometric Constraint Management to Support Low Clearance Virtual Manual Assembly. *Journal of Mechanical Design*, 132(8):081002, July 2010.

[108] Oculus VR, LLC. Oculus Avatar SDK 1.15.0. `https://developer.oculus.com/downloads/package/oculus-avatar-sdk/1.14.0/`, 2017.

[109] Andreas Roever. Burr Tools 0.6.3. `http://burrtools.sourceforge.net/`, 2015.