STARS

Electronic Theses and Dissertations, 2004-2019

2019

The Effects of Presence and Cognitive Load on Episodic Memory in Virtual Environments

Paul Barclay

University of Central Florida

University of
Central Florida

STARS
Showcase of Text, Archives, Research & Scholarship

THE EFFECTS OF PRESENCE AND COGNITIVE LOAD ON EPISODIC MEMORY IN
VIRTUAL ENVIRONMENTS

by

PAUL ANTHONY BARCLAY
B.S. University of Houston, 2009
M.A. University of Houston – Clear Lake, 2011

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Psychology
in the College of Science
at the University of Central Florida
Orlando, Florida

Summer Term
2019

Major Professor Valerie K. Sims

ABSTRACT

Episodic memory refers to an individual's memory for events that they have experienced in the past along with the associated contextual details. In order to more closely reflect the way that episodic memory functions in the real world, researchers and clinicians test episodic memory using virtual environments. However, these virtual environments introduce new interfaces and task demands that are not present in traditional methodologies. This dissertation investigates these environments through the lenses of Presence and Cognitive Load theories in order to unravel the ways that basic technological and task differences may affect memory performance. Participants completed a virtual task under High and Low Immersion conditions intended to manipulate Presence and Single-Task, Ecological Dual-Task and Non-Ecological Dual-Task conditions intended to manipulate cognitive load. Afterward they completed a battery of memory tasks assessing spatial, object, and feature binding aspects of episodic memory. Analysis through 2x3 ANOVA showed that performance for spatial memory is greatly improved by manipulation of Presence, where performance for object memory is improved by germane cognitive load. Exploratory analyses also revealed significant gender differences in spatial memory performance, indicating that improving Presence may offset the higher levels in male performance traditionally seen on spatial tasks. These results have practical implications for clinical memory assessment, as well as training paradigms and may serve to highlight the differences in the ways that memory is studied in the laboratory versus the way that it is employed in day-to-day life. Future studies based on this research should focus on linking these differences in memory performance to visuospatial and verbal strategies of memorization and determining whether the effects observed in this study replicate using other manipulations of presence and cognitive load.

ACKNOWLEDGEMENTS

I would like to thank my partner, Alice Watkins, for the support and assistance she has provided throughout my graduate career, up to and including this dissertation. It would be impossible to detail all the ways that she has helped make this document a reality, but if you're reading this just know that it wouldn't have existed without her. Spelling and grammatical errors are my own fault, though. She wanted me to make sure you know that.

I would also like to thank my advisor Clint Bowers and my dissertation chair Valerie Sims. Both have been great mentors (although, with opposite styles!) and both have provided me with a variety of opportunities, guidance, and mentorship throughout my time at UCF. My committee members, Florian Jentsch and Steve Fiore, also contributed a great deal to improving the quality and focus of this dissertation research.

I would also like to thank my peers (labmates, friends, and RAs), especially Brandon Matsumiya, Jason Parker, Johanna Hidalgo, Kierstin Rudl, Gaby Flores-Cruz, and Eleanor Didden. These individuals all contributed in their own ways, through feedback, building electronic measures, helping me find space and conduct data collection, and general emotional support. I hope you all understand how much I appreciate you.

Finally, I'd like to thank my parents, Dawn and Obie Fortenberry. With the exception of (maybe) Alice, these two have spent more time than anyone listening to me endlessly babble about VR, 3D printing, myoelectric sensors, video games, cognitive models, and a variety of other topics that I have been obsessed with over the years. Thank you for always supporting whatever weird pursuit I'm on about and thank you for never being shy about letting me know how proud of me you are.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER ONE:
INTRODUCTION

Episodic memory is the memory of events that one has experienced in the past, along with the contextual details associated with these events (Tulving, 1972). When studied in the laboratory, episodic memory is generally broken down into 3 components: object memory, source memory (sometimes referred to as spatial memory or spatial-location memory), and feature-binding. Object memory refers to the recollection of a specific item that one has encountered before, where source memory is the recollection of the details surrounding the encounter of that item, such as location (Tulving, 1983). Feature binding refers to the process through which object and source memories are associated with one another (Tulving, 1983).

The spatial component of episodic memory can further be broken down into two aspects, representing separate and distinct representations of spatial information (Golledge, Dougherty, & Bell,1995). Egocentric spatial representations (sometimes called route memory) refer to a first-person representation in memory where spatial information is encoded in relation to the observer. Allocentric spatial representations (sometimes called survey memory) refer instead to a complex system of object-to-object spatial relationships. This type of knowledge is often characterized as similar to an overhead map of the environment. Conversion from one of these spatial frames can be difficult and may result in poorer memory performance (Morganti, Stefanini, & Riva, 2013).

Clinical neuropsychologists have a special interest in episodic memory due to its association with several key markers of cognitive dysfunction. The specificity of episodic memories has been shown to be impaired in patients with depression (Lemogne et al, 2006) and difficulty with forming new episodic memories predicts the onset of dementia (Bäckman, Small, & Fratiglioni, 2001). Additionally, decline in spatial aspects of episodic memory without

similar decline in object-related aspects can be used to predict the development of Alzheimer's disease and differentiate Alzheimer's from other types of dementia (Pengas et al, 2010). Due to these and other clinical implications, high priority is placed on the development of more precise, more predictive methods of testing episodic memory.

The most common forms of episodic memory assessment present a series of words or other stimuli to a patient, followed by a recall or recognition memory task (cf. Parker et al, 1995). Other assessment techniques may incorporate pictures or video. However, as part of a recent push to improve the ecological validity of neuropsychological assessment, many approaches have arisen that leverage virtual reality to assess the formation of episodic memories in more complex environments (Parsons, 2016). These tasks immerse the participant into a fully-rendered virtual environment (VE) and expose them to in-environment objects or events (Burgess, Maguire, Spiers, & O'Keefe, 2002; Plancher et al, 2012). After participants explore the virtual environment, they complete a delayed recall task for those items and their context. These methods yield comparable findings to traditional memory testing while maintaining a higher level of ecological validity (Parsons & Rizzo, 2008; Arvind-Pala et al., 2014).

Recent technological and industrial advancements have helped drive the adoption of virtual assessment techniques. While in previous decades the cost to build and maintain a virtual reality system in a hospital would have been prohibitively expensive, the rise of virtual reality to the forefront of entertainment computing has driven down the cost for high-end display technologies, such as graphics cards and head-mounted displays. At the same time, graphical and computing power have advanced to a point where these virtual environments can be both highly controlled and highly realistic.

The same advances in technology that have enabled this move have also enabled episodic memory assessment to be presented in an interactive format, allowing for active

exploration of the virtual environment rather than passive viewing. The move from passive to active interactive styles has led to a literature that investigates the differences between the two modalities (Parsons, 2016). The general pattern of findings shows that active navigation enhances episodic memory (Plancher, 2017) with a few exceptions (Chrastil & Warren, 2012). Overall, these results have been taken to indicate that active navigation improves object memory, source memory, and feature binding so long as the mode of navigation does not make the task too demanding (Parsons, 2016).

However, the exact cognitive mechanisms that lead to this enhancement are not fully understood, raising questions about the predictive validity of this type of assessment. The variety of interactive methods, presentation methods, and task approaches used may introduce task differences that muddy the population norms necessary for accurate diagnosis. To disentangle this, it is useful to take a step away from purely clinical concerns and approach these assessments as virtual tasks in which the goal of the task is to learn and remember environmental details. When the evidence is examined from this perspective, two primary theories can be proposed explaining the differences seen between active and passive task approaches. The Presence theory argues that the interface mechanisms required for active navigation contribute to the experience of Presence, or the subjective feeling of being physically present in the virtual environment. These feelings of Presence result in improved memory for the virtual environment. On the other hand, Cognitive load theory argues that active navigation increases the cognitive load of the virtual task, thereby enhancing performance by eliminating cognitive underload. The following sections examine each theory in detail and the evidence supporting these claims.

Presence

Presence is defined as a user's feelings of being physically situated in a virtual environment. In the body of scientific literature, Presence is often conflated with Immersion-- a similar and related construct. Because of this tendency, it is necessary to clearly define and separate these constructs. For the purposes of this study, Presence refers to a user's subjective response to a virtual environment and has traditionally been measured using self-report questionnaires (Slater, 2003). Immersion, then, refers to the environmental factors of the VE that envelope the user's senses and promote a feeling of Presence. These factors include Breadth (the number of sensory modalities), Depth and Resolution (the intensity of presentation to a specific sensory modality), Motion (the ability of the environment to evoke natural motion), and Consistency (the ability of the user to predict outcomes) (Nash, Thompson, & Barfield, 2000).

Despite this conceptual separation, Immersion and Presence are demonstrably related to one another in that more immersive environments tend to evoke greater feelings of presence (Nash et al, 2000). For example, an environment presented to the user on a head-mounted display (HMD) would envelop a greater portion of the user's vision than an environment presented on a desktop monitor, increasing its Immersion. Users of this environment will generally experience higher levels of Presence, but users who are unaccustomed to a HMD may be much more sensitive to this change resulting in a high magnitude of Presence experience.

There are two essential elements to the theory that the differences in active and passive assessment of episodic memory may be explained by the Presence construct. First, active interaction with a virtual environment has repeatedly been shown to increase a user's sense of Presence (c.f. Welch et al, 1996). Second, higher levels of Presence are correlated with

improved task performance across a wide range of tasks performed in virtual environments (Stanney et al, 1998; Nash et al, 2000). Several recent studies have argued that this observed correlation may be specifically related to episodic memory, in that feelings of Presence may encourage users to encode the environment in episodic, rather than semantic, memory (Cheng, She, & Annetta, 2015).

Several studies have yielded results that support this theory. For example, Bakdash, Augustyn, and Proffitt (2006) asked participants to freely explore a virtual environment presented either on a small or large screen (20" x 15" and 58" x 43", respectively). After the exploration phase, participants were asked to indicate the location of several environmental landmarks in the VE using a tracked wand. Participants who had the environment presented on the larger screen showed significantly reduced errors in the pointing task, indicating improved spatial memory for the virtual environment.

Wallet et al (2011) performed a similar experiment, this time in tandem with an active/passive manipulation. In this experiment, the researchers looked at the interaction between visual fidelity (detailed or undetailed) and navigation mode (active or passive). In the detailed condition, participants explored a virtual city with detailed colors and textures. In the undetailed version participants explored a virtual city comprised entirely of greyscale boxes that matched the city buildings in size and shape. Participants were asked to perform wayfinding and sketch mapping tasks assessing their spatial memory of the environment. The researchers found a main effect for both manipulations and an interaction indicating an additive effect in which participants who actively navigated the detailed environment performed best and participants who passively navigated an undetailed environment performed worst. Both the Bakdash et al (2006) and Wallet et al (2011) experiments demonstrate that manipulating the Depth and Resolution factors of Immersion improve spatial memory.

Another experiment conducted by Gaunet, Vidal, Kemeny, and Berthoz (2001) explored the differences in passive, snapshot, and active exploration on spatial memory for a virtual city environment. The active and passive conditions in this study were similar to those described in other experiments (self-guided exploration with a joystick vs pre-recorded video). However, in the snapshot condition participants were guided through the environment in a series of still images. Participants were then asked to perform a sketch mapping task. Remarkably, participants in the snapshot condition showed significantly impaired performance even when compared to those in the passive condition. This experiment demonstrates that manipulation of the Motion factor of Immersion impacts spatial memory, as the snapshot condition represents a less natural form of motion than passive or active conditions.

Finally, in an experiment conducted by Carassa, Geminiani, Morganti, and Varotto (2002), participants were asked to explore a VE under self-governed and avatar-guided conditions. In both conditions, participants used a modified joystick to explore the virtual environment. However, in the avatar-guided condition, participants were required to follow a virtual guide, removing the decision-making component from navigation. Participants then completed a wayfinding task, a pointing task, and a sketch-mapping task. Participants in the self-governed group showed higher performance across memory measures when compared to the avatar guided group. Similar results establishing the importance of decision-making control have also been demonstrated by subsequent researchers (Bakdash, Linkenauger, & Proffitt, 2008; Plancher, Barra, Orriols & Piolino, 2013; Jebara et al, 2014). These experiments demonstrate that changes in the Consistency factor of Immersion affect spatial memory, as both active control and decision-making control allow the user to make better predictions about the environment.

In all the experiments discussed here, manipulating a factor of Immersion resulted in improvements in the level to which users were able to recall details of the virtual environment. These relationships suggest a Presence theory of episodic memory for virtual environments. In this theory, it is purported that the addition of active engagement with the VE is a technological means of increasing the Immersion of the VE, thereby increasing the user's subjective experience of Presence. When users feel present and emotionally engaged with the environment, they are more likely to encode the environment as a self-relevant experience in episodic memory. It is notable, however, that evidence supporting this theory is drawn almost entirely from a literature concerned with spatial memory rather than episodic memory in general. Additionally, as none of the supporting studies used a direct measure of Presence, the effect of these various manipulations is assumed rather than demonstrated.

Cognitive Load

Cognitive Load theory is used to describe the total amount of working memory resources that have been devoted to a particular task under a variety of task conditions (Paas, Renkl, & Sweller, 2004). As the mental demands of a task increase, the individual performing that task must devote a greater level of working memory to that task, thus increasing the task's cognitive load. Generally, the relationship of performance to cognitive load follows a U-shaped distribution in which task performance will be degraded at excessively low and excessively high levels of cognitive load (Teigen, 1994).

However, it is not simply the case that either the introduction of additional cognitive load to a low-load task nor that the reduction of cognitive load from a high-load task will improve task performance. This is due to differences not only in raw level of cognitive load, but also the type of cognitive load being added or removed. Sweller (2010) proposes that there are three types of cognitive load: intrinsic, extraneous, and germane cognitive load. Intrinsic cognitive load is

created by the inherent complexity of the task that is being conducted. It represents a minimum value below which cognitive load of a given task cannot be reduced. Extraneous cognitive load is defined as additional cognitive load that is introduced through the ways that individuals interface with the task, such as instructional methods and technological abstraction. Traditionally, this type of cognitive load negatively impacts task performance and can be reduced by optimization of instruction or task interfaces. In contrast, germane cognitive load represents the total amount of cognitive load that is dedicated to essential learning materials. In this conception, the addition of cognitive load may increase task performance if, as a result, the individual performing the task devotes a greater proportion of mental resources toward the essential task rather than extraneous elements.

This differentiation serves to better explain why higher levels of cognitive load are sometimes associated with performance improvements and other times associated with performance decrements (Sweller, 2010). However, it introduces a secondary issue of how to determine whether additional cognitive load is germane or extraneous without relying on a circular argument (i.e. that germane cognitive load increases performance, so if performance is increased then germane cognitive load must have been added). One potential metric for this question involves the computation of Instructional Efficiency (IE; Paas & Van Merriënboer, 1993). The calculation of IE involves standardizing cognitive load ratings and performance metrics within a given experiment and taking a ratio of performance to cognitive load. This metric provides a relative measure of performance output per unit of cognitive load and enables researchers to determine whether improved performance comes at a cost of disproportionate increases in cognitive load.

The application of Cognitive Load theory to episodic memory assessment was originally proposed by Plancher and Piolino (2017) to explain a variety of results found across studies that

used near-identical virtual environments with different populations of participants. In 2010, Plancher, Gyeslinck, Nicolas, and Piolino conducted an experiment using a virtual city environment in order to compare their virtual task with classic neuropsychological test. The study used both young adult and older adult samples and manipulated instructions for encoding style. Participants in both groups were assigned to either encode the environment intentionally (received instructions to remember the environment) or incidentally (did not receive instructions to remember the environment). Afterward, participants were assessed on their memory of objects and events within the virtual environment using Free Recall. Overall, participants in the intentional encoding group remembered more environmental details. However, there was an unexpected interaction effect in which older adults assigned to the incidental group outperformed younger adults.

In a series of follow-ups (Plancher et al, 2012; Plancher et al, 2013), the researchers looked at passive, decision-control, and full motor-control in young adults, healthy aging, and patients with mild cognitive impairment. Their results showed that both decision-control and full motor control resulted in improved recall of environmental details in young adults. In older adults, however, decision-control and passive conditions were equivalent and a decline in recall was observed in the full motor-control condition. In patients with mild cognitive impairment, performance was best in the passive condition, with both decision-control and motor-control conditions resulting in diminished recall performance.

A similar set of experiments demonstrated this same pattern of results in a different VE using a recognition memory task. Sauzeon et al (2011) immersed a sample of young adults in a virtual apartment environment that held a controlled number of modeled objects. These participants were exposed in either passive or active conditions and then completed a recognition memory task that contained an equal number of environmental objects and foil

objects. The active group recognized significantly more objects and reported significantly fewer

false recognitions of foil objects. Subsequent follow-ups by Arvind-Pala et al (2014), Sauzeon et

al (2014), and Sauzeon et al (2016) repeated these procedures with clinical populations,

seeking to find associations between episodic memory performance and working memory. Their

results showed similar patterns, with active conditions improving performance for young adults

but impairing performance for older adults. However, performance for both groups had a strong

correlation with measures of working memory in which participants with greater working memory

capacity and higher levels of executive control were able to recognize more environmental

objects. These studies provide evidence that the contradictory results of previous experiments

are associated with individual differences in working memory.

These explanations are supported by the results of Jebara et al (2014). In this

experiment, both older and younger participants were divided across 4 groups, Passive,

Itinerary Control, Low Control, and High Control. In the Passive condition, participants were

passengers in a virtual car. In the Itinerary Control group, participants were also passengers but

chose the directions and route. In the Low Control group, participants drove the vehicle but did

not choose the route and in the High Control condition participants drove the vehicle and chose

the route. The results showed that memory performance, as measured by a recall task, was

universally improved in the Itinerary Control and Low Control Groups. However, in the High

Control group, young adults outperformed older adults and memory performance had a

moderate correlation with measures of working memory capacity and executive function.

In a later book chapter written by Plancher and Piolino (2017), the authors summarize

these results. In younger adults, being passively exposed to the virtual environment does not

add enough workload to overcome cognitive underload. Therefore, adding an active element

improves task performance by increasing cognitive load. However, older adults tend to

experience reductions in capacity as they age. As a result, interactive elements can push cognitive load past capacity limits and degrade task performance. It is important to note that this explanation synthesizes across a diverse range of potential sources of task demand, resulting in a unitary conception of cognitive load. Across their results, the authors assume that workload is added by 1) directing attention (intentional vs. incidental encoding) and 2) addition of a manual task (manipulation of the control device), and 3) the information processing required to make decisions. Additionally, they assert that working memory capacity, and therefore workload capacity, is reduced by age-related changes to the brain.

Based on these results, a Cognitive Load explanation of the active/passive enhancement effect would provide a parsimonious explanation covering the performance of both younger and older adults. This theory would suggest that there is nothing specific to actively controlling one's movement through the virtual environment that improves performance on a virtual episodic memory task and that similar results could be found through other means of raising the user's Cognitive load. It is important to note, however, that all of the experiments used to support this theory focus on object memory and feature binding. This raises some question concerning the applicability of this theory to episodic memory as a whole. Additionally, while this theory infers that the association between working memory and episodic memory outcomes come from differences in Cognitive load between the two conditions, none of these studies provide measures of workload to confirm this.

Most importantly, however, this explanation does not strictly account for levels of extraneous vs germane cognitive load. With the supporting data provided, it is equally possible to argue that younger adults, having more experience with virtual technologies, are better able to convert the additional cognitive load of active navigation into germane cognitive load. Older adults, on the other hand, may be spending more of the additional cognitive load grappling with

the interface itself and bridging the level of abstraction that it creates. It is therefore clear that any investigation of the effects of cognitive load on episodic memory performance must try to differentiate between germane and extraneous categories of cognitive load.

Interaction

The Presence and Cognitive load theories presented here have been drawn from experiments with different goals that study episodic memory using different memory components. For this reason, in the episodic memory literature there is a strong dissociation between results that are attributed to Presence and results that are attributed to Cognitive load. However, in the general body of VE research several experiments have been conducted that investigate these two constructs together. These experiments leave some question concerning the level to which these two constructs are separable and demonstrate the possibility that there may be some level of interaction between the two.

Ma and Kaber (2006), for example, conducted an experiment in which participants played a virtual game of basketball. In this experiment, Immersion was manipulated using environmental factors (perspective and depth of visual and auditory presentation) and task difficulty was manipulated with a secondary visual monitoring task. They then measured subjective levels of presence and workload and found a strong positive correlation between the two constructs. However, follow-up analysis showed no significant effect of the Immersion manipulation on workload and no significant effect of task difficulty on Presence. This means that even when presence and workload are manipulated separately, there remains a relationship between the two. Therefore, there is a possibility that this relationship may express itself as a mediating or moderating relationship with task performance. For example, it is possible that experiencing Presence is an active process that requires working memory, which

could mean that other sources of Cognitive load may diminish the user's capacity to experience presence.

Gender Differences

Previous research (Herlitz & Rehnman, 2008) using traditional, pen-and-paper measures of episodic memory has also suggested potential gender differences in episodic memory depending on the type of outcome measure that is used. Generally, women tend to perform better on measures of verbal and object memory while men tend to perform better on measures of spatial memory (relative size and route). These differences are demonstrably associated with similar gender differences in spatial vs verbal measures of working memory span (Lewin, Wolgers, & Herlitz, 2001).

However, even these general findings come with some caveats. For example, while male participants are shown to demonstrate better performance on route memory tasks in tasks that present few landmarks (Astir, Ortiz, & Sutherland, 1998), these differences tend to disappear when testing is conducted in environments where landmarks, objects, or other items are present that can support navigation strategies that rely on verbal rather than visuospatial skills (Crook, Youngjohn, & Larrabee, 1993).

It is therefore worthwhile to consider whether the ecologically-driven design of most contemporary VE-based episodic memory testing may advantage one gender over another. While previous evidence suggests that there should be a significant difference in performance based on gender, it is also possible that environments based on the real world may not adequately prevent female users from leveraging verbal strategies and/or male users from leveraging visuospatial strategies. For example, the simple existence of doors, windows, and furniture in an ecologically valid environment may allow female participants to make route judgments based on environmental objects rather than visuospatial skills.

Present Study

The present study aims to examine the applicability of Presence theory and Cognitive Load theory to virtual assessment of episodic memory. Presence theory suggests that the differences in performance can be explained by the level to which the user is made to feel physically situated in the virtual environment. Cognitive Load theory, on the other hand, would suggest that (in younger adults) an increase in cognitive load, particularly if that load is germane, will improve memory performance over lower-load conditions. The conflict and ambiguity between these two theories provides an opportunity to address several gaps in the evidence supporting each theory.

Foremost, while both theories offer an explanation for the existing literature, no studies have been conducted to specifically test either theory and no studies have sought to measure either Presence or Cognitive Load directly. This means any claim that a specific condition or manipulation supports a theory can be called into question because there is no basis for knowing which manipulations produced a measurable effect on a given construct. Additionally, while each of the two theories makes a claim regarding episodic memory, each theory is largely supported by evidence about only a specific component of episodic memory. Presence theory is supported almost entirely by evidence of its effect on the spatial component of episodic memory, while Cognitive Load theory is supported entirely by research looking at object memory and feature binding. It is therefore possible that both theories are correct, but only for a specific memory component, rather than episodic memory as a whole.

Additionally, these constructs have not been investigated together in an episodic memory paradigm. This means that there is a possibility of interaction between the two. Previous studies in virtual environments have shown even when Presence and Cognitive load are manipulated separately, the two constructs maintain a moderately powerful association (Ma

& Kaber, 2006). This suggests the possibility that both Presence and Cognitive load contribute independent improvements to task performance in episodic memory tasks that are difficult to dissociate from one another or that one construct's effect is mediated by the other.

Finally, while a fair amount of research using traditional measures of episodic memory have shown that gender differences may exist for certain subtypes of episodic memory, it remains to be seen if these gender differences persist when assessed using ecologically-focused virtual environments. This study will aim to assess these differences to determine whether the reduction in rigid control of strategy diminishes the observable effects of gender and their related verbal vs visuospatial strategies.

The present study aims to address these gaps to better understand the strengths and weaknesses of the two existing theories. First, this study was structured to use canonical manipulations of Presence and Cognitive Load to compare direct manipulation of the construct in question with memory performance. Second, following the virtual task participants completed traditional self-report measures of Presence and Cognitive load intended to corroborate the assumed effects of these manipulations. Third, each component of episodic memory was assessed separately to determine the level to which each theory generalizes to episodic memory, and the level to which each theory makes specific component-level predictions. Finally, this study used a factorial design that enabled analysis of both theoretical main effects and interactions between the Presence and Cognitive Load constructs.

Hypotheses

While episodic memory itself can be divided almost endlessly into various components and subcomponents, this study was restricted specifically to the components of episodic memory upon which the Presence and Cognitive Load theories are based. These components are spatial memory, object memory, and feature binding. As each theory is supported by a

distinct literature that is concerned with a specific component memory type, it is worthwhile to break down this study's hypotheses as they relate to these components.

The clear majority of studies which focused on measuring Spatial Memory found improved memory as a result of manipulations that were likely to increase users' experience of Presence. This pattern of results spans measures that assessed spatial layout and relative size, both of which can be accurately assessed by applying different types of analysis to the results of a sketch-mapping task. This study manipulated Immersion at two levels in order to create conditions that will increase the experience of presence for the High Immersion group. Therefore, Hypothesis 1 can be broken down into the two following components:

H1a: Participants in High Immersion conditions will show better memory for spatial layout.

H1b: Participants in High Immersion conditions will show better route learning for the path taken in the virtual environment.

Object memory, on the other hand, has typically shown mixed results regarding manipulation of conditions. However, the Cognitive Load theory proposed by Plancher and Piolino (2017) was derived from a series of experiments that used recall and recognition memory tasks. These tasks primarily asked participants to recall objects from the virtual environment, such as grocery store names or types of fruit. As their primary outcome measure, the researchers used a raw score for number of correctly recalled or recognized environmental objects. The relationship of cognitive load to performance is typically thought to follow an inverted U-shaped curve, with the highest levels of performance occurring in conditions with

moderate levels of cognitive load. Based on this, the present study aims to manipulate Cognitive load using single-task and dual-task conditions. However, in order to account for differences in germane and extraneous cognitive load, this study will use both ecological and non-ecological dual-task conditions that are designed to increase cognitive load but also manipulate the allocation of cognitive load between germane and extraneous conditions.

As this study was conducted using a sample of young adults, it was not expected that the dual-task condition would be sufficiently difficult to increase cognitive load to unmanageable levels. Therefore, Hypothesis 2, which predicts that object memory is governed by cognitive load, can be broken down into the two following components:

H2a: Participants in Ecological Dual-Task conditions will be able to name more environmental objects during the Free Recall task, while participants in Non-Ecological Dual-Task conditions will name fewer.

H2b: Participants in Ecological Dual-Task conditions will successfully identify more environmental objects during the Recognition Task, while participants in Non-Ecological Dual-Task conditions will name fewer.

The Feature Binding component of episodic memory presents a unique challenge because it involves elements of both object memory and spatial memory. As a result, the evidence supporting each theory is similarly mixed. Plancher and Piolino (2017) report similar effects of Cognitive load on their measures of Feature Binding, that involved a subsequent Free Recall task in which participants listed the associated details of each object they had correctly recalled. This task captures a wide range of contextual information, such as size, color, and

location of objects. However, studies that support the Presence theory have shown presence-related improvements on measures that only require the binding of object and spatial memory, such as Landmark, Pointing, and Object Location tasks. Therefore, it is reasonable to suggest that Feature Binding measures that focus on object-spatial binding would behave similarly to measures of pure spatial memory, and that measures that capture object-context binding broadly will behave similarly to measures of pure object memory. Therefore, Hypothesis 3 can be broken down into the two following components:

H3a: Participants in Ecological Dual-Task conditions will be able to name more environmental details during the Free Recall task, while participants in Non-Ecological Dual-Task conditions will name fewer.

H3b: Participants in High Immersion conditions and Ecological Dual-Task conditions will show greater accuracy in the Object Location portion of the Recognition and Object Location task. There will also be an interaction effect of Immersion and Task groups, in which participants will perform best in the High Immersion, Ecological Dual-Task and perform worst in Low Immersion, Non-Ecological Dual-Task conditions.

Exploratory Analyses

Based on previous research conducted using traditional measures of episodic memory, such as pen-and-paper tasks, there is a consistent effect of gender. For measures that emphasize visuospatial skills, male participants tend to outperform female participants and for measures that emphasize verbal skills, female participants tend to outperform male participants.

This gender difference is generally considered to result from differences in visuospatial and verbal working memory (Lewin, Wolgers, & Herlitz, 2001). Based on this, Hypothesis 4 predicts:

H4a: Male participants will perform better on measures of spatial memory (both layout and route).

H4b: Female participants will perform better on measures of object memory (both recall and recognition).

H4c: Male and Female participants will show similar levels of performance on measures of feature binding (details per object and object location).

In order to determine the general level of germane vs extraneous workload represented by each condition, it will also be necessary to consider the level of instructional efficiency afforded by each condition for each performance metric. Based on our performance hypotheses, Hypothesis 5 predicts that IE will follow a similar pattern of outcomes as performance. By this measure:

H5a: Participants in the High Immersion condition will show higher IE for spatial layout and route learning.

H5b: Participants in the Ecological Dual-Task condition will show higher IE for object recall and object recognition than participants in either the non-ecological dual-task or the single-task conditions.

H5c: There will be an Immersion x Task interaction, through which IE for feature binding measures will be highest when High Immersion is combined with an ecological dual-task and lowest when low immersion is combined with a non-ecological dual task.

CHAPTER TWO:
METHOD

Sample

An a priori power analysis was conducted using G*Power 3.1.9.2 to ensure sufficient

power to detect an effect in this experimental procedure. This power analysis used an alpha

level of .05 and an effect size of η_p^2 = .06. This effect size was determined by examining the

effect size for the 10 previous studies that made direct comparisons of active vs passive

navigation and taking the lowest observed (η_p^2 ranged from 0.06 to 0.15). In order to obtain a

power level of 1 - β = 0.80, it was determined that we would require a total sample size of at

least 162.

183 participants were recruited through the University of Central Florida (UCF) SONA

system, an online utility that allows undergraduate students enrolled in psychology courses at

UCF to volunteer for research experience for course credit. Of these participants, 3 stopped the

virtual task early due to simulation sickness and 12 did not meet adequate performance on the

secondary task. These 15 participants were excluded from the analysis. The final sample used

in analysis included a total of 168 participants, randomly assigned to 6 experimental groups.

This left a total of 28 participants per group. Male and female participants were assigned to

each group in equal numbers in order to maintain group homogeneity of spatial and verbal

ability and evenly distribute potential gender-based differences in strategy. Of the 168 cases

analyzed, participants ranged in age from 18 to 29 (M=19.32, SD=1.90).

Design

This study used a between-subjects, 2x3 factorial design to compare the contributions of

virtual presence and workload to the enhancement of task performance on a virtual episodic

memory task. Independent variables (IV) were Immersion (High vs Low) and Cognitive Load

(Single Task vs Ecological Dual-Task vs Non-Ecological Dual-Task). The effects and interactions of these IVs were analyzed across 6 dependent variables (DV): two measures of Spatial Memory (survey and route), two measures of Object Memory (Recall and Recognition), and two measures of Feature Binding (Recall for details and Object Location). IV conditions are detailed in the Conditions section and DV measures are detailed in the Measures section.

Virtual Environment

The experimental VE was a virtual hospital environment. This environment was modeled from still photos and video of a real-world hospital in Orlando, Florida with alterations made to fit with the outcome measures used in this study. The experimental VE modeled the ground floor of this hospital, which is around 18,600 sq. feet and comprised of 8 distinct areas: a reception area, waiting room, triage area, ward, playroom, cafeteria, pharmacy and atrium. Each room contained a controlled number of modeled objects (10) for participants to recall. Additionally, participants were given a heads-up display of a timer indicating how much time they have left in the virtual task. The VE was built in Unreal Engine and was rendered to run on a high-end gaming PC built for virtual reality. It was designed to be presented on both a flat screen monitor and on the Oculus Rift head-mounted display. Audio was presented through Microsoft Life noise-reduction headsets.

Two versions (A and B) of the virtual environment were built in order to ensure that target objects and object foils presented in the Recognition and Object Location task (see Measures) were sufficiently counterbalanced. Both versions maintained the same physical layout with regard to rooms, room size, and object location. However, in Version B half of the target objects were replaced with alternate objects of similar size and shape and placed in identical locations (for example, a coffee mug might be replaced with a soda can, but they would

both be centered on the same location within the VE). This meant that Versions A and B

contained 40 common objects that were present in both version and 40 version-specific objects.

Conditions

Immersion (High vs Low). Presence was manipulated using immersive technologies

that have been shown to increase presence. In the Low Immersion condition, participants had

the experimental VE presented to them on a 23-inch widescreen LCD monitor and normal

computer headphones. In the High Immersion condition, participants had the experimental VE

presented to them through the Oculus Rift head-mounted display.

Figure 1: Screenshot from the Experimental VE

Cognitive Load (Single vs Ecological Dual-Task vs Non-Ecological Dual-Task).

This comparison is designed to manipulate the effects of cognitive load on task performance. In

the Single Task condition participants explored the environment aiming to learn the spatial

23

layout of the VE and the objects. Participants were instructed that they were a doctor tasked

with examining the way the hospital is currently using space and that their job was to remember

as much as they could, including rooms, layout, and objects in the environment. They were

given full control over their movement through and investigation of the VE for the duration of the

task.

In the Dual Task conditions, participants were asked to simultaneously complete both

the virtual episodic memory task and a secondary task that was designed to increase cognitive

load. For this task, participants heard a series of number pairs (e.g. 140 over 80) presented

through audio headphones. Participants were asked to monitor these number pairs and respond

by pressing the space bar when either the first number exceeded 160 or the second number

exceeded 120. For example, if the participant heard "150 over 130" or "170 over 100" they

would respond, where if they heard "120 over 80" they would not. The stimuli heard by both

dual-task groups was identical, as were the computational requirements. Based on this, it was

expected that increases in cognitive load should be similar in magnitude. Over the course of the

10-minute task, participants heard a total of 20 number pairs (one every 30 seconds), 10 of

which required response and 10 of which did not. To be included in the final sample,

participants needed to score at least 75% accuracy for the 20 trials (15 correct).

For the Ecological Dual-Task condition, participants were instructed that the number

pairs were blood pressure readings that they were meant to monitor while investigating the

hospital. This manipulation was intended to allow participants to integrate the additional task

load with the environmental setting and primary task. Based on the cognitive load explanations

of episodic memory performance, this should have increased levels of germane cognitive load

rather than extraneous cognitive load. For the Non-Ecological Dual-Task condition, participants

were instead instructed that the secondary task was a math task that was not connected to the

primary task. This should have led to similar levels of increased cognitive load due to the

additional processing requirements but should not integrate the secondary task with the primary

task. According to the cognitive load explanation of episodic memory performance, the

additional cognitive load would be extraneous rather than germane.

These conditions resulted in a total of six potential experimental groups. These groups

are shown below in Figure 2.

	Single Task	Ecological Dual Task	Non-Ecological Dual-Task
High Immersion	HI-ST	HI-EDT	HI-NDT
Low Immersion	LI-ST	LI-EDT	LI-NDT

Figure 2: Table of Experimental Groups

Measures

The four primary dependent variables in this study were measured using iconic

assessment tasks that have been used in a variety of previous studies on episodic memory. All

of these measures were presented through the Qualtrics system.

Room Size Task. For the Room Size task, participants were presented with the name

and a brief description of each of the eight rooms present in the environment. This task was

designed to assess participants' allocentric (survey) knowledge of the environment after their

exploration. Descriptions were carefully written to not describe the rooms based on size,

location, or the objects contained therein (these descriptions can be reviewed in Appendix 1).

Participants were then asked to rank the rooms in order from smallest to largest room based on

the room's area. These rankings were then scored based on accuracy using both exact and

relative scoring methods. Scoring method for this task is explained in the Analyses section.

Room Order Task. This task was designed to assess participants' egocentric (route) knowledge of the virtual environment, similar to the methods used in a virtual route learning task (cf. Pengas et al, 2010). The presentation and scoring of this task were identical to the Room Size task, with the exception that instead of ranking the rooms in order from smallest to largest, participants were asked to place the rooms in the order that they encountered them within the virtual environment. Scoring method for this task is detailed in the Analyses section.

Free Recall Task. In the Free Recall task, participants were presented with a series of blank spaces (up to 120). They were instructed to list as many objects from the environment as they could remember. After the participant had listed as many objects as they remembered, they were presented with a list of their responses. They were instructed, for each item recalled, to provide any additional details that they remembered about the object, such as size, color, or location.

Recognition and Object Location Task. For each trial, participants were presented with a picture of an object that had been rendered through the same engine as the virtual environment. They were asked to provide a judgment on whether they remembered seeing that object from the environment. Participants saw a total of 160 objects, half of which were objects taken from the virtual environment and the other half objects which were not present in the environment. For each object that participants reported having seen, they were subsequently shown a map of the virtual environment and asked to click the object's location on the map. For objects that the participant did not report having seen, they moved on to the next trial. Figure 3 shows a pictorial representation of this task.

Additional Measures. Subjective experience of presence was measured using the Presence Questionnaire, version 3 (PQ; Witmer and Singer, 1998). The PQ is a highly reliable measure of virtual presence (Chronbach's alpha = 0.91) that is widely used in studies

26

conducted with virtual environments (Witmer, Jerome, & Singer, 2005). Scores on the PQ have also consistently shown a weak, but significant, positive correlation with task performance across virtual tasks (Witmer & Singer, 1998). Subjective cognitive load was assessed using the single item measure of cognitive load recommended by Sweller (2010).
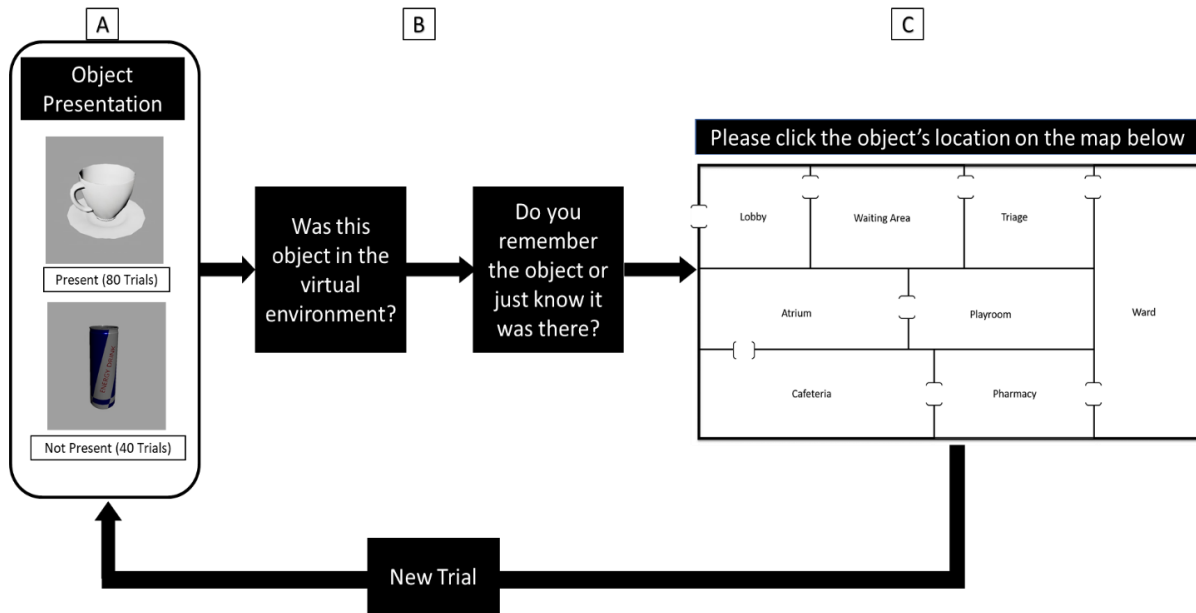

Figure 3: (A) Participants will be presented with an image of an object. Half of the objects presented will be taken from the environment and the other half will be distractor items which were not present. (B) Participants will then be asked to identify whether or not the presented object was present in the VE during the virtual task. If the participants responds "No," a new trial will begin and a new object will be presented. (C) If the participant responds "Yes," a map of the environment will be presented and the participant will be asked to click the object's location on the map.

Figure 3: Pictorial Representation of the Recognition and Object Location Task

Individual differences in visuospatial and verbal working memory were measured using matrix span and reading span tasks. These brief measures of working memory were administered with a JavaScript package developed by Stone and Towse (2015). The matrix span task displays a 4x4 grid of squares. Participants were shown a sequence of squares within the grid during the presentation phase and were asked to click that sequence during the recall phase. These sequences increased in length as the task progressed. In the reading span task, in the presentation phase participants were shown a number and then a sentence. They then made a judgment regarding whether the content of that sentence was true or false. Afterward

27

they were shown additional numbers and sentences in sequences of increasing length. In the recall phase, participants were asked to recall the numbers that they were shown prior to each sentence in sequential order.

Procedure

Participants recruited through the SONA system attended a brief, 1-hour, in-person experimental session. Each participant completed an informed consent process during which they had the experimental procedures detailed and given the opportunity to ask any questions that they may have. Following the consent procedure, participants were briefed on the virtual task.

The virtual task involved learning the details of the experimental VE as they traveled through it. During the virtual task, participants were immersed in the experimental VE according to their assigned experimental group. Participants explored the environment aiming to learn the spatial layout of the VE and the objects. Participants were instructed that they were doctors tasked with examining the way the hospital was currently using space and that at the end of the task they would be asked to report on layout and contents of VE, including the size and layout of rooms and the objects housed within those rooms. This instruction was intended to eliminate the known confound that exists between intentional and incidental encoding strategies for episodic memory, as young adults have been found to exhibit sub-clinical performance in incidental encoding paradigms (Plancher et al, 2010). Participants assigned to a dual task condition also received instructions for their secondary task, under the constraints outlined in the Conditions section. The total time to complete the virtual task was 10 minutes.

Following the virtual task, participants completed a Qualtrics survey that included the following measures, in order: Simulator Sickness Questionnaire, Single-Item Cognitive Load Measures, PQ, Room Size Task, Room Order Task, Free Recall Task, and Recognition and

28

Object Location Task. Following this questionnaire, participants completed Matrix Span and Reading Span tasks through the Tatool Java Applet (Stone & Towse, 2015).

Analyses

Manipulation, Version, and Group Checks: Manipulation checks were conducted using two 2x3 (Immersion x Task) factorial ANOVAs. The first ANOVA used raw cognitive load as a dependent variable. If the manipulation of cognitive load was successful, then there would be a main effect of Task and post-hoc comparisons should show that participants in the two dual-task conditions reported greater levels of cognitive load. The second ANOVA analyzed scores on the PQ as a dependent variable. If this manipulation was successful, there would be a main effect of group with the High Immersion group showing higher scores on the PQ.

Version checks were conducted for each outcome measure that included Object Memory or Feature Binding to ensure that the differences between VE Version A and Version B did not create a confound. These checks were conducted using one-way ANOVA for Object Recall, Object Recognition, Details per Object, and Object Location (distance score and correct room score) grouped by VE version. Similarly, group checks were conducted to ensure that baseline levels of spatial and verbal ability were evenly distributed between the 6 conditions. These checks were conducted using 2 one-way ANOVA for max span on the matrix span and reading span tasks analyzed by group assignment.

Hypothesis Testing: Hypotheses 1-3 were tested using a series of one-way, 2x3 (Immersion x Task) factorial ANOVAs with post-hoc comparisons using Tukey's Least Significant Difference (LSD). Hypothesis 1 was tested using measures of Spatial Memory. These analyses are described in the following paragraphs and summarized in Figure 4.

Hypothesis 1 was tested using measures of performance on the Room Size and Room Order tasks. 2x3 factorial ANOVAs were conducted using outcome scores for both relative and

exact scoring methods. For the exact scoring method, participants were awarded one point for each room that they provided the exact room ranking. So, for example, participants would only be awarded a point if they rated the largest room (Atrium) as number eight in the rankings. This method allowed for a minimum score of 0 points and a maximum score of 8.

For the relative scoring method, the rank provided by the participant for each room was compared to the rank provided for each other room. Participants were then awarded a point for each correct room-room comparison. For example, if a participant gave the largest room (Atrium) a rank of seventh largest room, they would be awarded 6 points based on their ranking that largest room as larger than six other rooms. This scoring method allowed for a more nuanced picture of participant room size comparison. Scores using this method ranged from 0 points to 28 points.

Hypothesis 2 was tested using performance on measures of Object Memory, the first portion of the Free Recall task and the Recognition portion of the Recognition and Object Location task. For the Free Recall task, two experimenters coded the number of correctly recalled objects from the virtual environment. For the Recognition task, participant scores were based on the number of correctly identified objects (objects that were both present in the environment and were endorsed as having been seen). Hypothesis H2a was tested using an ANOVA for number of correct objects on the Free Recall task and Hypothesis H2b was tested using an ANOVA for number of correctly identified objects from the Recognition task.

Hypothesis 3 was tested using performance on our measures of Feature Binding, the second portion of the Free Recall task and the Object Location portion of the Recognition and Object Location Task. For the Free Recall task, two experimenters coded the number of correctly recalled details associated with each correctly recalled object. The mean number of details recalled per object was then taken to provide a measure of binding for general details

(FB$_g$). For the Object Location portion of the Recognition and Object Location task, a distance score (DS) was calculated for each object based on the distance between the reported location and the objective location. This calculation used the X and Y coordinates on the provided map to calculate DS according to the following formula:

$$DS = \sqrt{\left(X_{reported} - X_{objective}\right)^2 + \left(Y_{reported} - Y_{objective}\right)^2}$$

The mean DS was then calculated across all correctly identified objects (objects that were in the environment that were endorsed as having been seen) to provide a measure of feature binding that was specific to the binding of object and spatial information (FB$_s$). Hypothesis H3a was tested using an ANOVA for FB$_g$, which is predicted to yield a main effect for Interaction and Workload. Hypothesis H3b was tested using an ANOVA for FB$_s$, which was predicted to yield a main effect for Interaction and an interaction of Presence and Workload.

During the experiment, it became apparent that the order of these measures may have influenced outcomes. Several participants reported that, despite the instructions provided in the Object Location task, the preceding Room Size and Room Order tasks had led them to answer with which room they remembered rather than an exact location. On this basis, a second scoring method was used for the Object Location data that compared reported locations to known room boundaries. Participants were awarded one point for each object for which the participant reported that they had seen the object and for which they placed the object in any location within the correct room.

For Hypothesis 4, gender differences were investigated through eleven 2x2x3 (Gender x Immersion x Task) factorial ANOVAs. A separate ANOVA was conducted for each of the following outcome variables: Room Size score (relative and exact), Room Order score (relative and exact), correctly recalled objects, falsely recalled objects, correctly recognized objects,

falsely recognized objects, details per object, object distance score, and correct room judgements.

For Hypothesis 5, comparable levels of instructional efficiency were analyzed using 2x3 (Immersion x Task) factorial ANOVA. Instructional efficiency (IE) for each participant was calculated for each of the 11 main outcome measures. Raw scores for cognitive load (C) and performances metrics (P) were converted to standardized z-scores. Then IE was calculated using the formula suggested by Paas and Van Merriënboer (1993):

$$Where\ P > C, IE = |P - C| \div \sqrt{2}$$

$$Where\ P < C, IE = \left(|P - C| \div \sqrt{2}\right) - 1$$

This method produces a positive value in instances where the ratio of performance to cognitive load was high and a negative value where the ratio of performance to cognitive load was low. This value suggests the performance that was generated by each unit of cognitive load.

CHAPTER THREE:
RESULTS

Manipulation, Version, and Group Checks

Manipulation Checks. A 2x3 (Immersion x Task) factorial ANOVA was conducted using PQ scores as the dependent variable. There was a significant main effect of Immersion (F=15.218, p<.0005, η_p^2=0.086). No main effect was observed for Task condition (F=0.011, p=0.989) and no interaction was observed for Immersion and Task conditions (F=0.731, p=0.483). Mean PQ for High Immersion groups (M=129.49, SD=15.109) was observed to be greater than mean PQ for Low Immersion groups (M=119.60, SD=16.332).

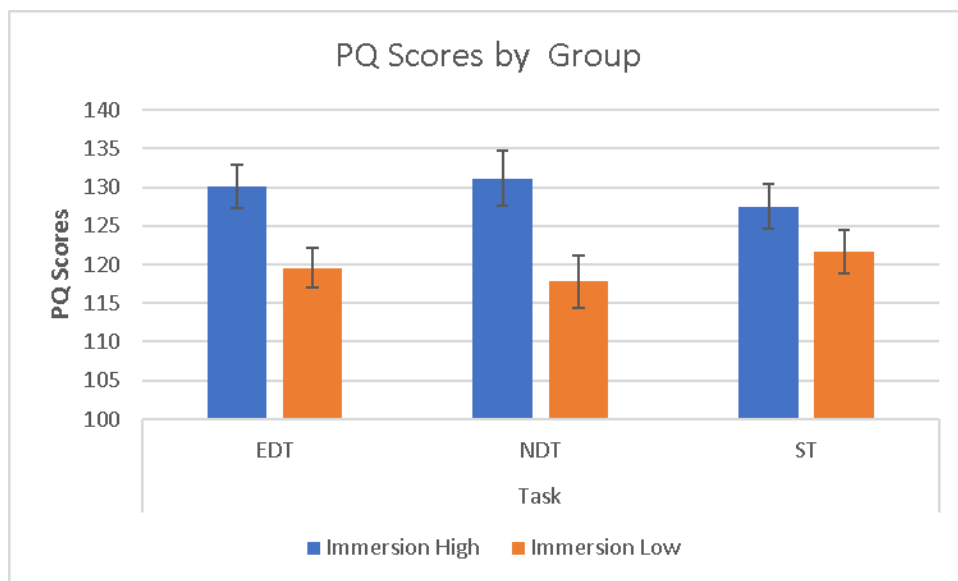| Measure	Source	df	MS	F	p	η_p^2
	Immersion	1	4110.482	15.218	<.0005*	.983
PQ	Task	2	3.042	.011	.989	.000
	Immersion x Task	2	197.482	.731	.483	.009

Figure 4: ANOVA for PQ

Figure 5: PQ Scores by Group

A 2x3 (Immersion x Task) factorial ANOVA was conducted using subjective cognitive load scores as the dependent variable. A main effect of Task was observed (F=39.51, $p<.0005$, η_p^2=.328). No main effect was observed for Immersion (F=2.660, p=0.105) and no significant interaction was observed between Immersion and Task groups (F=0.210, p=0.810). Post hoc comparisons showed a significant difference in cognitive load ratings between the two dual-task groups (EDT: M=4.02, SD=1.359, NDT: M=4.88, SD=1.562) and the single task group (M=2.45, SD=1.295). However, the difference between EDT and NDT was not shown to be significant.

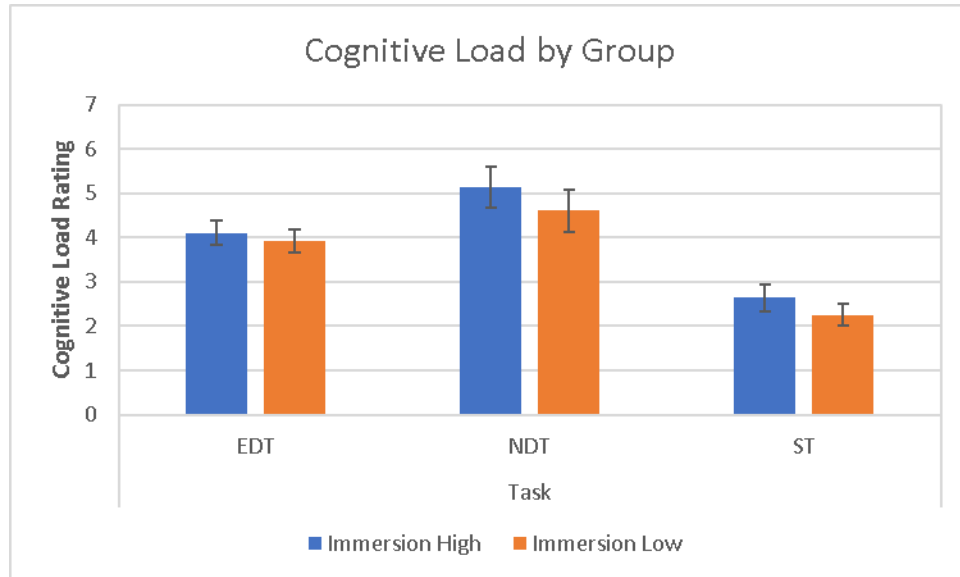| Measure	Source	df	MS	F	p	η_p^2
Single-Item Cognitive Load	Immersion	1	5.720	2.660	.105	.016
	Task	2	84.952	39.510	<.0005*	.328
	Immersion x Task	2	.452	.210	.810	.003

Figure 6: ANOVA for Cognitive Load

Figure 7: Cognitive Load by Group

Version Checks. A total of 7 independent sample t-tests were conducted using VE version as the grouping variable. A t-test for correctly remembered object on the free recall task did not show a significant difference between the two versions (t=0.223, p=0.824). A t-test for

correctly classified objects on the recognition task did not show significant differences between the two versions (t=-0.257, p=0.797). A t-test for details remembered per object did not show significant differences between versions (t=0.801, p=.425). A t-test for distance score on the object location task did not show significant differences between versions (t=1.622, p=.107). A t-test for number of correct room judgements on the object location task did not show significant differences between versions (t=-0.946, p=0.345). A t-test for falsely remembered object on the free recall task did not show significant differences between versions (t=0.778, p=0.438). A t-test for falsely recognized objects on the recognition task did not show significant differences between versions (t=0.778, p=0.885).

Group Checks. Two one-way ANOVAs were conducted using group assignment as the independent variable. A one-way ANOVA for max span on the matrix span task did not reveal any significant differences in the 6 groups (F=0.690, p=0.631). A one-way ANOVA for max span on the reading span task did not reveal any significant differences between groups (F=1.023, p=0.406).

Episodic Memory Measures

Spatial Memory. A 2x3 (Immersion x Task) factorial ANOVA was conducted using scores on the Room Size task (Relative scoring method) as the dependent variable. There was a significant main effect for Immersion (F=13.194, p<0.0005, η_p^2=0.075). No main effect was observed for Task (F=1.353, p=0.261), but a significant interaction of Immersion and Task was observed (F=3.249, p=0.041, η_p^2=0.039). Performance was higher in the High Immersion conditions and highest in the High Immersion, E-DT condition (M=21.04, SD=5.77). Performance was lower in Low Immersion groups and lowest in the Low Immersion, N-DT condition (M=15.39, SD=5.94).

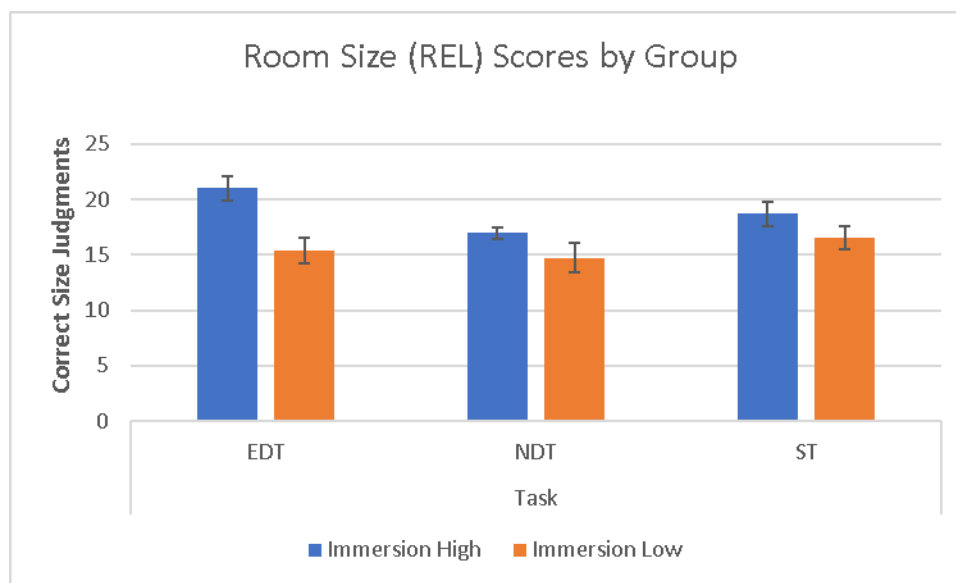| Measure	Source	df	MS	F	p	η_p^2
Room Size (Relative)	Immersion	1	418.006	13.194	<.0005*	.075
	Task	2	42.863	1.353	.261	.016
	Immersion x Task	2	102.935	3.249	.041*	.039

Figure 8: ANOVA for Room Size (Relative)

Figure 9: Room Size (REL) by Group

A 2x3 (Immersion x Task) factorial ANOVA was conducted using scores on the Room Size task (exact) as a dependent variable. This analysis showed a main effect of Immersion (F=3.904, p=.040, η_p^2=0.050). No main effect of Task (F=1.415, p=0.246) or interaction of Immersion and Task (F=2.875, p=0.059) were observed. Participants in High Immersion conditions (M=2.30, SD=1.656) were observed to have higher mean scores that participants in Low Immersion conditions (M=1.88, SD=1.555).

Measure	Source	df	MS	F	p	η_p^2
Room Size (Exact)	Immersion	1	8.292	3.904	(.04)*	.050
	Task	2	3.554	1.415	(.246)	.017
	Immersion x Task	2	7.220	2.875	(.059)	.034

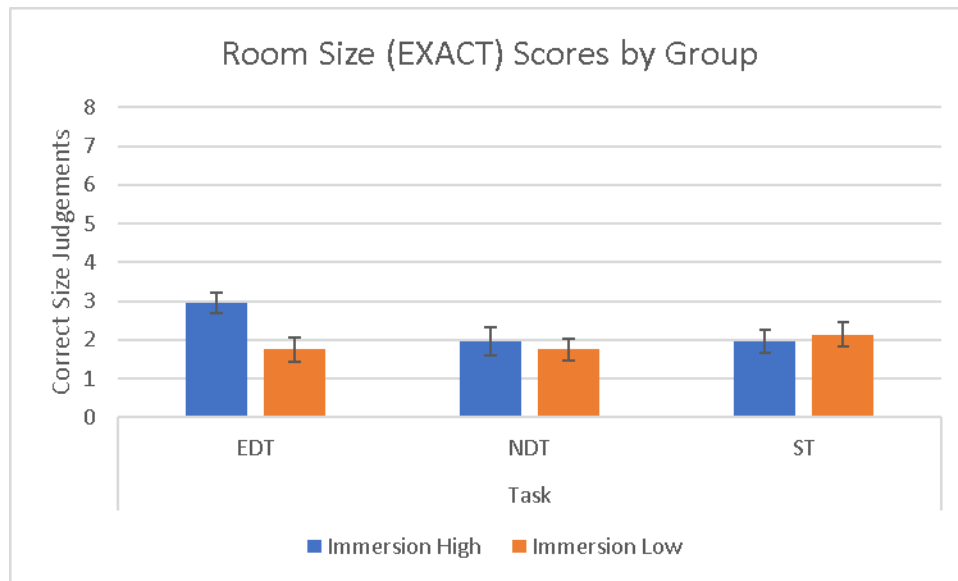*Figure 10: ANOVA for Room Size (Exact)*

Figure 11: Room Size (EXACT) Scores by Group

A 2x3 (Immersion x Task) factorial ANOVA was conducted using scores from the Room Order task (relative scoring method) as a dependent variable. There was a main effect for Immersion (F=115.363, p<0.0005, η_p^2=0.416) and for Task (F=4.324, p=0.015, η_p^2=0.051). A significant interaction of Immersion and Task (F=3.593, p=0.030, η_p^2=0.042) was also observed. Participants' scores were observed to be highest in the High Immersion condition with little differentiation (M=25.23, SD=3.283). In the Low Immersion condition (M=18.27, SD=5.189), participants' scores were lower with the worst score observed in the Low Immersion, Non-Ecological Dual-Task condition (M=15.75, SD=6.264).

Measure	Source	df	MS	F	p	η_p^2
Room Order (Relative)	Immersion	1	2030.095	115.363	(<.0005)*	.416
	Task	2	76.089	4.324	(.015)*	.051
	Immersion x Task	2	63.220	3.593	(.030)*	.042

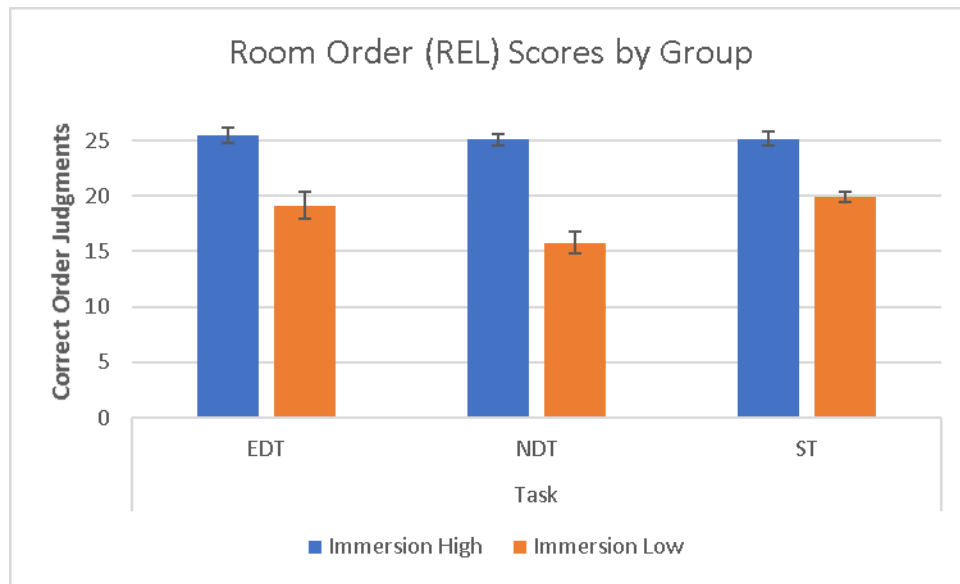*Figure 12: ANOVA for Room Order (Relative)*

Figure 13: Room Order (REL) Scores by Group

A 2x3 (Immersion x Task) factorial ANOVA was conducted using scores on the Room Order task (exact scoring) as a dependent variable. There was a significant main effect of Immersion (F=163.962, p<0.0005, η_p^2=0.503). No significant main effect of Task (F=1.262, p=0.286) was observed and there was no significant interaction of Immersion and Task (F=2.222, p=0.122). Higher scores were observed in the High Immersion group (M=6.71, SD=1.247) than in the Low Immersion group (M=3.74, SD=1.750).

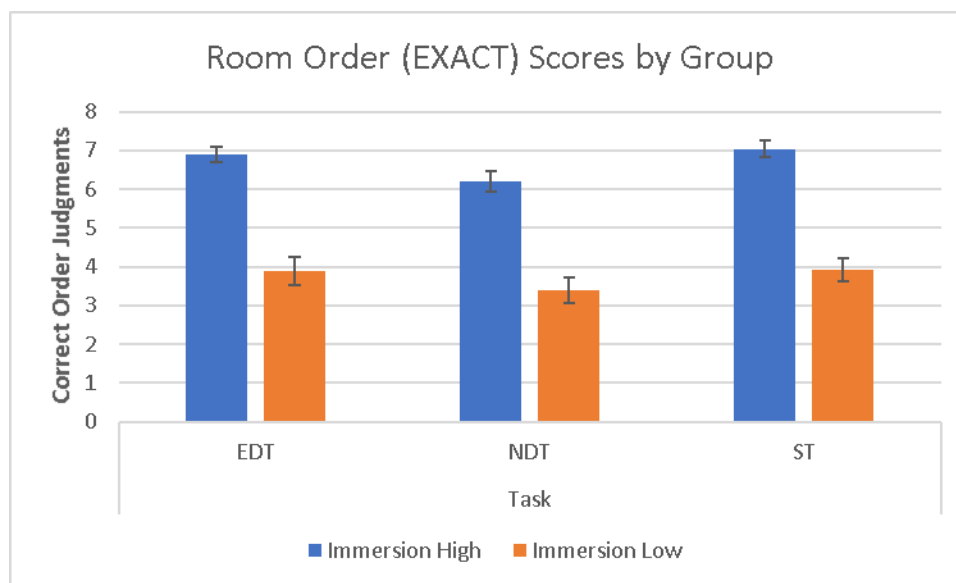| Measure	Source	df	MS	F	p	η_p^2
Room Order (Exact)	Immersion	1	372.024	163.962	(<.0005)*	.503
	Task	2	2.863	1.262	(.286)	.015
	Immersion x Task	2	5.042	2.222	(.112)	.027

Figure 14: ANOVA for Room Order (Exact)

Figure 15: Room Order (EXACT) Scores by Group

Object Memory. A 2x3 (Immersion x Task) factorial ANOVA was conducted using correctly remembered objects on the free recall task as a dependent variable. There was a main effect of Task (F=5.315, p=0.006, η_p^2=0.062). No main effect was observed for Immersion (F=1.703, p=0.194) and there was no significant interaction of Immersion and Task (F=2.577, p=0.079). Post hoc testing showed that participants assigned to EDT groups recalled significantly more objects (M=23.750, SD=11.672) than either ST (M=19.75, SD=10.261) or NDT (M=17.25, SD=10.307) groups.

39

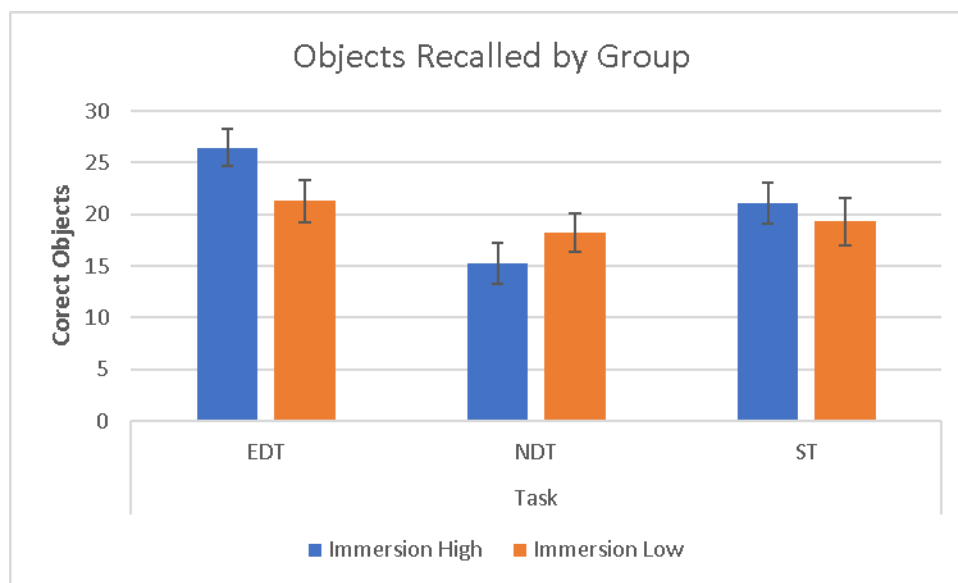| Measure	Source	df	MS	F	p	η_p^2
Object Recall (Correct)	Immersion	1	192.857	1.703	.194	.010
	Task	2	602.000	5.315	.006*	.062
	Immersion x Task	2	291.929	2.577	.079	.031

Figure 16: ANOVA for Object Recall (Correct)

Figure 17: Objects Recalled by Group (Correct)

A 2x3 (Immersion x Task) factorial ANOVA was conducted using falsely remembered objects on the free recall task as a dependent variable. No significant main effects were found for Immersion ($F=0.032$, $p=0.859$) or for Task ($F=0.088$, $p=0.916$). There was also no significant interaction of Immersion and Task ($F=0.664$, $p=.516$). Falsely remembered objects were few across all groups, with the average participant reporting fewer than one ($M=0.76$, $SD=1.288$).

Measure	Source	df	MS	F	p	η_p^2
Object Recall (FA)	Immersion	1	.054	.032	.859	.000
	Task	2	.149	.088	.916	.001
	Immersion x Task	2	1.125	.664	.516	.008

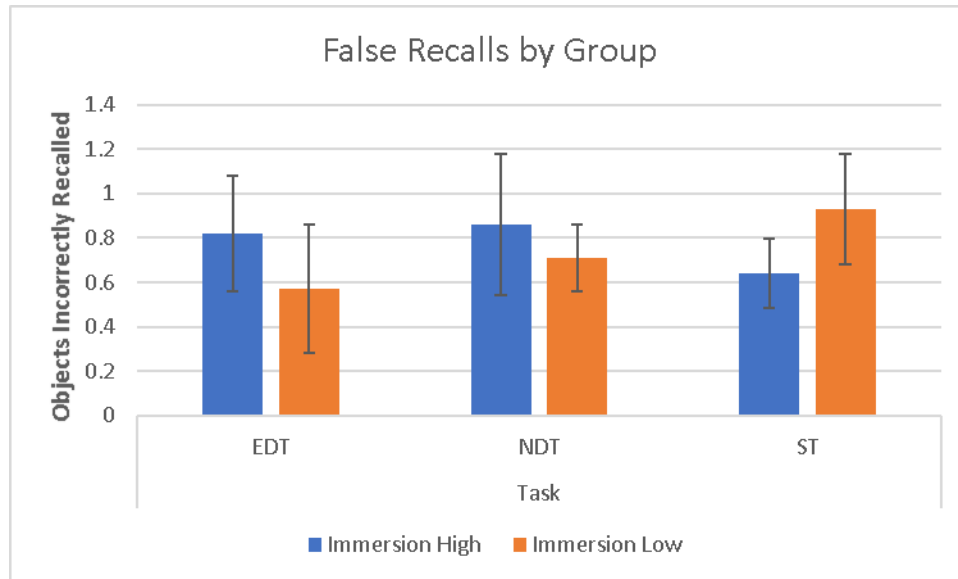*Figure 18: ANOVA for Object Recall (FA)*

Figure 19: False Recalls by Group

A 2x3 (Immersion x Task) factorial ANOVA was conducted using correct identifications from the object recognition task as a dependent variable. There was a main effect of Immersion (F=7.300, p=0.008, η_p^2=0.043) and main effect of Task (F=8.453, p<0.0005, η_p^2=0.094). There was also a significant interaction between Immersion and Task (F=3.710, p=0.026, η_p^2=0.044). Mean number of objects recognized was highest in the Low Immersion, EDT condition (M=88.23, SD=13.206) and lowest in the High Immersion, NDT condition (M=70.07, SD=13.968).

Measure	Source	df	MS	F	p	η_p^2
Object Recognition (Correct)	Immersion	1	1090.381	7.300	.008*	.043
	Task	2	1262.625	8.453	<.0005*	.094
	Immersion x Task	2	555.542	3.719	.026*	.044

Figure 20: ANOVA for Object Recognition

41
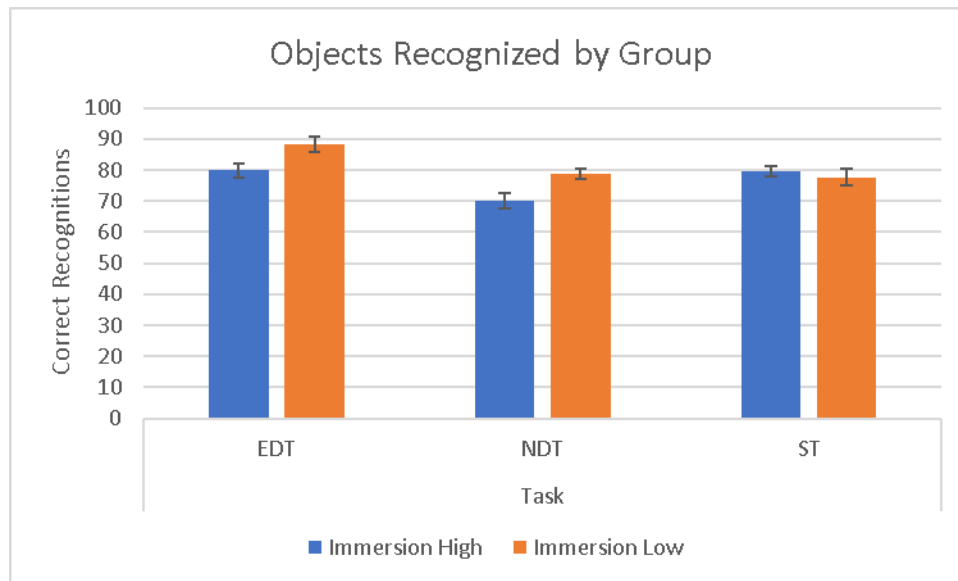
Figure 21: Correct Recognitions by Group

A 2x3 (Immersion x Task) factorial ANOVA was conducted using falsely recognized objects on the Object Recognition task as a dependent variable. No significant main effects were observed for Immersion (F=2.009, p=0.158) or Task (F=1.220, p=0.298), nor was there a significant interaction of Immersion and Task (F=0.131, p=0.877). The mean number of falsely recognized objects for the sample was 11.09(7.995).

Measure	Source	df	MS	F	p	η_p^2
Object Recognition (FA)	Immersion	1	128.625	2.009	.158	.012
	Task	2	78.125	1.220	.298	.015
	Immersion x Task	2	8.375	.131	.877	.002

Figure 22: ANOVA for Object Recognition (FA)

42
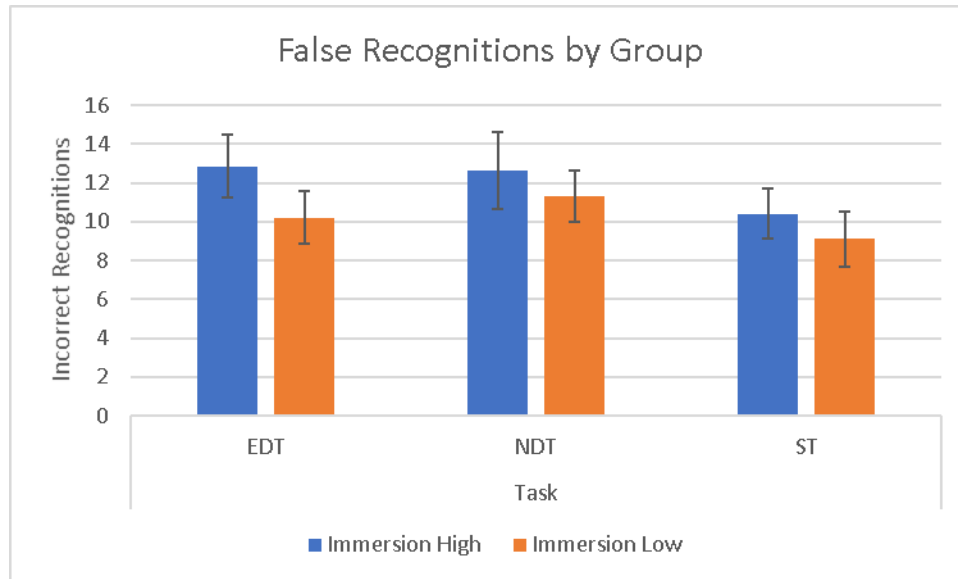
Figure 23: False Recognitions by Group

Feature Binding. A 2x3 (Immersion x Task) factorial ANOVA was conducted using reported details per object from the free recall task as a dependent variable. No main effect was found for Immersion (F=0.242, p=0.624) or Task (F=0.252, p=0.133). There was also no significant interaction of Immersion and Task (F=2.042, p=0.133). The mean number of details per object for this sample was 1.44(0.62).

Measure	Source	df	MS	F	p	η_p^2
Details Per Object	Immersion	1	.094	.242	.624	.001
	Task	2	.098	.252	.777	.003
	Immersion x Task	2	.792	2.042	.133	.025

Figure 24: ANOVA for Details Per Object
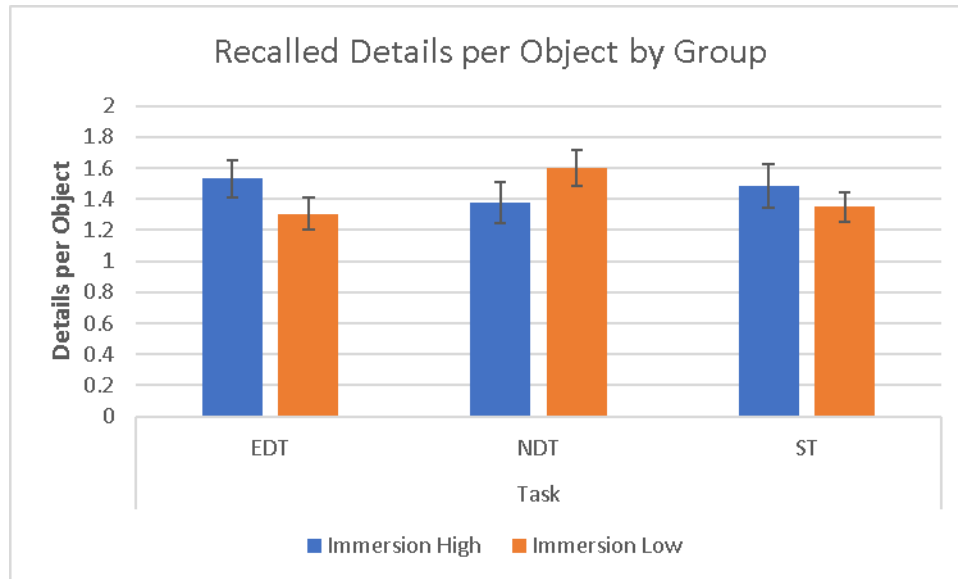
Figure 25: Recalled Details per Object by Group

A 2x3 (Immersion x Task) factorial ANOVA was computed using participant's mean distance score on the Object Location task as a dependent variable. There was a significant main effect of Task (F=3.818. p=0.024, η_p^2=0.045). However, no significant main effect of Immersion (F=3.179, p=0.076) and no significant interaction of Immersion and Task (F=2.444, p=0.090) were observed. Post hoc testing showed that there was a significant difference between EDT (M=156.114, SD=86.741) and NDT (M=199.547, SD=102.674) conditions, but not ST (M=183.420, SD=75.590).

Measure	Source	df	MS	F	p	η_p^2
Object Location (Distance Score)	Immersion	1	22475.647	3.179	.076	.019
	Task	2	26993.206	3.818	.024*	.045
	Immersion x Task	2	17284.640	2.444	.090	.029

Figure 26: ANOVA for Object Location (Distance Score)
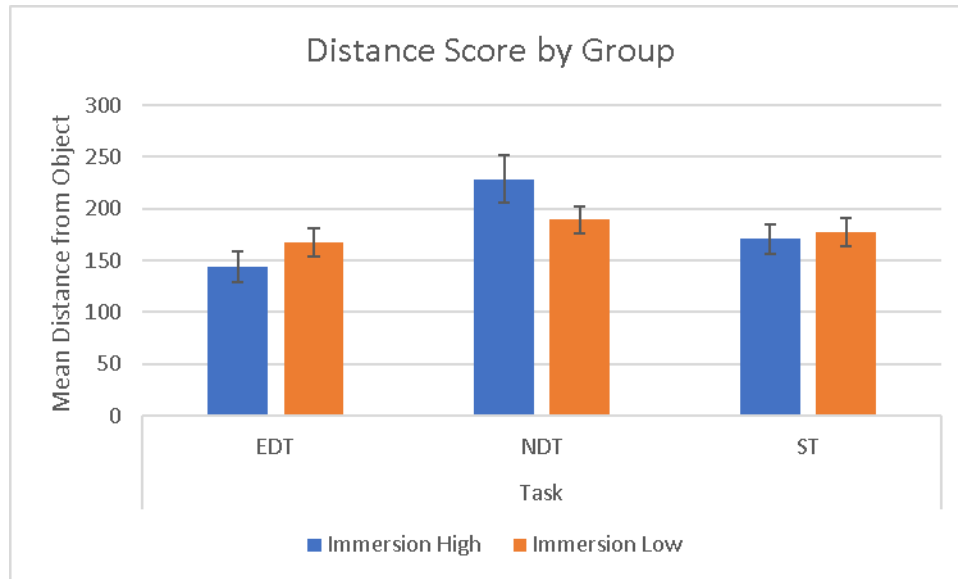
Figure 27: Distance Score by Group

A 2x3 (Immersion x Task) factorial ANOVA was conducted using number of correct room judgements on the Object Location task as a dependent variable. A significant main effect was observed for both Immersion (F=5.441, p=0.021, η_p^2=0.032) and Task (F=3.731, p=0.026, η_p^2=0.044). There was also a significant interaction of Immersion and Task (F=5.992, p=0.003, η_p^2=0.069). Participants in the High Immersion (M=48.39, SD=15.675) and EDT (M=51.195, SD=16.802) provided the most correct room judgements, with participants in the HI-EDT group providing the highest number (M=55.39, SD=16.082) and participants in the LI-NDT group providing the fewest (M=38.29, SD=13.478).

Measure	Source	df	MS	F	p	η_p^2
Object Location (Correct Room)	Immersion	1	1070.095	5.441	.021*	.032
	Task	2	733.649	3.731	.026*	.044
	Immersion x Task	2	1178.435	5.992	.003*	.069

Figure 28: ANOVA for Object Location (Correct Room)

45
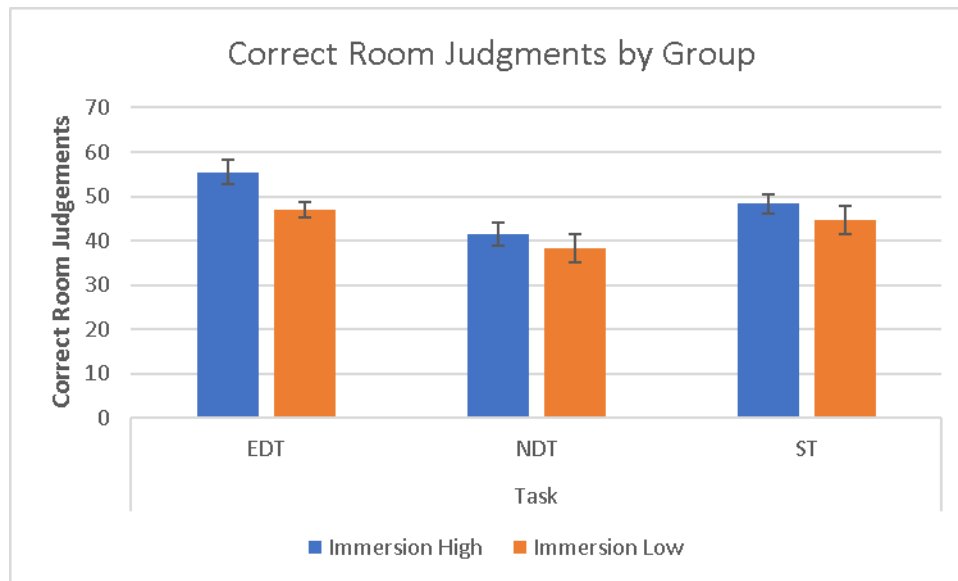
Figure 29: Correct Room Judgements by Group

Know/Remember Ratings. A 2x3 (Immersion x Task) ANOVA was conducted to assess differences in participants' subjective vividness of remembering by condition. No significant differences were observed for Immersion (F=0.128, p=0.635) or Task (F=2.79, p=0.064) and no significant Immersion by Task interaction was observed (F=1.077, p=0.343). However, posthoc analysis using the LSD method did show a significant difference between Single Task (M=1.65, SD=0.749) and Ecological Dual Task (M=1.96, SD=0.747) groups when collapsed across Immersion conditions.

Exploratory Analyses

Gender Differences. A 2x2x3 (Gender x Immersion x Task) factorial ANOVA was conducted using performance on the room size task (relative scoring) as a dependent variable. Significant main effects of Immersion (F=13.725, p<0.0005, η_p^2=0.081) and Gender (F=6.689, p=0.011, η_p^2=0.041) were observed. There was also a significant interaction of Task and Immersion (F=3.380, p=0.037, η_p^2=0.042). No significant main effect of Task (F=1.407, p=0.248) was observed, nor were there any significant interaction effects for the Gender by Task

(F=0.149, p=0.862) or Gender by Immersion by Task (F=2.154, p=0.119) terms. Male participants (M=18.24, SD=5.529) showed higher average scores for this task than did female participants (M=16.04, SD=6.112).

Measure	Source	df	MS	F	p	η_p^2
	Gender	1	203.720	6.689	.011*	.041
	Immersion	1	418.006	13.725	<.0005*	.081
	Task	2	42.863	1.407	.248	.018
Room Size (Relative)	Gender x Immersion	1	37.149	1.220	.271	.008
	Gender x Task	2	4.542	0.149	.862	.002
	Task x Immersion	2	102.935	3.380	.037*	.042
	Gender x Task x Immersion	2	65.613	2.154	.119	.027

Figure 30: ANOVA for Room Size (Relative) With Gender

Figure 31: Room Size Scores (REL) by Group and Gender

A 2x2x3 (Gender x Immersion x Task) factorial ANOVA was conducted using performance on the room size task (exact scoring) as a dependent variable. No significant main effects were observed for Gender (F=2.929, p=0.089), Immersion (F=2.929, p=0.089), or Task (F=1.427, p=.243). There were also no significant interaction terms for Gender by Immersion

47

(F=1.265, p=.262), Gender by Task (F=0.232, p=0.793), Task by Immersion (F=2.900, p=.058), or Gender by Task by Immersion (F=1.380, p=0.255).

Measure	Source	df	MS	F	p	η_p^2
Room Size (Exact)	Gender	1	7.292	2.929	.089	.018
	Immersion	1	7.292	2.929	.089	.018
	Task	2	3.554	1.427	.243	.018
	Gender x Immersion	1	3.149	1.265	.262	.008
	Gender x Task	2	0.577	0.232	.793	.003
	Task x Immersion	2	7.220	2.900	.058	.036
	Gender x Task x Immersion	2	3.435	1.380	.255	.017

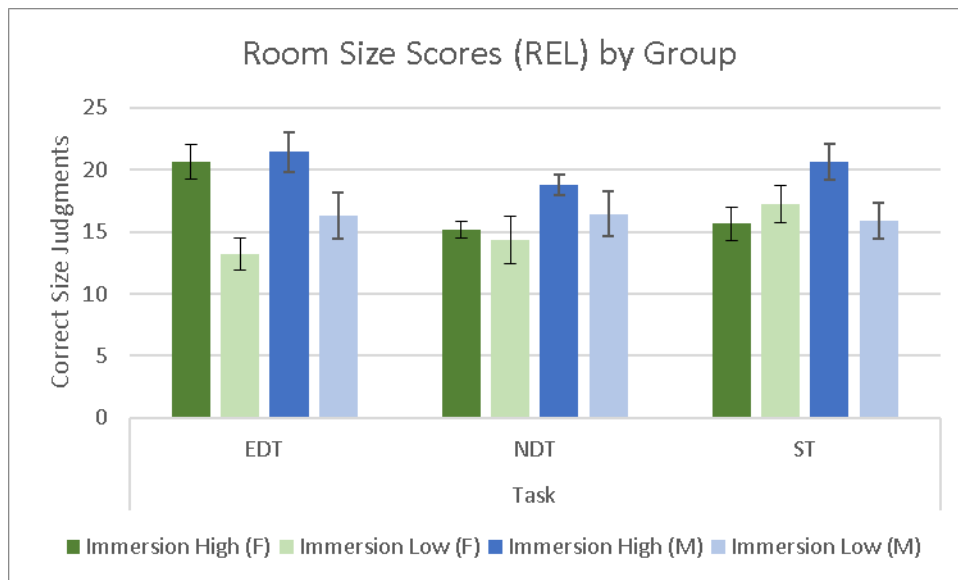*Figure 32: ANOVA for Room Size (Exact) with Gender*

Figure 33: Room Size Scores (EXACT) by Group and Gender

A 2x2x3 (Gender x Immersion x Task) factorial ANOVA was conducted using performance on the room order task (relative scoring) as a dependent variable. There was a significant main effect for Gender (F=6.582, p=0.011, η_p^2=0.040), Immersion (F=121.366, p<0.0005, η_p^2=0.438), and Task (F=4.549, p=0.012, η_p^2=0.065). A significant interaction of Immersion and Task was also observed (F=3.780, p=0.025, η_p^2=0.046). However, no significant

interaction was observed between Gender and Immersion (F=0.241, p=0.624) or Gender and Task (F=2.334, p=0.100), and no significant three-way interaction was observed (F=1.469, p=0.233). Generally, male participants scored higher (M=22.56, SD=4.427) where female participants scored lower (M=20.94, SD=6.422).

Measure	Source	df	MS	F	p	η_p^2
Room Order (Relative)	Gender	1	110.095	6.582	.011*	.040
	Immersion	1	2030.095	121.366	<0.0005*	.438
	Task	2	76.089	4.549	.012*	.055
	Gender x Immersion	1	4.024	.241	.624	.002
	Gender x Task	2	39.042	2.334	.100	.029
	Task x Immersion	2	63.220	3.780	.025*	.046
	Gender x Task x Immersion	2	24.577	1.469	.233	.018

Figure 34: ANOVA for Room Order (Relative) with Gender
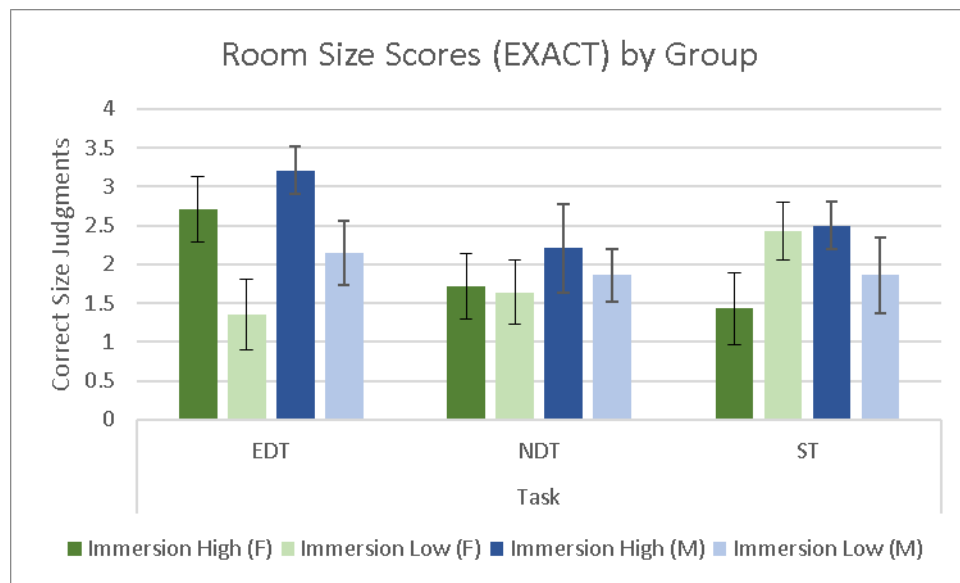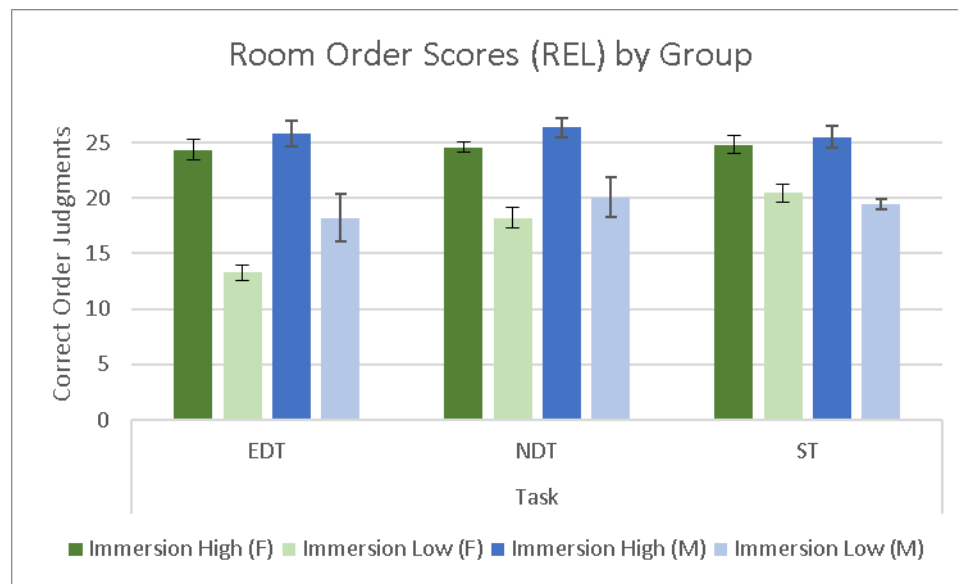

Figure 35: Room Order Score (EXACT) by Group and Gender

A 2x2x3 (Gender x Immersion x Task) factorial ANOVA was conducted using scores on the room order task (exact scoring) as a dependent variable. There was a significant main effect of Immersion (F=182.995, p<0.0005, η_p^2=0.540). No main effect was observed for Gender

(F=3.385, p=0.068) or Task (F=1.408, p=0.248). There was no significant interaction of Gender

and Immersion (F=4.024, p=0.161) nor was there a significant interaction of Task and

Immersion (F=2.480, p=0.087). However, a significant interaction was observed between

Gender and Task (F=3.552, p=0.031, η_p^2=0.044) as well as a significant three-way interaction of

Gender, Immersion, and Task (F=6.169, p=0.003, η_p^2=0.073). Between group comparisons

show that while High Immersion groups performed homogenously, in Low Immersion groups

there was a gender difference in response to task conditions. For these conditions, female

participants performed best in the single task group (M=4.75, SD=2.048), where male

participants performed worst (M=3.36, SD=1.499).

| Measure	Source	df	MS	F	p	η_p^2
Room Order (Exact)	Gender	1	6.881	3.385	.068	.021
	Immersion	1	372.024	182.995	<.0005*	.540
	Task	2	2.863	1.408	.248	.018
	Gender x Immersion	1	4.024	1.979	.161	.013
	Gender x Task	2	7.220	3.552	.031*	.044
	Task x Immersion	2	5.042	2.480	.087	.031
	Gender x Task x Immersion	2	12.542	6.169	.003*	.073

Figure 36: ANOVA for Room Order (Exact) with Gender
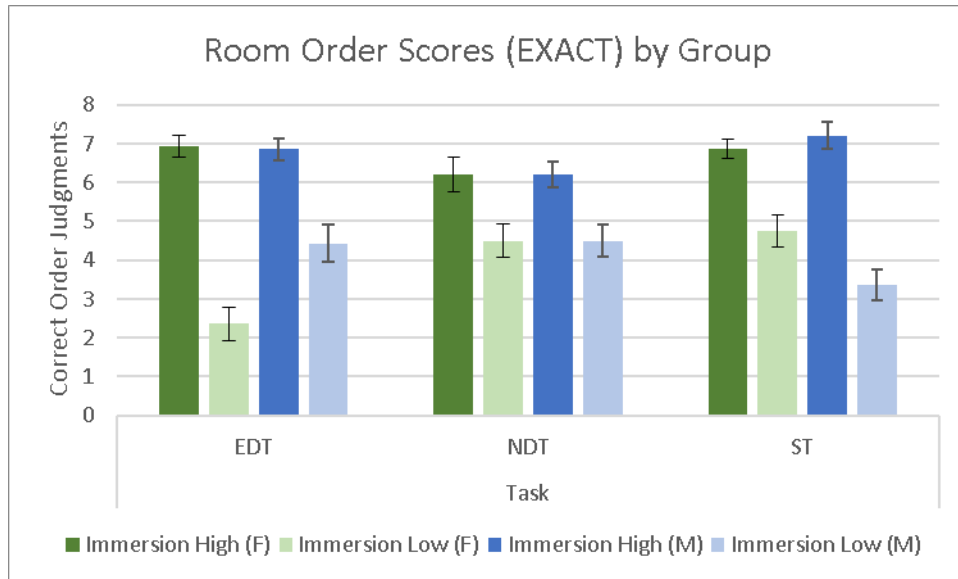
Figure 37: Room Order Scores (EXACT) by Group and Gender

A 2x2x3 (Gender x Immersion x Task) factorial ANOVA was conducted using correct objects from the free recall task as a dependent variable. There was a significant main effect of Task (F=5.396, p=0.005, η_p^2=0.065). However, no significant main effect was observed for Gender (F=2.008, p=0.158) or Immersion (F=1.729, p=0.190). No significant interactions were observed for Gender and Immersion (F=0.669, p=0.415), Gender and Task (F=2.109, p=0.076), or Task and Immersion (F=2.617, p=0.076). Additionally, no significant three-way interaction was observed (F=0.792, p=0.455).

51

Measure	Source	df	MS	F	p	η_p^2
Object Recall (CORRECT)	Gender	1	224.024	2.008	.158	.013
	Immersion	1	192.857	1.729	.190	.011
	Task	2	602.000	5.396	.005*	.065
	Gender x Immersion	1	74.667	0.669	.415	.004
	Gender x Task	2	235.310	2.109	.125	.026
	Task x Immersion	2	291.929	2.617	.076	.032
	Gender x Task x Immersion	2	88.310	0.792	.455	.010

Figure 38: ANOVA for Object Recall (Correct) with Gender

Figure 39: Objects Recalled by Group and Gender

A 2x2x3 (Gender x Immersion x Task) factorial ANOVA was conducted using falsely remembered objects from the free recall task as a dependent variable. No significant main effect was observed for Gender (F=3.530, p=0.062), Immersion (F=0.033, p=0.856), or Task (F=0.092, p=0.912). There was no significant interaction of Gender and Immersion (F=0.092, p=0.762) or of Task and Immersion (F=0.694, p=0.501) and no significant three-way interaction of Gender, Immersion, and Task (F=0.180, p=0.835). However, there was a significant interaction of Gender and Task (F=4.676, p=0.011, η_p^2=0.057), where female participants falsely

recalled more objects than male participants, but in EDT conditions (female: M=1.357, SD=1.83, male: M=0.21, SD=0.499).

Measure	Source	df	MS	F	p	η_p^2
Object Recall (FALSE)	Gender	1	5.720	3.530	.062	.022
	Immersion	1	0.054	0.033	.856	.000
	Task	2	0.149	0.092	.912	.001
	Gender x Immersion	1	0.149	0.092	.762	.001
	Gender x Task	2	7.577	4.676	.011*	.057
	Task x Immersion	2	1.125	.694	.501	.009
	Gender x Task x Immersion	2	0.292	0.180	.835	.002

Figure 40: ANOVA for Object Recall (False) with Gender

Figure 41: False Objects Recalled by Group and Gender

A 2x2x3 (Gender x Immersion x Task) factorial ANOVA was conducted using correct recollections from the object recall task as a dependent variable. There was a significant main effect of Immersion (F=7.623, p=0.007, η_p^2=.046) and of Task (F=8,711, p<0.0005, η_p^2=0.100). No significant main effect was observed for Gender (F=2.994, p=0.086). There was a significant interaction of Immersion and Task (F=3.833, p=0.024, η_p^2=0.047) but no significant interaction

was observed between Gender and Immersion (F=3.598, p=0.060) or between Gender and Task (F=1.936, p=0.148). There was also no significant three-way interaction of Gender, Immersion, and Task (F=0.244, F=0.783).

Measure	Source	df	MS	F	p	η_p^2
Object Recognition (CORRECT)	Gender	1	433.929	2.994	.086	.019
	Immersion	1	1090.381	7.523	.007*	.046
	Task	2	1262.625	8.711	<.0005*	.100
	Gender x Immersion	1	521.524	3.598	.060	.023
	Gender x Task	2	280.661	1.936	.148	.024
	Task x Immersion	2	555.542	3.833	.024*	.047
	Gender x Task x Immersion	2	35.435	0.244	.783	.003

Figure 42: ANOVA for Object Recognition (Correct) with Gender
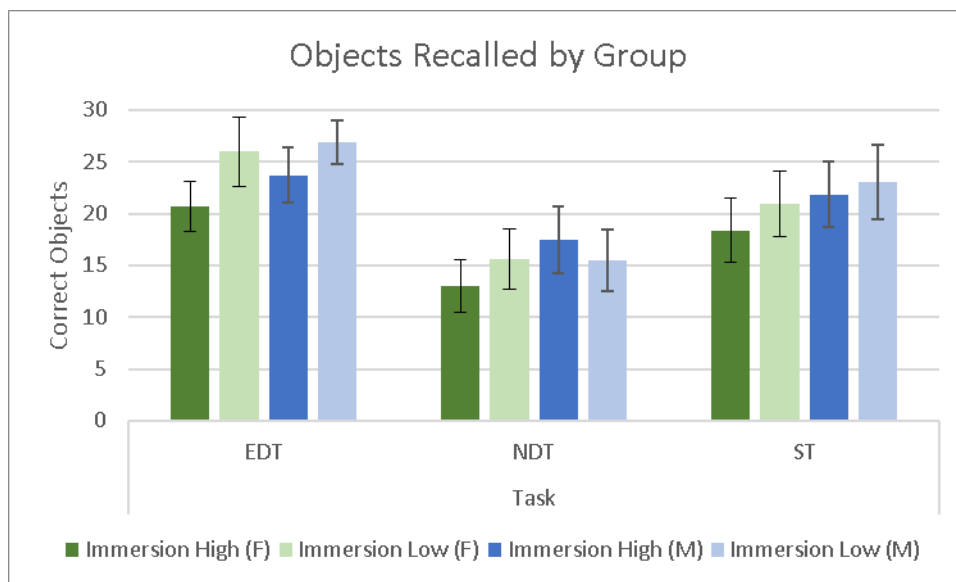

Figure 43: Objects Recognized by Group and Gender

A 2x2x3 (Gender x Immersion x Task) factorial ANOVA was conducted using falsely recognized objects from the object recognition task as a dependent variable. No significant main effect was observed for Gender (F=0.033, p=0.856), Immersion (F=1.967, p=0.136), or Task (F=1.194, p=0.306). Additionally, no significant interactions were observed between Gender and

54

Immersion ($F=1.809$, $p=0.181$), Gender and Task ($F=0.096$, $p=0.909$), or Immersion and Task ($F=0.128$, $p=0.880$). No significant three-way interaction was observed ($F=0.271$, $p=0.763$).

Measure	Source	df	MS	F	p	η_p^2
Object Recognition (FALSE)	Gender	1	2.149	0.033	.856	.000
	Immersion	1	128.625	1.967	.136	.012
	Task	2	78.125	1.194	.306	.015
	Gender x Immersion	1	118.339	1.809	.181	.011
	Gender x Task	2	6.256	0.096	.909	.001
	Task x Immersion	2	8.378	0.128	.880	.002
	Gender x Task x Immersion	2	17.696	0.271	.763	.003

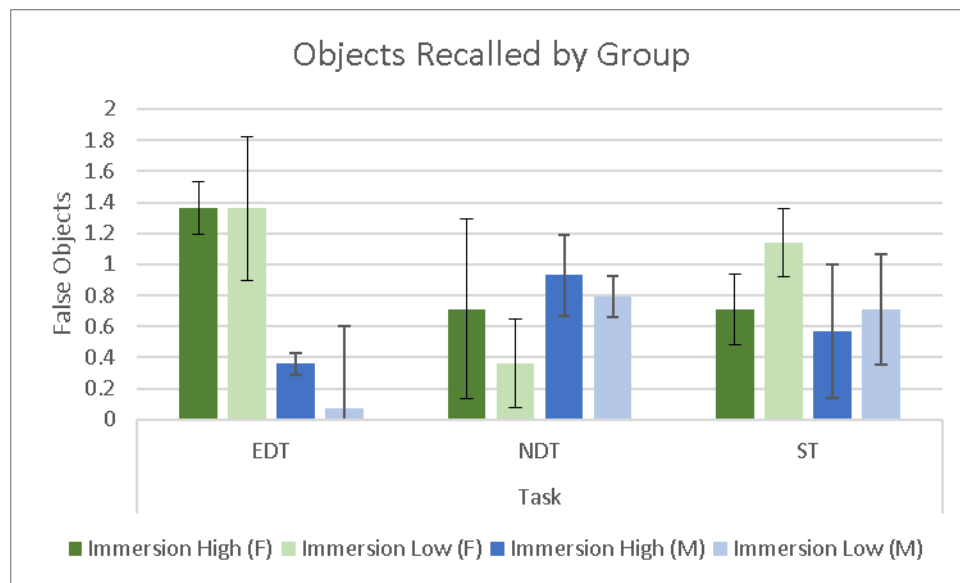*Figure 44: ANOVA for Object Recognition (False) with Gender*
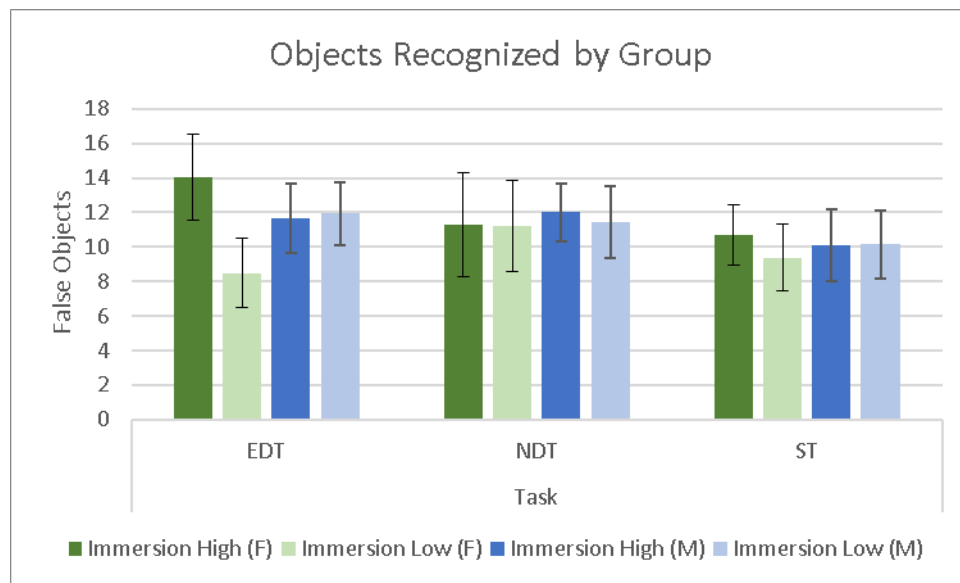
Figure 45: False Objects Recognized by Group and Gender
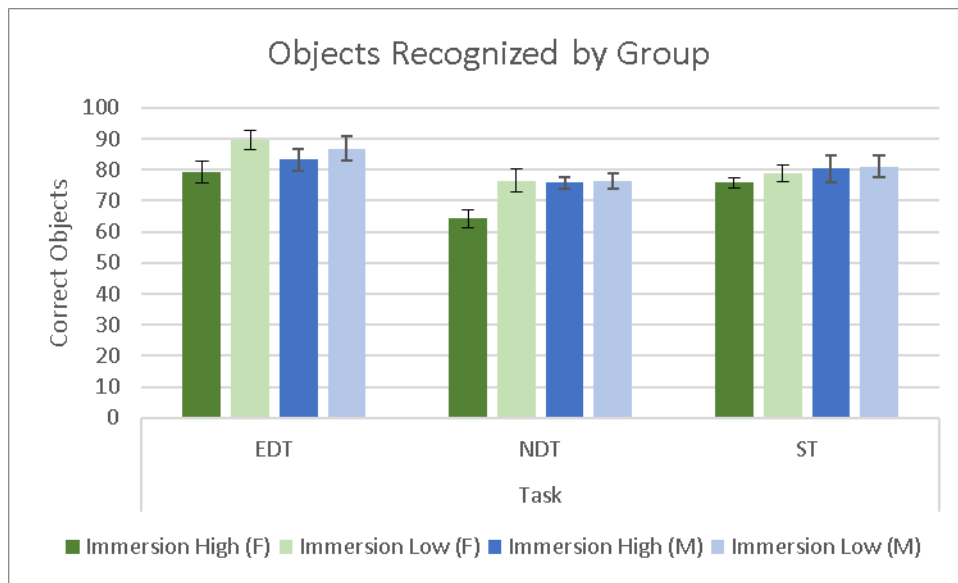
A 2x2x3 (Gender x Immersion x Task) factorial ANOVA was conducted using falsely recognized objects from the object recognition task as a dependent variable. No significant main effect was observed for Gender ($F=0.341$, $p=0.560$), Immersion ($F=0.240$, $p=0.625$), or Task ($F=0.250$, $p=0.779$). Additionally, no significant interactions were observed between Gender and

Immersion (F=0.423, p=0.516), Gender and Task (F=1.283, p=0.280), or Immersion and Task (F=2.026, p=0.135). No significant three-way interaction was observed (F=0.690, p=0.503).

Measure	Source	df	MS	F	p	η_p^2
	Gender	1	0.134	0.341	.560	.002
	Immersion	1	0.134	0.240	.625	.002
	Task	2	0.098	0.250	.779	.003
Binding Recall (Details per Object)	Gender x Immersion	1	0.166	0.423	.516	.003
	Gender x Task	2	0.502	1.283	.280	.016
	Task x Immersion	2	0.792	2.026	.135	.025
	Gender x Task x Immersion	2	0.270	0.690	.503	.009

Figure 46: ANOVA for Binding Recall (Details per Object) with Gender
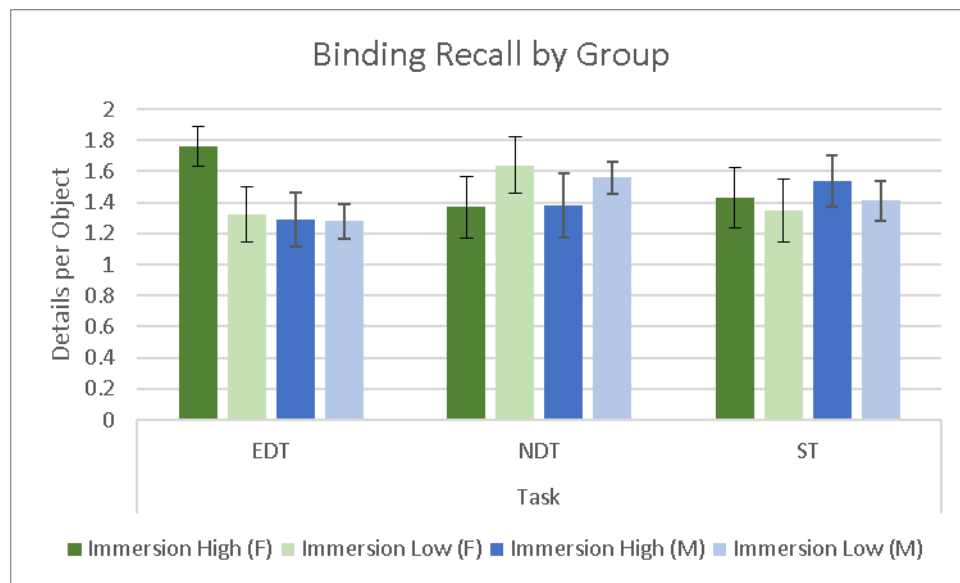
Figure 47: Binding Recall by Group and Gender

A 2x2x3 (Gender x Immersion x Task) factorial ANOVA was conducted using average distance score on the object location task as a dependent variable. A significant main effect of Gender (F=13.213, p<0.0005, η_p^2=0.078) and Task (F=4.183, p=0.017, η_p^2=0.051) was observed. However, there was no significant main effect of Immersion (F=3.483, p=0.064). No significant interaction of Gender and Immersion (F=1.342, p=0.248), Gender and Task

56

(F=2.892, p=0.058), or Task and Immersion (F=2.678, p=0.072) were observed and there was

no significant three-way interaction (F=0.575, p=0.564).

Measure	Source	df	MS	F	p	η_p^2
Object Location (Distance Score)	Gender	1	85273.896	13.213	<.0005*	.078
	Immersion	1	22475.647	3.483	.064	.022
	Task	2	26993.206	4.183	.017*	.051
	Gender x Immersion	1	8660.335	1.342	.248	.009
	Gender x Task	2	18665.400	2.892	.058	.036
	Task x Immersion	2	17284.640	2.678	.072	.033
	Gender x Task x Immersion	2	3712.886	.575	.564	.007

Figure 48: ANOVA for Object Location (Distance Score) with Gender
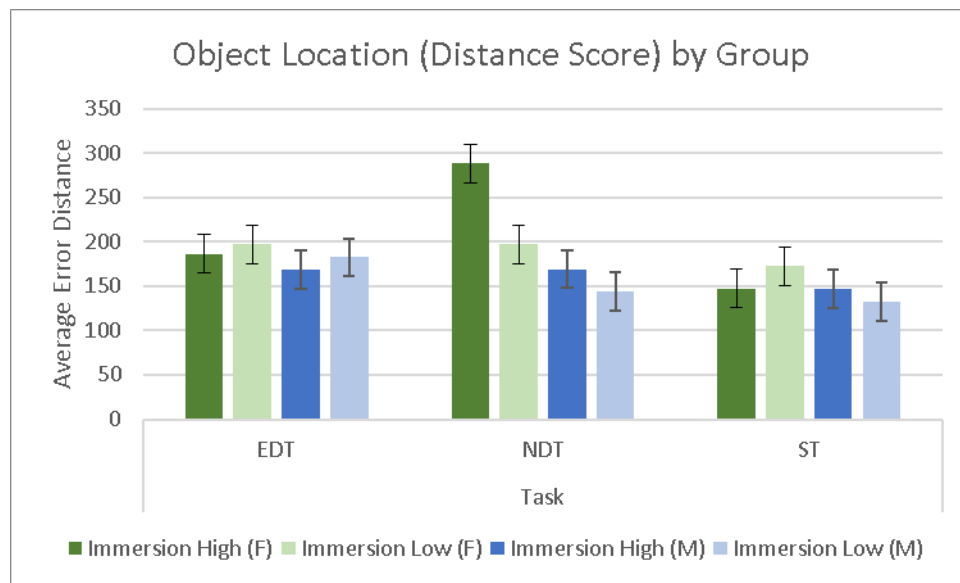
Figure 49: Object Location (Distance Score) by Group and Gender

A 2x2x3 (Gender x Immersion x Task) factorial ANOVA was conducted using correct

room ratings from the object location task as a dependent variable. A significant main effect was

observed for both Immersion (F=5.414, p=0.027, η_p^2=0.034) and Task (F=3.712, p=0.027,

η_p^2=0.045). However, there was no significant main effect for Gender (F=1.973, p=0.162). There

was a significant interaction of Task and Immersion (F=5.962, p=0.003, η_p^2=0.071). However,

no significant interaction was observed between Gender and Immersion (F=0.642, p=0.424) or between Gender and Task (F=0.560, p=0.573). No significant three-way interaction was observed (F=0.719, p=0.489).

Measure	Source	df	MS	F	p	η_p^2
Object Location (Correct Room)	Gender	1	390.095	1.973	.162	.012
	Immersion	1	1070.095	5.414	.021*	.034
	Task	2	733.649	3.712	.027*	.045
	Gender x Immersion	1	126.881	0.642	.424	.004
	Gender x Task	2	110.613	0.560	.573	.007
	Task x Immersion	2	1178.435	5.962	.003*	.071
	Gender x Task x Immersion	2	142.185	0.719	.489	.009

Figure 50: ANOVA for Object Location (Correct Room) with Gender
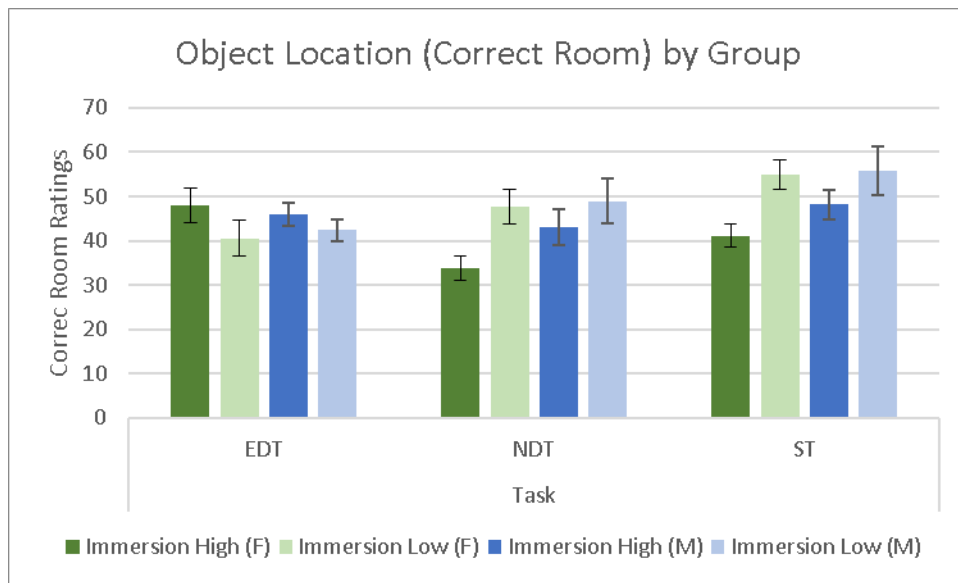
Figure 51: Object Location (Correct Room) by Group and Gender

Instructional Efficiency. As an exploratory measure, for each outcome measure that showed a significant main effect or interaction, instructional efficiency of each condition was calculated and analyzed. A 2x3 (Immersion x Task) ANOVA was conducted using instructional efficiency

58

for the Room Size (relative scoring) as a dependent variable. There was a main effect of Task (F=15.006, p<0.0005, η_p^2=0.156). However, there was no significant main effect for Immersion (F=2.516, p=0.115) and no significant interaction of Immersion and Task conditions (F=1.153, p=0.318). Post hoc testing showed that IE was greatest in ST conditions and lowest in NDT conditions.

Measure	Source	df	MS	F	p	η_p^2
Instructional Efficiency for Room Size (Relative)	Immersion	1	2.221	2.516	.115	.015
	Task	2	13.247	15.006	<.0005*	.156
	Immersion x Task	2	1.018	1.153	.318	.014

Figure 52: ANOVA for Instructional Efficiency for Room Size (Relative)

A 2x3 (Immersion x Task) ANOVA was conducted using instructional efficiency for the Room Size (exact scoring) as a dependent variable. There was a main effect of Task (F=12.829, p<0.0005, η_p^2=0.137). However, there was no significant main effect for Immersion (F=0.059, p=0.809) and no significant interaction of Immersion and Task conditions (F=1.205, p=0.302). Post hoc testing showed that IE was greatest in ST conditions and lowest in NDT conditions.

Measure	Source	df	MS	F	p	η_p^2
Instructional Efficiency for Room Size (Exact)	Immersion	1	.052	.059	.809	.000
	Task	2	11.402	12.829	<.0005*	.137
	Immersion x Task	2	1.071	1.205	.302	.015

Figure 53: ANOVA for Instructional Efficiency for Room Size (Exact)

A 2x3 (Immersion x Task) ANOVA was conducted using instructional efficiency for the Room Order (relative scoring) as a dependent variable. There was a main effect of Task (F=26.260, p<0.0005, η_p^2=0.245) and a significant main effect for Immersion (F=35.458, p<0.0005, η_p^2=0.180). No significant interaction of Immersion and Task conditions (F=1.205,

59

p=0.302) was observed. IE was higher in High Immersion conditions than in Low Immersion and post hoc testing showed that IE was greatest in ST conditions and lowest in NDT conditions.

Measure	Source	df	MS	F	p	η_p^2
Instructional Efficiency for Room Order (Relative)	Immersion	1	22.829	35.458	<.0005*	.180
	Task	2	16.907	26.260	<.0005*	.245
	Immersion x Task	2	1.520	2.361	.098	.028

Figure 54: ANOVA for Instructional Efficiency for Room Order (Relative)

A 2x3 (Immersion x Task) ANOVA was conducted using instructional efficiency for the Room Order (exact scoring) as a dependent variable. There was a main effect of Task (F=27.692, p<0.0005, η_p^2=0.255) and a significant main effect for Immersion (F=46.054, p<0.0005, η_p^2=0.221). No significant interaction of Immersion and Task conditions (F=1.563, p=0.213) was observed. IE was higher in High Immersion conditions than in Low Immersion and post hoc testing showed that IE was greatest in ST conditions and lowest in NDT conditions.

Measure	Source	df	MS	F	p	η_p^2
Instructional Efficiency for Room Order (Exact)	Immersion	1	29.797	46.054	<.0005*	.221
	Task	2	17.917	27.692	<.0005*	.255
	Immersion x Task	2	1.011	1.563	.213	.019

Figure 55: ANOVA for Instructional Efficiency for Room Order (Exact)

A 2x3 (Immersion x Task) ANOVA was conducted using instructional efficiency for the object portion of the free recall task as a dependent variable. There was a main effect of Task (F=36.637, p<0.0005, η_p^2=0.311) and a significant main effect for Immersion (F=4.523, p=0.035, η_p^2=0.027). No significant interaction of Immersion and Task conditions (F=2.354, p=0.098) was observed. IE was higher in High Immersion conditions than in Low Immersion and post hoc testing showed that IE was greatest in ST conditions and lowest in NDT conditions.

Measure	Source	df	MS	F	p	η_p^2
Instructional Efficiency for Object Recall (Correct)	Immersion	1	3.400	4.523	.035*	.027
	Task	2	27.542	36.637	<.0005*	.311
	Immersion x Task	2	1.769	2.354	.098	.028

Figure 56: ANOVA for Instructional Efficiency for Object Recall (Correct)

A 2x3 (Immersion x Task) ANOVA was conducted using instructional efficiency for correct responses on the object recognition task as a dependent variable. There was a main effect of Task (F=36.796, p<0.0005, η_p^2=0.312) and a significant main effect for Immersion (F=8.803, p=0.003, η_p^2=0.052). No significant interaction of Immersion and Task conditions (F=2.720, p=0.069) was observed. IE was higher in High Immersion conditions than in Low Immersion and post hoc testing showed that IE was greatest in ST conditions and lowest in NDT conditions.

Measure	Source	df	MS	F	p	η_p^2
Instructional Efficiency Object Recognition (Correct)	Immersion	1	7.443	8.803	.003*	.052
	Task	2	31.108	36.796	<.0005*	.312
	Immersion x Task	2	2.300	2.720	.069	.032

Figure 57: ANOVA for Instructional Efficiency for Object Recognition (Correct)

A 2x3 (Immersion x Task) ANOVA was conducted using instructional efficiency for distance score on the object location task as a dependent variable. There was a main effect of Task (F=7.857, p=0.001, η_p^2=0.088). However, there was no significant main effect for Immersion (F=0.103, p=0.748) and no significant interaction of Immersion and Task conditions (F=0.933, p=0.396). Post hoc testing showed that IE was greatest in ST conditions and lowest in NDT conditions.

Measure	Source	df	MS	F	p	η_p^2
Instructional Efficiency for Object Location (Distance Score)	Immersion	1	.072	.103	.748	.001
	Task	2	5.464	7.857	.001*	.088
	Immersion x Task	2	.649	.933	.396	.011

Figure 58: ANOVA for Instructional Efficiency for Object Location (Distance Score)

A 2x3 (Immersion x Task) ANOVA was conducted using instructional efficiency for the object portion of the free recall task as a dependent variable. There was a significant main effect of Task (F=30.052, p<0.0005, η_p^2=0.271) and a significant main effect for Immersion (F=7.774, p=0.006, η_p^2=0.046). A significant interaction of Immersion and Task conditions (F=4.355, p=0.014, η_p^2=0.051) was observed. IE was higher in High Immersion conditions than in Low Immersion and post hoc testing showed that IE was greatest in ST conditions and lowest in NDT conditions. The condition with the greatest instructional efficiency was the High Immersion, Single task condition.

Measure	Source	df	MS	F	p	η_p^2
Object Location (Correct Room)	Immersion	1	6.316	7.774	.006*	.046
	Task	2	24.415	30.052	<.0005*	.271
	Immersion x Task	2	3.538	4.355	.014*	.051

Figure 59: ANOVA for Instructional Efficiency for Object Location (Correct Room)

CHAPTER FOUR:
DISCUSSION

Manipulation Checks

The first manipulation check was conducted to determine that the Immersion manipulation increased subjective experience of presence, but that the Task manipulation did not. There was a main effect of Immersion on PQ scores where high immersion groups reported experiencing a greater degree of presence than low immersion groups. However, there was no main effect of Task or interaction term. It is evident that the Immersion manipulation was successful in increasing participants' experience of presence and, based on this, any performance difference between the two groups will be interpreted as a result of differences in presence.

The second manipulation check was conducted to determine that the Task manipulation increased cognitive load, but that the Immersion manipulation did not. There was a main effect of Task where participants reported higher levels of cognitive load in the two dual-task groups. However, there was no main effect of Immersion or interaction effect. Additionally, post-hoc testing showed that while the two dual-task groups experienced significantly greater cognitive load than the single-task groups, EDT groups did not differ significantly from NDT. It is evident that the Task manipulation was successful in increasing cognitive load. Additionally, because there was no significant difference in cognitive load between the two dual-task groups, any differences in performance can reasonably be attributed to differences in type of load (germane vs extraneous) rather than amount of load.

Neither manipulation check showed significant evidence that either of the two manipulations affected the target construct of the other. Immersion condition did not affect cognitive load, Task condition did not affect presence, and there was no significant interaction

term in either analysis. Based on this, it is possible to determine that both constructs were

manipulated independently as well as successfully, as was shown to be the case in Ma and

Kaber (2006).

Version and Group Checks

The t-tests conducted for comparing the two versions of the VE did not show any

significant difference in performance between versions on any outcome measure that

incorporates object memory. Further analysis on an item-by-basis did not show any significant

difference between any of the object-foil pairs. This demonstrates that the two versions were

roughly equivalent and that the items from one version may be used as controlled foils for

participants who experienced the other version. A full accounting of each object and its alternate

version can be seen in Appendix 2: Object Codebook.

Additionally, the one-way ANOVAs conducted for max span on both the spatial and

verbal span tasks did not show any significant difference between the six experimental group.

This means that any group differences that are observed are most likely to be a result of the

group conditions rather than any underlying difference in group cognitive ability.

Hypothesis Testing

Hypothesis 1 (Spatial Memory). Hypothesis 1 predicted that participants in High

Immersion conditions will show better memory for spatial measures. Analysis of Room Size

scores (both relative and exact scoring methods) showed a significant main effect of Immersion

condition where high immersion groups made more correct judgments than low immersion

groups. Analysis of Room Order scores (both relative and exact scoring methods) showed a

similar pattern, with a main effect of Immersion condition and with high immersion groups

making more correct judgments than low immersion groups. Based on these results and on the

successful demonstration that the Immersion manipulation increased subjective experience of

presence, it is reasonable to conclude that Hypothesis 1 is strongly supported by the observations of this study.

However, in addition to these supportive findings, these four analyses also provided some unexpected results. For both Room Size and Room Order tasks, using the relative scoring method resulted in a significant interaction of Task and Immersion manipulations. This interaction term is no present when using the exact scoring methods. These results suggest that there may be a difference in the underlying cognitive mechanisms that support performance as measured through a particular scoring method. Additionally, when comparing the main effect of the Immersion manipulation between the Room Size and Room Order task, it is clear that the effect size of this manipulation is considerably larger for Room Order than it is for Room Size.

Hypothesis 2 (Object Memory). Hypothesis 2 predicted participants in the Ecological Dual-Task group would perform best on measures of object memory while participants in the Non-Ecological Dual Task groups would perform worst. Analysis of participant performance on the Free Recall task showed a main effect of Task on the number of objects remembered with participants recalling the largest number of objects in the EDT groups and the smallest number of objects in the NDT groups. Likewise, analysis of participant performance on the Object Recognition task showed a significant main effect of Task with participants in EDT groups remembering the greatest number of objects and NDT groups remembering the fewest. Because our manipulation check showed that there was no significant difference in subjective cognitive load between EDT and NDT groups, it is reasonable to conclude that these differences in performance are driven by differences in germane vs extraneous cognitive load.

However, in addition to these supportive findings there were also several unexpected findings on these measures. First, neither analysis of falsely remembered objects (from the Free Recall task and the Object Recognition task) showed any significant main effect or interaction of

the manipulations used in this experiment. This may mean that presence and cognitive load primarily work on the encoding of things that are present but do not affect the formation of false memories. Alternately, this finding may simply be due to a floor effect, as the number of falsely remembered items on each of these measures was extremely low.

Analysis of the Object Recognition task also revealed unexpected effects. In addition to the predicted main effect for task, there was also a significant main effect of Immersion and an interaction of Immersion and Task. Post hoc testing showed that across conditions, EDT outperformed NDT and ST groups. However, it also showed that in the two dual-task conditions, participants in high immersion conditions remembered significantly fewer objects than participants in low immersion conditions. This means that when a secondary task was present, higher levels of immersion actually resulted in worse memory of objects.

Hypothesis 3 (Feature Binding). Hypothesis 3 made diverging predictions about feature binding measures. For the Recall task, which allowed for qualitative details of all types, it was predicted that participants in the EDT group would perform best while participants in the ND group would perform worst. However, for the Object Location task, it was predicted that the equal weighting of object and spatial components would result in performance driven by both Immersion and Task group, resulting in a strong interaction term.

Contrary to this hypothesis, analysis of details recalled per object on the Free Recall task did not show any significant main effects or interactions. Analysis of participant performance on the Object Location task showed somewhat mixed results. When analyzed based on raw error distance (distance score), there was a main effect of Task condition but no other main effect or interaction. Post hoc testing showed that participants in EDT groups provided object locations which were significantly closer to the objects' actual location than

66

participants in other groups. Neither of these analyses aligned with the predictions made by Hypothesis 3.

However, during the course of the experiment, participants provided qualitative feedback indicating that some portion of them may have provided responses to the Object Location task that indicated the room that the object was located rather than the exact location. On this basis, we also analyzed the Object Location data based on the number of cases where the participant reported a location that was in the correct room. When analyzed based on number of correct room judgments, there was a main effect of both Immersion and Task conditions as well as a strong significant interaction term. This interaction term, in fact, showed a greater effect size than either of the two main effects. Post-hoc comparisons showed that participants performed best in high immersion and EDT conditions, with the HI-EDT group performing best and the LI-NDT group performing worst. These results align precisely with the specific predictions made in Hypothesis 3(b). However, overall the evidence supporting Hypothesis 3 remains somewhat weak.

Exploratory Measures

Hypothesis 4 (Gender Differences). Hypothesis 4 outlined the expected gender differences in performance for each component of episodic memory investigated in this experiment. For spatial memory (H4a), male participants were expected to show better performance than female participants. Analysis of performance on the Room Size and Room Order task showed that there was a significant main effect of Gender for relative scoring methods but not for exact scoring methods. In both cases, male participants provided significantly more correct judgments of size and order than did female participants. These findings provide some support for H4a.

However, there was an unexpected additional finding present in these analyses. When performance on the Room Order task was analyzed based on performance using the exact scoring method, there was no main effect of gender. However, there was a significant interaction of Gender and Task and a very strong three-way interaction of Gender, Immersion, and Task. In this interaction, participants in the high immersion groups performed homogenously well. In low immersion conditions, Task group affected performance differentially based on gender. For male participants, the worst performance was seen in the single-task group and the best performance in the EDT group. However, for female participants, the best performance was seen in the single task group and the worst performance seen on the NDT group.

For object memory (H4b), female participants were expected to show better performance than male participants. Analysis of performance on object memory measures revealed no main effect of Gender on either Object Recall or Object Recognition tasks. However, there was a small but significant interaction of Gender and Task for falsely recalled objects on the object recall task. Post hoc testing showed that this interaction existed primarily within the EDT group, where female participants provided significantly more false recalls than did male participants. However, despite this interaction, H4b remains mostly unsupported.

For feature binding (H4c), there was expected to be no significant difference in performance between male and female participants. For analyses based on both details per object from the Free Recall task and correct room judgements on the Object Location task, there were no significant main effects of Gender or Gender interactions. These two null findings provide some support for H4c, though caution should always be taken when inferring support based on null findings.

However, in addition to these supportive findings, there was one unexpected result that contradicts H4c. When analyzed using distance score on the Object Location task, there was a main effect of gender where male participants provided significantly closer object locations than did female participants. While it is tempting to dismiss this finding on the basis that many participants reported providing correct room judgments rather than exact location judgments, this gender difference shows the largest effect size of any found in this experiment ($\eta_p^2=0.078$).

Hypothesis 5 (Instructional Efficiency). Hypothesis 5 made predictions about the relative levels of instructional efficiency for each condition based on outcome measures. These hypotheses roughly corresponded to the predictions made in Hypotheses 1-3. For spatial measures (5a) it was predicted that IE would be higher in High Immersion groups than Low Immersion groups. For object measures (5b) it was predicted that IE would be higher for groups assigned to perform the secondary task, and for feature binding (5c) it was expected that there would be an Immersion by Task interaction where the LI-ST condition would have the lowest IE and the HI-EDT condition would result in the highest IE.

For every outcome measure tested in this experiment, single task conditions had greater IE than either dual-task condition. With regard to spatial measures, Immersion condition did not have any significant effect on IE for the Room Size measure. High Immersion conditions did show greater IE on the Room Order task, but only the single task condition had a positive value for IE. However, it is important to note that post-hoc testing showed a significant difference between the IE of ST and NDT groups, but did not show a significant difference between the IE of ST and EDT groups. This means that the EDT groups performed significantly better and did not lose a significant degree of instructional efficiency.

Limitations

When drawing conclusions based on this experiment, it is important to note several limitations in the design and measures used. First, it is important to note that the two major constructs measured here cannot be manipulated directly. This may be especially problematic in terms of our manipulation virtual presence. While the manipulation checks demonstrated that increased levels of technological immersion did increase subjective levels of presence there is no way in the present study to demonstrate that the increases in presence were the cause of any differences in performance. For example, it might also be valid to argue that improvements in one's memory of the spatial aspects of a virtual environment led participants to subjectively report higher levels of virtual presence rather than the reverse.

It is also important to note that, while significant, the manipulation of presence may have been weaker than other studies. This study used a head-mounted display as the primary immersive technology but did not supplement this with immersive motion controls, standing VR or Room Scale VR. The addition of even more immersive technology may have demonstrated either enhanced effects of additional levels of presence or diminishing returns. Additionally, PQ scores demonstrated that even participants assigned to the Low Immersion condition rated their subject experience of presence near the highest point of the scale. It is therefore important to remember that the results of this study represent a small slice of the experience of presence, and that the addition of even less immersive conditions may have revealed additional important data.

Further complicating this limitation is the muddy nature of immersion and presence as represented in the literature. As noted across many sources (for an extensive review, see Procci, 2015), there is comparatively little consensus among researchers on what constitutes immersion and what constitutes presence. This study chose to focus on the dichotomy

proposed by Slater (2003) because this relationship between objective technological aspects (immersion) and subjective experience (presence) most closely matched the existing evidence for the given problem space. However, researchers using different definitions of immersion and presence (and using different measures for these constructs) may come to different conclusions even from the same set of results.

Another potential limitation on this experiment stems from the order that the measures were administered. As noted in the Methods section, it was necessary to present Room Size and Room Order tasks prior to the administration of the Recognition and Object Location task because the latter task presented participants with a correct allocentric representation of the environment. However, this order of presentation may have influenced the way that participants chose to answer the object location task by priming them to process the environment in terms of room contents rather than in pure spatial terms. While results were processed both in exact object location terms and on a correct room basis, there is no way to know for sure the level to which the order may have influenced participants. Future experiments may choose to present only the Object Location task to determine if this may be the case.

The lack of notable gender differences on spatial and verbal span tasks suggests a further limitation on the results of this study. For these tasks, order may have also played a role as the span tasks were presented at the end of the session after the participants had already completed the virtual task and a lengthy questionnaire. Participants may therefore have been fatigued and not performed to their maximum capacity. Additionally, there has been some nascent research that suggests that gender differences in route and survey spatial representations may be at their greatest in outdoor and natural environments. Additionally, some types of gender differences in spatial memory can be mitigated by the provision of external verbal information, such as landmarks (Crook, Youngjohn, & Larrabee, 1993). As the

71

virtual environment in this study was set inside a hospital environment, the indoor nature of the

task may have suppressed potential differences between male and female participants.

Exploratory findings of gender differences should likewise be taken with some caution.

There is growing cultural awareness of a differentiation of the terms "gender" and "sex" that has

also begun to shape the way that scientific study is conducted (Oertelt-Prigione & Regritz-

Zagrosek, 2011). "Sex" now typically refers to the biological sex of the individual and is

generally associated with some biological difference (such as chromosomes, gonads, or

hormones). "Gender" now typically refers to psychosocial differences between men and women,

generally thought to be largely socially and culturally constructed. While this dissertation broadly

speaks in terms of "gender differences," the literature that it draws upon to craft hypotheses and

conclusions broadly refers to sex differences. While it may be possible to determine whether a

particular study (for example, Lewin, Wolgers, & Herlitz, 2001) asked its participants for sex or

for gender, there is no way of determining whether those participants were adequately capable

of differentiating the two. For this dissertation, participants were explicitly asked to provide their

gender and a further exploratory question was used to determine whether the participant

considered themselves transgender. All participants in this study provided a male or female

response to the gender question and no participants indicated that they were transgender. This

is the basis on which all gender difference analyses were made, and it is therefore possible that

a different approach may yield different results. For example, it is possible that if gender and sex

(assigned at birth) were both collected and analyzed, each might show a different pattern of

interaction.

Finally, the sample used in this study was composed entirely of young, college educated

adults in the United States. While this decision was made to provide a baseline for the effects of

presence and cognitive load on a group of healthy participants, there may be some issues with

generalization of these results to other populations. For example, a great deal of the foundational evidence used in the formulation of these theories was synthesized across a body of research that included healthy older adults, young adults with neurological complaint, and older adults with dementia. These studies have shown that certain manipulations (e.g. active vs passive control) may have differential effects that depend on the age of the participant, their cognitive status, and their experience with technology (Plancher & Piolino, 2017). Additionally, a great deal of this foundational research was conducted using outdoor environments and with European populations (Barclay, Parker, & Sims, 2018). Very little research has been done distinguishing the effects of native language on episodic memory, nor is there a great deal of research directly comparing episodic representation of outdoor vs indoor environments. Either of these two factors may play an unexplored role in the reaction of our participants to the manipulations of this study.

Conclusions

Theoretical Considerations. The combination of Presence theory and Cognitive Load theory within a single experiment provides an interesting basis for discussion about the two theories. While each theory is concerned with performance, they also each approach performance from a differing perspective that may lead to a problematic understanding of the results of this experiment. On this basis, it is important that we take a step back and consider some of the fundamental assumptions of each theory as it relates to this experiment.

While Presence theory predicts that task performance on virtual tasks will increase as participants experience greater levels of presence, the underlying principle of the theory can be characterized as a subtractive theory rather than an additive one. Presence is usually described as a feeling of being physically situated in a virtual environment and accepting that virtual environment as one's primary operation environment. This conception implies that a non-virtual

73

environment, where one is physically situated and operating, is strictly superior to a virtual one on the basis that bridging that abstraction is not necessary. Such an argument would be supported by a good deal of research that indicates that more immersive technologies tend to be those that make the virtual task more similar to the real world, such as the mounting evidence that participants employ similar patterns in visual search in a virtual environment (whether by mouse, joystick, or HMD) as they would in a natural environment (cf. Kit et al, 2014).

The experiment presented here manipulated presence using the Oculus head-mounted display, that allowed users to use their head, neck and eyes to deploy these same visual search patterns. When provided this ability, participants' subjective ratings of presence increased, as did their performance on measures of spatial memory. However, this enhancement largely did not extend to measures of object memory, perhaps indicating that presence is a construct that relates better to a person's experience of space than it does to generalized performance. This line of inquiry bears further research, as it could have strong implications for how the results of virtual tasks (such as the virtual route learning task, used in the early identification of Alzheimer's disease) can be applied to real life.

Cognitive Load theory, on the other hand, is much more concerned with the distillation of learning into its intrinsic elements. Rather than adding new elements to a task to increase the level to which it matches reality, as is the aim of Presence theory, Cognitive Load theory would suggest that the best approach is to remove extraneous elements and leave only that which is helpful. In this study, it seems that this approach is moderately effective. While the introduction of new elements, such as secondary ecological tasks, were shown to sometimes enhance performance at a raw level, measures of instructional efficiency indicated that for all but a few cases the increases in cognitive load outweighed the performance gains.

However, the relationship of Cognitive Load theory to the current drive for ecological validity seems to be at odds. Cognitive Load theory would suggest that we strip away the unnecessary elements of the memory task. Ecological validity advocates, on the other hand, would argue that stripping away unnecessary elements runs counter to the goal of producing test environments that reflect real-life performance. Parsons and McMahon (2015), for example, argue that in real-life environments people very rarely perform only one task at a time and that even when they do, the real world is filled with distractors. Perhaps when it comes to real-world memory performance and ecological validity, it is the case that the appropriate application of Cognitive Load theory is to systematically account for these distractors and multitask situations. Moreover, the results of this experiment suggest that cognitive load may only be influential in tasks where there is some level of object memory being measured. This means that such an approach may not even be appropriate for situations where researchers are purely concerned with spatial aspects of memory.

Unexpected Results. Several of the results found in this study were not predicted by the hypotheses presented at the start of this dissertation. While most of these results do not run counter to these hypotheses, accounting for them may provide greater depth to the observations presented here. First, when a relative scoring method was used, there was a significant interaction of Immersion and Task on both measures of spatial memory. Second, the main effect of Immersion had a much larger effect size on our measure of route memory than it did on our measure of survey memory. Third, there was a main effect of Immersion and an interaction effect on the Object Recognition task but not the Free Recall task. Fourth, there was a very strong three-way interaction between Gender, Immersion, and Task on our measure of route learning.

There are two potential explanations for these observations. The first is the simplest, that both the use of more immersive technologies and the implementation of germane cognitive load are ways of directing the participants' attention. This explanation follows from the observations of Wilson and Peruch (2002), in which participants (both active and passive) were given instructions to either attend to spatial layout or target objects. Those who were directed to attend to spatial aspects performed better on spatial measures and those directed to attend to objects performed better on measures of object memory, with no detectable difference between active and passive participants. On this basis, it is possible to conclude that our presence and cognitive load manipulations served to direct participants attention to particular aspects of the environment. For example, the addition of a head-mounted display may have enhanced the level to which the participants' attention was directed to spatial aspects of the environment as they were necessary to navigate. This would explain the difference in effect size between survey and route measures, as the first-person perspective of the VE combined with the immersion of the HMD would have served to strongly direct participants' attention in a manner that is more conducive to an egocentric frame.

This explanation, however, does not explain the anomalous findings observed in the present experiment. For example, while the framing of the ecological secondary task could arguably direct participants to encode objects that are related to the medical context of the task it does not explain why it would also cause participants to perform better on a spatial measure. Nor does the supposition that a head-mounted display directs attention to spatial layout explain why those participants would later show enhanced performance on an object recognition task. Therefore, while this explanation is elegant, it may also be drastically oversimplified.

The second explanation requires the consideration of a body of work centered on coordinate and categorical representations (Kosslyn, Chabris, Marsolek & Koenig, 1992). In this

representation, coordinate tasks are fine-grained and exact tasks that typically call upon what we think of as spatial skills where categorical tasks are less exact, and only require information about the relationship of elements involved with the task. While this may seem like the traditional distinction between spatial and verbal tasks, there are many tasks that are considered spatial tasks that, by this framework, rely more heavily on verbal skills/strategies than on visuospatial skills. For example, consider a task that shows participants a dot on a computer screen and subsequently asks them to recall the location of the dot. If this task were to ask the participant to respond with which quadrant the dot was located, rather than an exact location, it would be considered a categorical spatial task (Landau & Jackendoff, 1993).

On this basis, it is likely the case that many of the tasks presented in this experiment represent either some mixture of coordinate and categorical processing, or at the very least, that they are able to be completed using either strategy. For example, our experiment found that Task condition had a significant main effect on memory for Room Order, but only when that task is scored based on *relative* room position. When scored based on *exact* room position, the effect of Immersion condition remains but the effect of Task condition disappears. A similar case can be made for the Recognition task, which showed a main effect of both Immersion and Task conditions. In this case, participants would primarily rely on categorical skills (i.e. was *object* in *environment*). However, as each object image was presented from an overhead, isometric angle, participants were likely to need to employ some degree of mental rotation, a coordinate spatial skill.

Most importantly, this conception helps to better explain the strong three-way interaction of Gender, Immersion, and Task observed on the Room Order task. Previous research on gender differences in episodic memory have shown that female participants are more likely to rely on verbal (categorical) strategies where male participants are more likely to rely on

77

visuospatial (coordinate) strategies (Herlitz and Rehnman, 2008). By this logic it can be argued that the high immersion condition directed all participants' attention to visuospatial strategies, leading to generally higher levels of performance. However, in low immersion conditions, the secondary task may have hindered female participants' performance by directly taxing their verbal skills. Male participants, on the other hand, did not see this effect because they were more likely to rely upon visuospatial skills instead. This explanation holds an additional level of weight when considering that the secondary task (both EDT and NDT) required participants to make categorical judgments (is X greater than Y?).

Based on this interpretation, initial theoretical interpretations would need to be revised considering the observations of this study. Where previously the evidence suggested that Presence and Cognitive Load altered performance based on the traditional separation of Spatial Memory and Object memory, respectively, the evidence in this study leads us to revise these expectations. Instead, we would suggest that Presence may be a larger factor in coordinate tasks and those tasks that require visuospatial strategies, where Cognitive Load may be a larger factor in categorical tasks and tasks that require verbal strategies. In tasks that require both types of processing, there is likely to be a degree of interaction between Presence and Cognitive Load that will vary depending on the level to which participants utilize each strategy.

Future Studies

The findings and explanations presented here lead to a new set of hypotheses that can be tested, each with interesting implications of their own. For example, while the traditional division of Spatial vs Object memory requires a focus on semi-separable streams of processing, a division into visuospatial and verbal strategies may require a focus on cerebral lateralization as visuospatial and coordinate tasks are generally believed to be lateralized to the right hemisphere where verbal and categorical tasks are believed to be lateralized to the left

hemisphere (Vogel, Bowers, & Vogel, 2003). If this is the case, then it is possible that issues that require lateralization measures (for example, focal epilepsy) may be differentially affected by the addition of a secondary task.

Additionally, the results of this experiment can be beneficial to the development of population norms for cognitive testing. As virtual environments become a more and more popular means of testing for everyday function, it is important to have a robust knowledge of what performance normally looks like in both impaired and unimpaired populations. However, due in part to a lack in standardization in the ways that these virtual environments are developed, built, and presented, there is a vast amount of data that cannot necessarily be directly compared. This study indicates that, even for a relatively homogenous population, performance on a variety of outcomes is affected by both technological factors and task factors. Future studies may be necessary for clinicians to be able to be confident in the level to which they can declare that impairments or improvements that their patient is showing are due to actual cognitive changes rather than differences in testing mechanisms.

Future studies on the application of Presence and Cognitive Load theories on episodic memory of virtual environments should expand upon the groundwork laid here. As this study only tested these effects using young adults, future studies might explore similar paradigms with older adults and patient populations to determine the extent to which the effect observed here generalizes across the lifespan and how these factors might influence the deficits experienced by patients with various forms of memory complaint. Additionally, the combination of improved performance and improved instructional efficiency for route learning in High Immersion conditions suggests that the introduction of more immersive technologies may provide a means to improve learners' route learning with little downside.

Future research along this path might also seek to devise new strategies for manipulation of germane vs extraneous workload. While our task manipulation did result in enhanced performance of ecological dual task participants over other groups, the analysis of instructional efficiency suggests that there were elements of both germane and non-germane load. Therefore, future experiments may seek to develop secondary tasks that more strongly support the primary task in order to increase the ratio of germane/extraneous load. For example, one might imagine developing a modified n-back task that asks participants about previously seen objects as they progress through the environment. Further refinement of this kind of methodology will enable us to more definitively differentiate between an underload-load-overload model of cognitive load and a conception of cognitive load that compartmentalizes between source and type of load. This question is especially salient when considering the arguments made by researchers such as Parsons and McMahan (2017) that distraction and multitasking are key parts of everyday functioning.

The goal of this dissertation was to investigate the role of presence and cognitive load in virtual tasks that purport to be ecological tests of episodic memory. By manipulating and measuring each construct, this study demonstrated that technological and task factors play a role in episodic memory performance, potentially limiting the level to which results on these tasks can be generalized across the category. Additionally, by measuring multiple components of episodic memory using different canonical approaches, we were able to show that even tasks that have traditionally purported to measure the same component of memory may be differentially affected by task and technological factors. Finally, in synthesizing the results of this study as a whole, several new and potentially impactful lines or research become apparent that may help to drive this burgeoning category of research in the future.

APPENDIX A:
TASK SCRIPTS

[Group 1]

You have been assigned to a group that completes the virtual task using a standard computer monitor. Once the task begins, you will have 10 minutes to explore the environment before the task ends. During this time, you will wear an audio headset. Through this headset, you will hear numbers. These numbers are used for a secondary task that you do not need to complete, so feel free to ignore them. Are you ready to begin?

[Group 2]

You have been assigned to a group that completes the virtual task using a standard computer monitor. Once the task begins, you will have 10 minutes to explore the environment before the task ends. During this time, you will wear an audio headset that you will use to complete a secondary task at the same time as you explore. You will hear number pairs, for example "one hundred and twenty over eighty." **These numbers are blood pressure readings, and your job is to monitor them.** When either the first number is over 160 OR the second number is over 140, you should respond by pressing the spacebar. Can you repeat that back to me? [Pause to let the participant respond]. Are you ready to begin?

[Group 3]

You have been assigned to a group that completes the virtual task using a standard computer monitor. Once the task begins, you will have 10 minutes to explore the environment before the task ends. During this time, you will wear an audio headset that you will use to complete a secondary task at the same time as you explore. You will hear number pairs, for example "one hundred and twenty over eighty." **Your job is to monitor these numbers.** When either the first number is over 160 OR the second number is over 140, you should respond by pressing the spacebar. Can you repeat that back to me? [Pause to let the participant respond]. Are you ready to begin?

[Group 4]

You have been assigned to a group that completes the virtual task using the Oculus head-mounted display. Once the task begins, you will have 10 minutes to explore the environment before the task ends. During this time, you will also wear an audio headset. Through this headset, you will hear numbers. These numbers are used for a secondary task that you do not need to complete, so feel free to ignore them. Are you ready to begin?

[Group 5]

You have been assigned to a group that completes the virtual task using the Oculus head-mounted display. Once the task begins, you will have 10 minutes to explore the environment

before the task ends. During this time, you will wear an audio headset that you will use to complete a secondary task at the same time as you explore. You will hear number pairs, for example "one hundred and twenty over eighty." **These numbers are blood pressure readings, and your job is to monitor them.** When either the first number is over 160 OR the second number is over 140, you should respond by pressing the spacebar. Can you repeat that back to me? [Pause to let the participant respond]. Are you ready to begin?

[Group 6]

You have been assigned to a group that completes the virtual task using a standard computer monitor. Once the task begins, you will have 10 minutes to explore the environment before the task ends. During this time, you will wear an audio headset that you will use to complete a secondary task at the same time as you explore. You will hear number pairs, for example "one hundred and twenty over eighty." **Your job is to monitor these numbers.** When either the first number is over 160 OR the second number is over 140, you should respond by pressing the spacebar. Can you repeat that back to me? [Pause to let the participant respond]. Are you ready to begin?

End of Session

Thank you again for participating. Please take your copy of the informed consent document. You should receive your SONA credit for this study within 24-48 hours. If you do not, please email the primary investigator, Paul Barclay, at the email listed at the bottom of first page of the informed consent. Have a great semester!

APPENDIX B:
OBJECT CODEBOOK
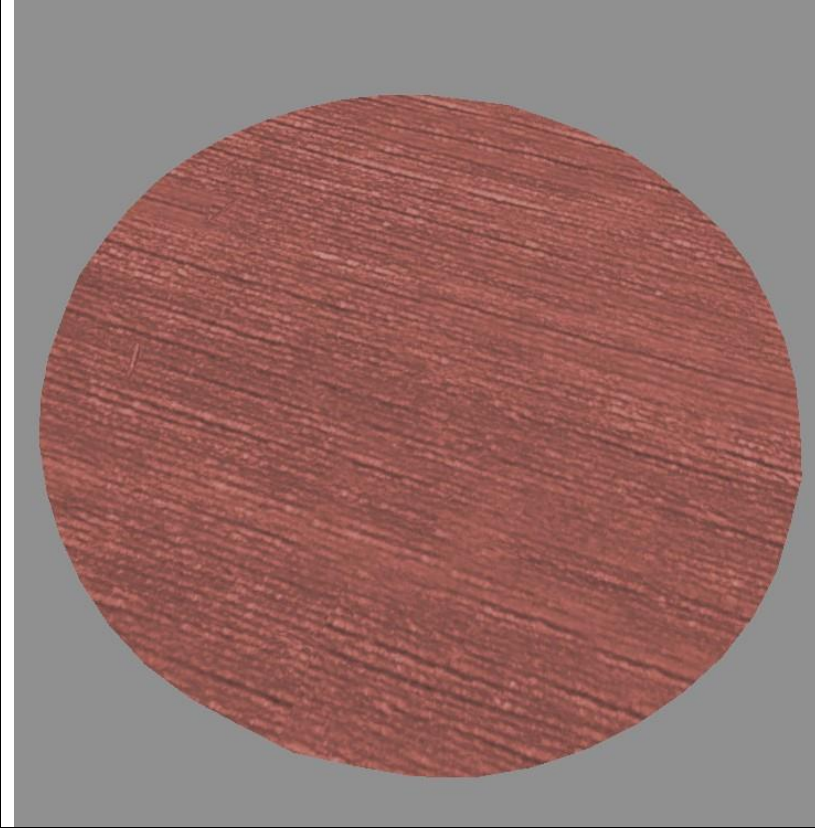
Object 1

| A | B |

Object 2

| A | B |

Object 3

| A | B |

Object 4

| A | B |

Object 5

| A | B |

Object 6

| A | B |

Object 7

| A | B |

Object 8

A	B

Object 9
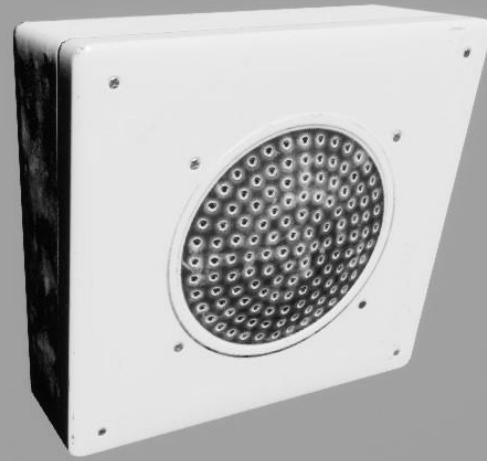
| A | B |

Object 10

| A | B |

Object 11

| A | B |

95

Object 12

| A | B |

Object 13

| A | B |

97

Object 14

| A | B |

98

Object 15

| A | B |

99
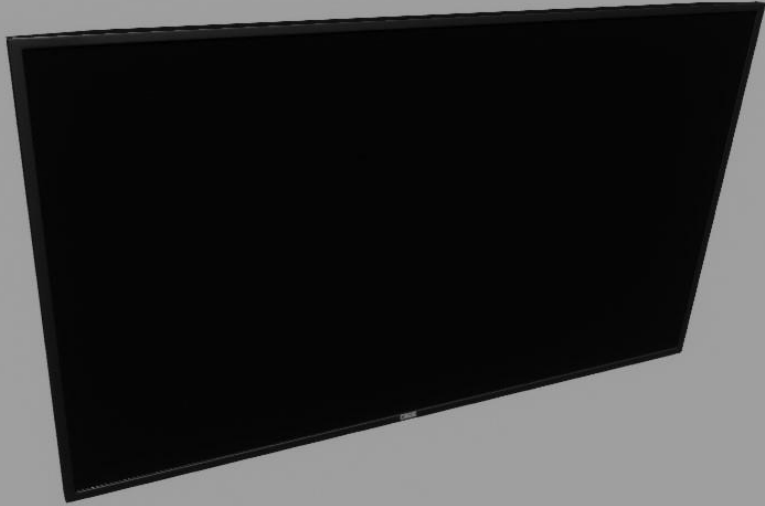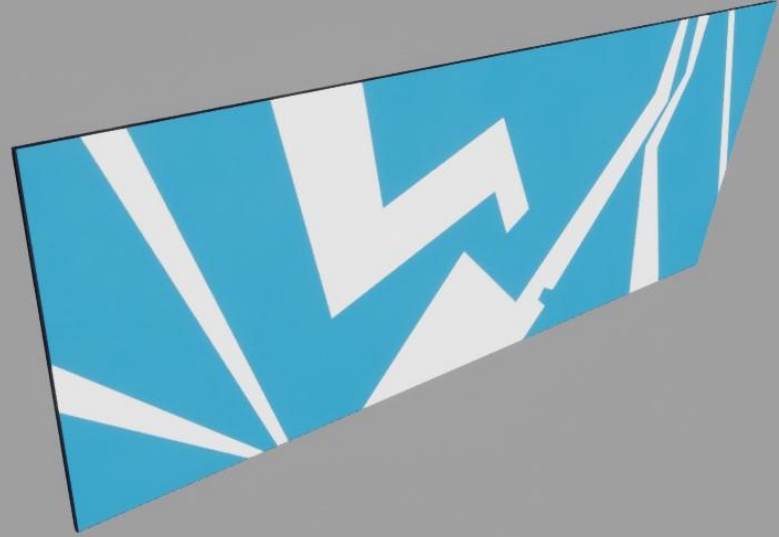
Object 16

| A | B |

Object 17

| A | B |

101
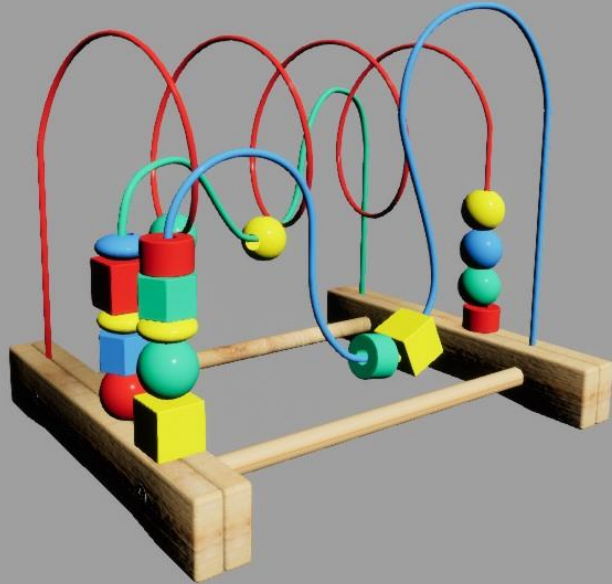
Object 18

| A | B |

Object 19

A	B

Object 20

| A | B |

Object 21

| A | B |

105

Object 22

A	B

Object 23

A	B

Object 24

A

B

108

Object 25

| A | B |

Object 26

| A | B |

Object 27

A	B

Object 28

A

B

Object 29

| A | B |

113

Object 30

| A | B |

Object 31

| A | B |

115

Object 32

| A | B |

116
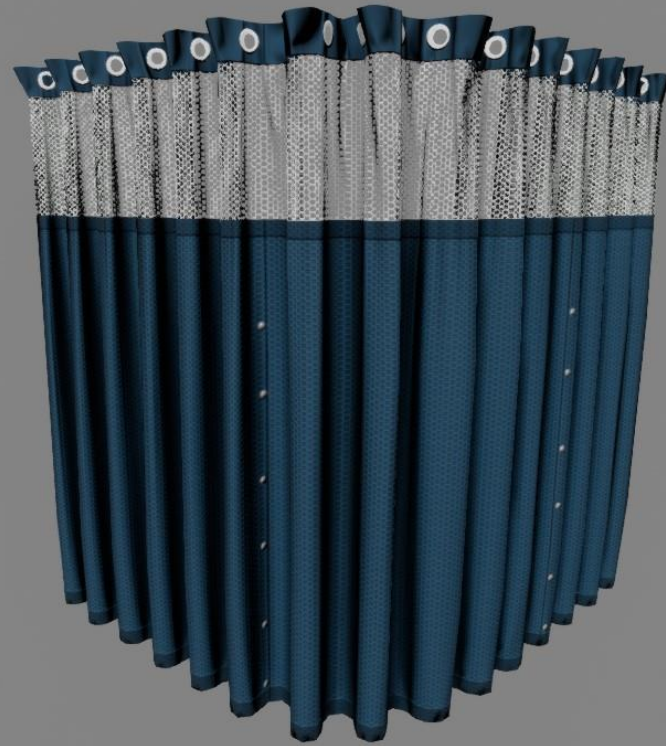
Object 33

| A | B |

Object 34

A	B

Object 35

| A | B |

119
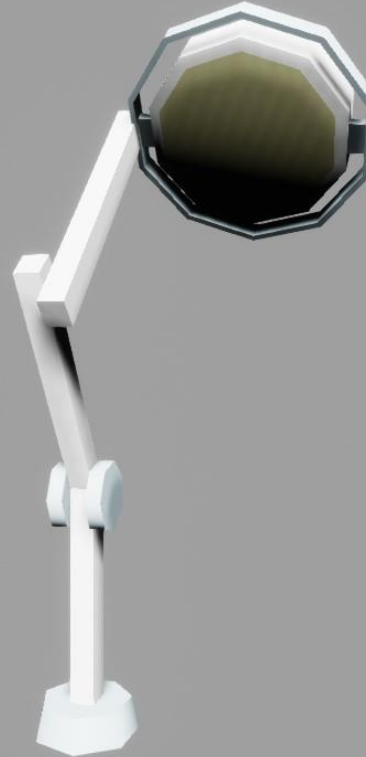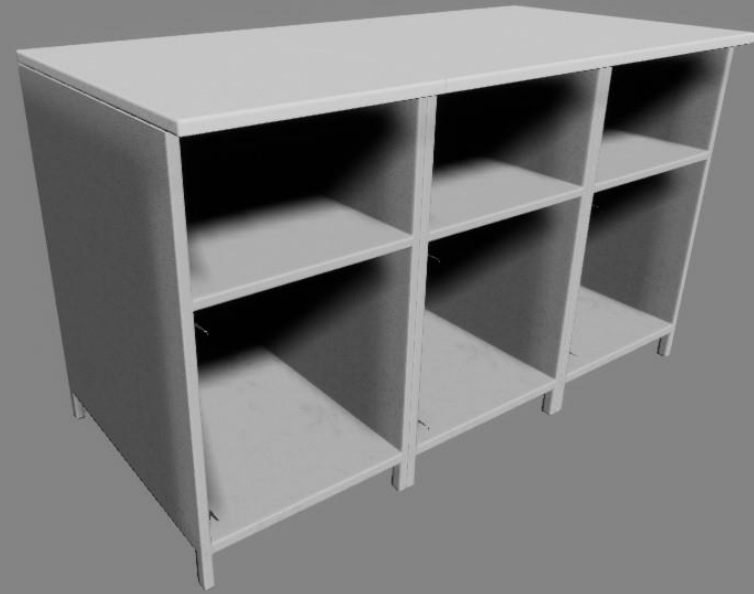
Object 36

| A | B |

Object 37

| A | B |

Object 38

| A | B |

122

Object 39

A

B

Object 40

| A | B |

124

Object 41

A

B

Object 42

A

B

Object 43

A

B

Object 44

A
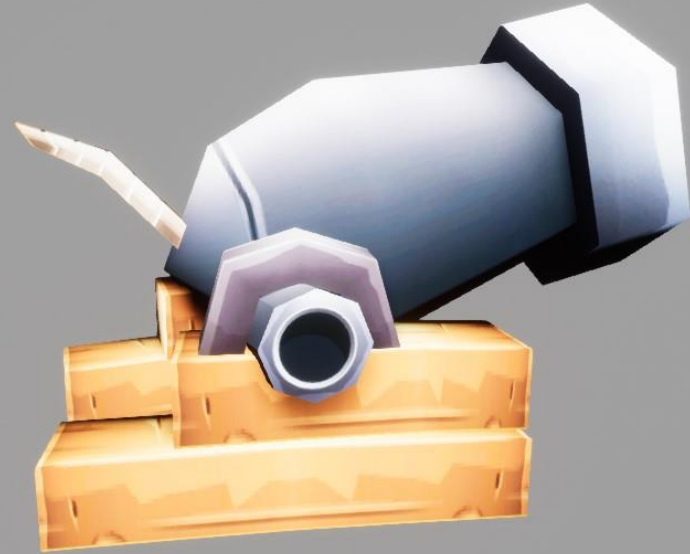
B

Object 45

A

B

Object 46

| A | B |

Object 47

| A | B |

Object 48

| A | B |

Object 49

A

B

Object 50

A	B

Object 51

| A | B |

Object 52

| A | B |

Object 53

A

B

Object 54

| A | B |

138
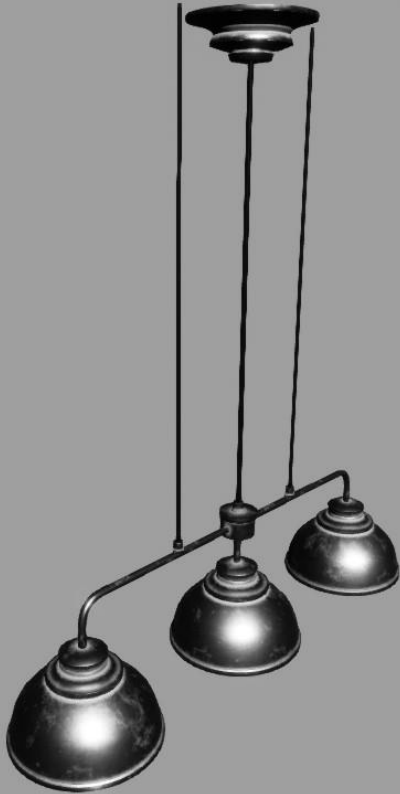
Object 55

A
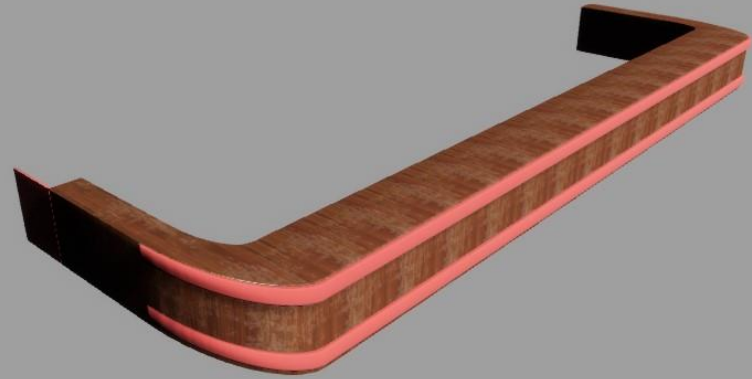
B

Object 56
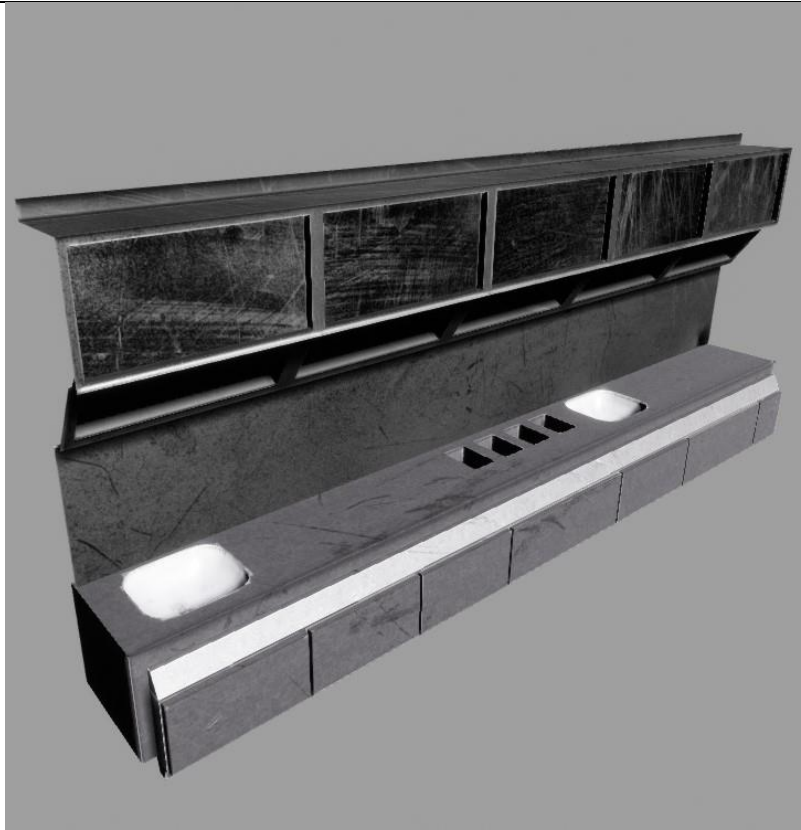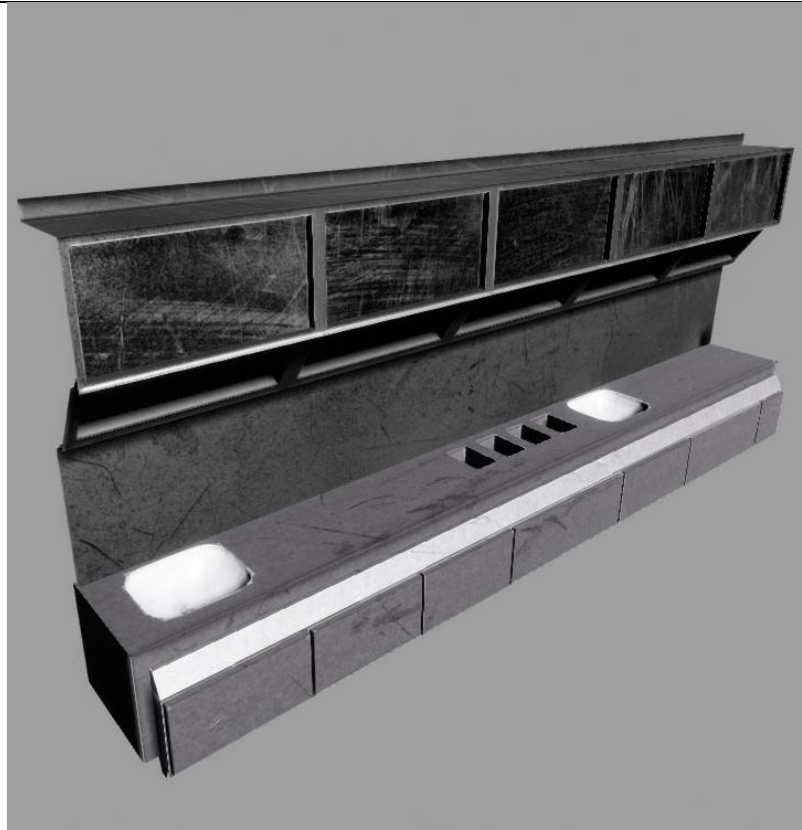
| A | B |

Object 57
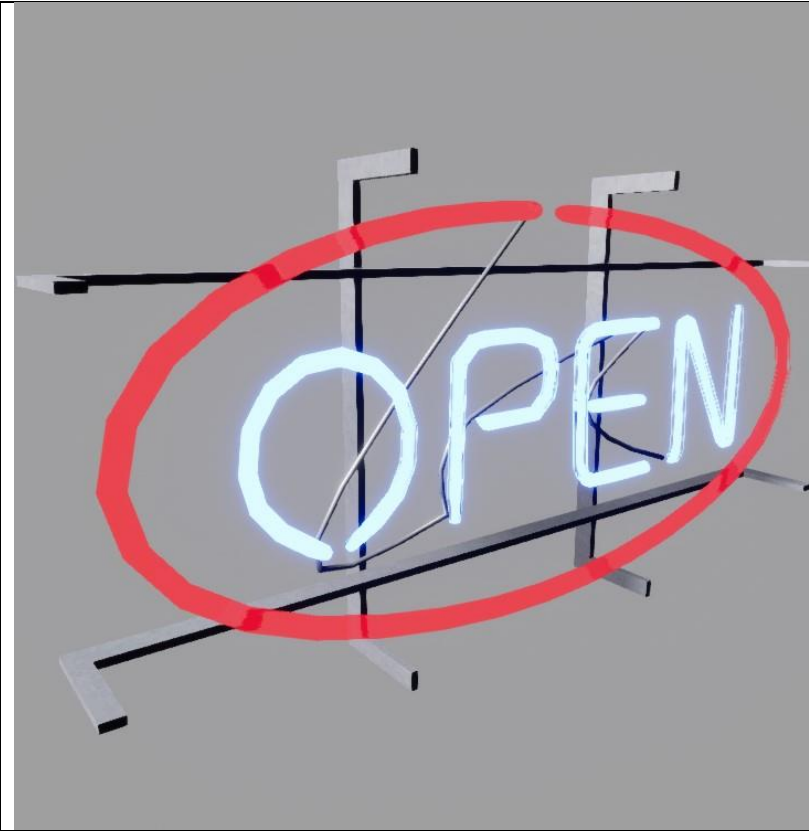
| A | B |

Object 58

A	B

Object 59

| A | B |

143

Object 60

| A | B |

Object 61
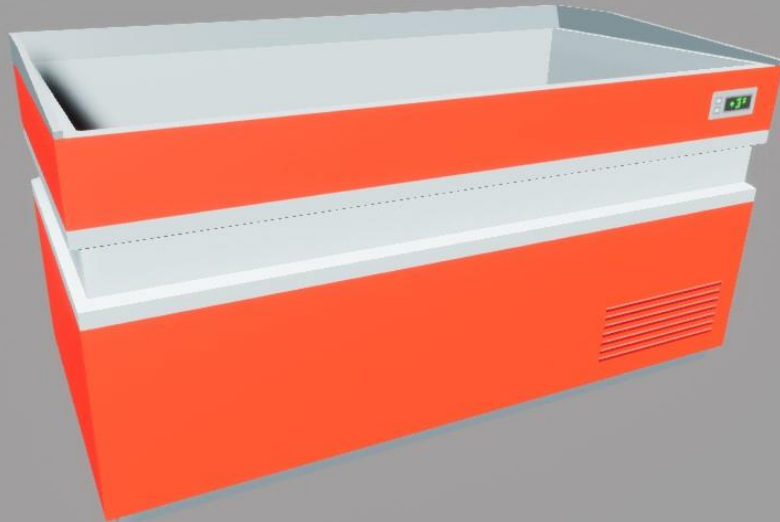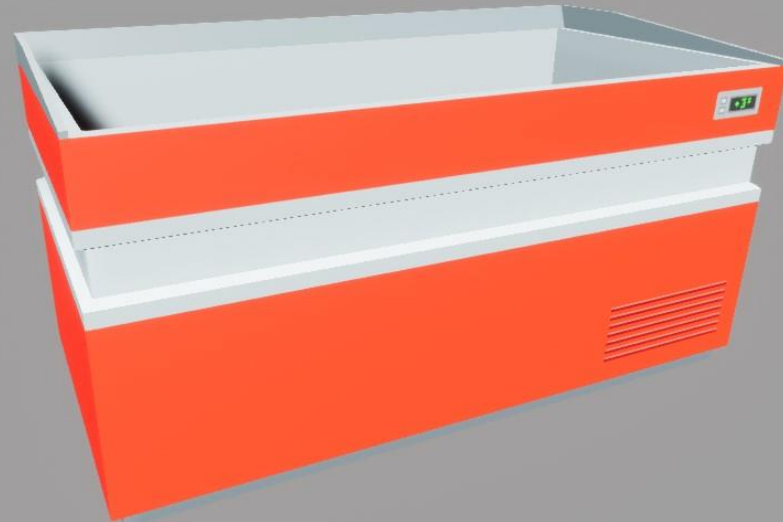
| A | B |

Object 62

| A | B |

Object 63

| A | B |

Object 64

| A | B |

Object 65

| A | B |

149
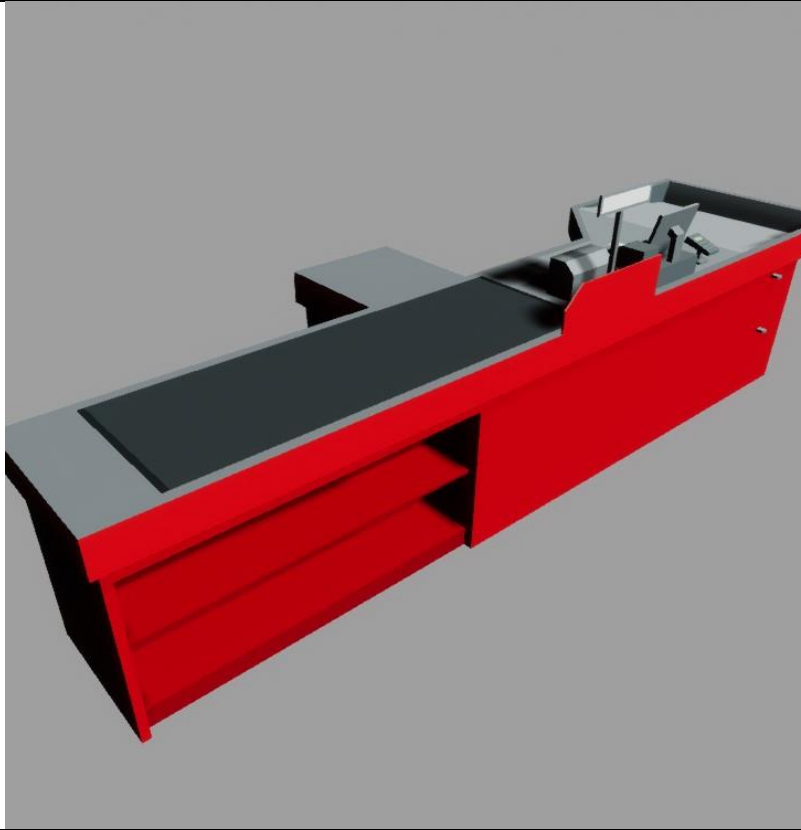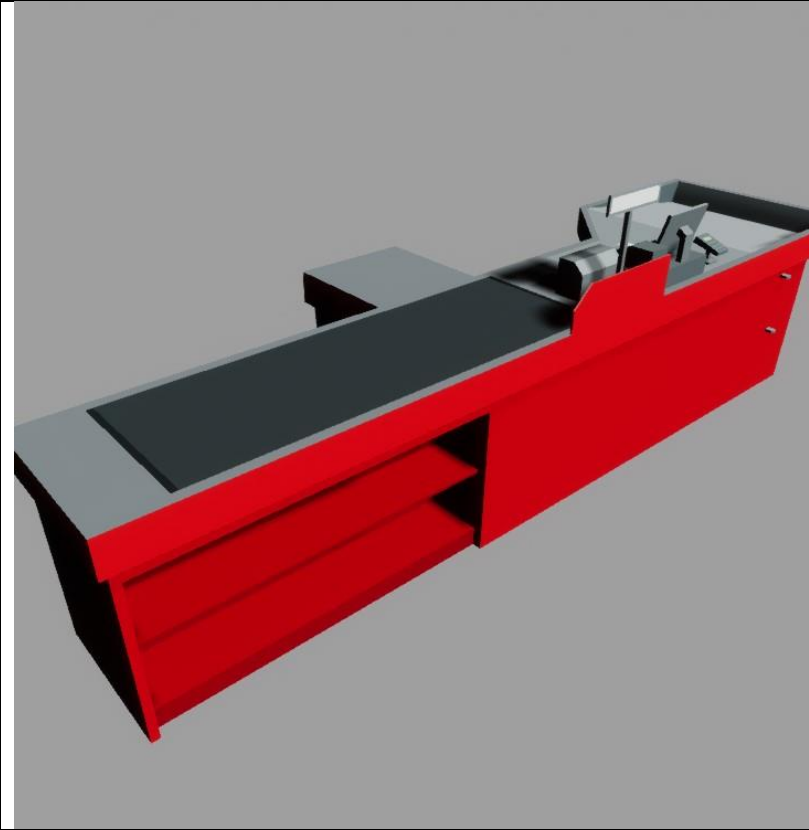
Object 66

| A | B |

Object 67

| A | B |

Object 68

A

B

Object 69

| A | B |

153

Object 70

A

B

Object 71

| A | B |

Object 72

| A | B |

Object 73

A

B

Object 74

| A | B |

158
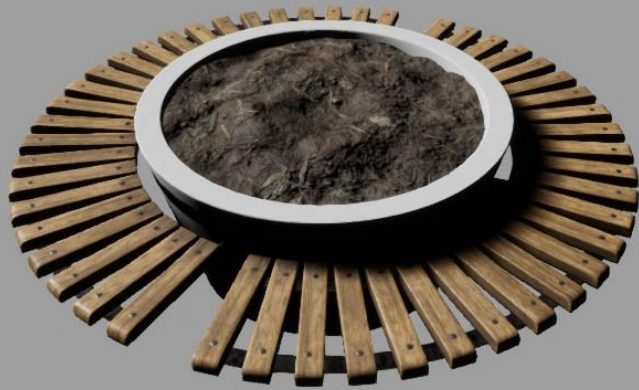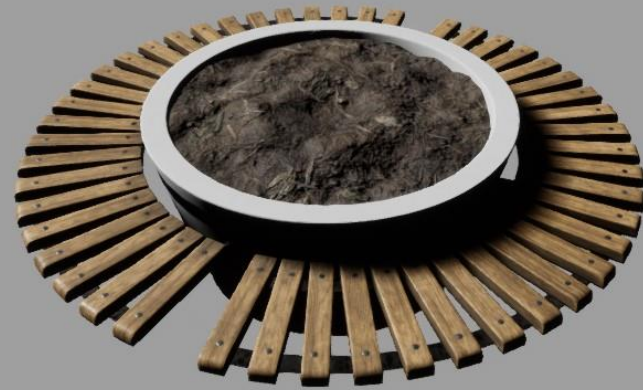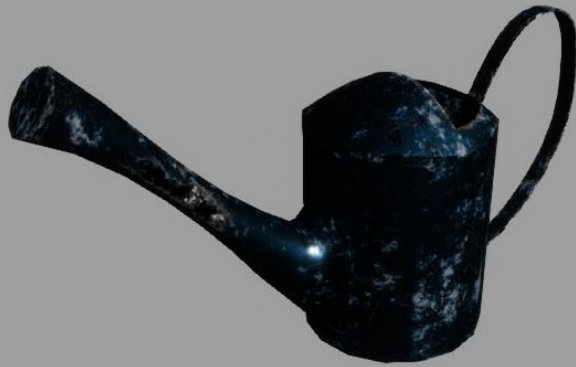
Object 75

A

B

Object 76

| A | B |

Object 77

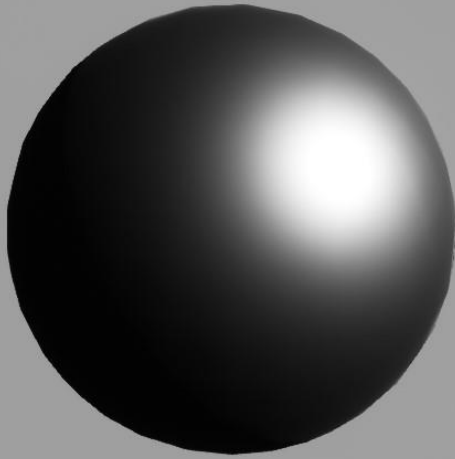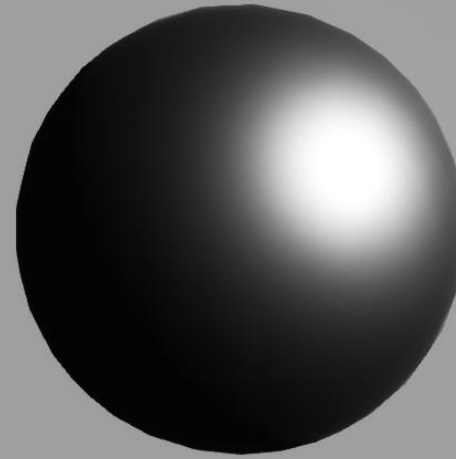| A | B |

161

Object 78

A

B

Object 79

A

B

Object 80

| A | B |

164

APPENDIX C:
IRB APPROVAL LETTER

University of Central Florida Institutional Review Board
Office of Research & Commercialization
12201 Research Parkway, Suite 501
Orlando, Florida 32826-3246
Telephone: 407-823-2901 or 407-882-2276
www.research.ucf.edu/compliance/irb.html

Approval of Human Research

From: **UCF Institutional Review Board #1**
FWA00000351, IRB00001138

To: **Paul Barclay**

Date: **July 11, 2018**

Dear Researcher:

On 07/11/2018 the IRB approved the following human participant research until 07/10/2019 inclusive:

Type of Review:	UCF Initial Review Submission Form
	Expedited Review Category #6 & 7
Project Title:	Effects of Control Style, Workload, and Presence on Episodic
	Memory for Virtual Environments
Investigator:	Paul Barclay
IRB Number:	SBE-18-14153
Funding Agency:	
Grant Title:	
Research ID:	N/A

The scientific merit of the research was considered during the IRB review. The Continuing Review Application must be submitted 30days prior to the expiration date for studies that were previously expedited, and 60 days prior to the expiration date for research that was previously reviewed at a convened meeting. Do not make changes to the study (i.e., protocol, methodology, consent form, personnel, site, etc.) before obtaining IRB approval. A Modification Form **cannot** be used to extend the approval period of a study. All forms may be completed and submitted online at https://iris.research.ucf.edu .

If continuing review approval is not granted before the expiration date of 07/10/2019,
approval of this research expires on that date. When you have completed your research, please submit a Study Closure request in iRIS so that IRB records will be accurate.

Use of the approved, stamped consent document(s) is required. The new form supersedes all previous versions, which are now invalid for further use. Only approved investigators (or other approved key study personnel) may solicit consent for research participation. Participants or their representatives must receive a copy of the consent form(s).

All data, including signed consent forms if applicable, must be retained and secured per protocol for a minimum of five years (six if HIPAA applies) past the completion of this research. Any links to the identification of participants should be maintained and secured per protocol. Additional requirements may be imposed by your funding agency, your department, or other entities. Access to data is limited to authorized individuals listed as key study personnel.

In the conduct of this research, you are responsible to follow the requirements of the Investigator Manual.

This letter is signed by:

Page 1 of 2

Kamille Chap

Signature applied by Kamille Chaparro on 07/11/2018 12:14:06 PM EDT

Designated Reviewer

APPENDIX D:
INFORMED CONSENT DOCUMENT

University of
Central Florida

Effects of Control Style, Workload, and Presence on Episodic Memory for Virtual Environments

Informed Consent

Principal Investigator(s):   Paul Barclay, M.A.

Faculty Supervisor: Valerie Sims, Ph.D.

Research Assistants: Jason Parker, Johanna Hidalgo

Investigational Site(s):   Psychology Building
University of Central Florida
4111 Pictor Lane, Orlando, FL 32816

Why am I being invited to take part in a research study?

We invite you to take part in a research study because you are an undergraduate student at the University of Central Florida and have expressed interest in participating in psychological research.

You must be 18 years of age or older to participate in this study and you must not have a history of seizures or seizure disorder.

What should I know about a research study?

- Someone will explain this research study to you.
- Whether or not you take part is up to you.
- You can choose not to take part.
- You can agree to take part and later change your mind.
- Your decision will not be held against you.
- You can ask all the questions you want before you decide.

Who can I talk to?

If you have questions, concerns, or complaints, or think the research has hurt you, talk to the research team at 281-687-2073 or email the Principal Investigator at PABarclay@knights.ucf.edu.

This research has been reviewed and approved by an Institutional Review Board ("IRB"). You may talk to them at 407-823-2901or irb@ucf.edu if:

- Your questions, concerns, or complaints are not being answered by the research team.
- You cannot reach the research team.
- You want to talk to someone besides the research team.
- You have questions about your rights as a research subject.
- You want to get information or provide input about this research.

Why is this research being done?

This study is being done to measure the differences between passively observing a virtual environment (like watching a movie) and actively interacting with it (like playing a video game). We hope to understand why people remember their environments better when they interact with it.

How long will the research last?

We expect that you will be in this research study for approximately 1 hour.

How many people will be studied?

A maximum of 300 people will be in this research study.

What happens if I say yes, I want to be in this research?

- You will complete a brief informed consent process. During this process, a staff member will explain the study to you in full and you will have an opportunity to as any question you may have.
- You will complete a 1-hour in-person experimental session.
- You will be randomized into one of our six (6) potential experimental groups. Your group assignment will affect how you interact with the virtual task.
- You will complete a short (5-10 minute) virtual task. During this task you will be immersed in a virtual environment and will interact with the environment according to your group assignment. There are three (2) different factors which are being manipulated and you may be assigned to any combination of those three factors:
    1. High vs. Low Immersion: Participants in the high immersion condition will have the task presented using a virtual reality headset (Oculus Rift) and a set of noise cancelling headphone. Participants in the low immersion condition will have the task presented on a desktop LCD monitor and a set of disposable earbuds.
    2. Single-Task vs. Dual Task: Participants in the single-task condition will only be asked to explore the virtual environment. Participants in the dual-task condition will explore the environment while completing a secondary task. This secondary task requires the participant to press the space bar on the keyboard at specific times while exploring the environment.
- The condition you get will be chosen by chance, like flipping a coin. Neither you nor the experimenter will choose what treatment you get. You will have a one in six (6) chance of being placed in each condition.
- You will then complete a short questionnaire through Qualtrix.

How will I receive my SONA Credit?

SONA credit will be awarded within 48 hours of your scheduled appointment. You will receive one (1.0) SONA credits for this study. You will receive this credit whether or not you complete all study procedures and your credit will not be revoked if you later choose to withdraw from this study.

What happens if I do not want to be in this research?

Your participation in this study is voluntary. You are free to withdraw your consent and discontinue participation in this study at any time without prejudice or penalty. Your decision to participate or not participate in this study will in no way affect your continued enrollment, grades, employment or your relationship with the individuals who may have an interest in this study.

What happens if I say yes, but I change my mind later?

You can leave the research at any time it will not be held against you.

If you decide to leave the research, contact the Primary Investigator, Paul Barclay. All data collected from you will be deleted, but your SONA ID will remain listed under your appointment slot in the SONA system to ensure that you continue to receive credit for having participated.

Is there any way being in this study could be bad for me?

There is a small risk that people who take part will develop what is ordinarily referred to as simulator sickness. It occurs once in a while to people who are exposed to prolonged continuous testing in simulated environments. Symptoms consist of nausea and a feeling of being light-headed. The risk is minimized as a result of the short duration of each session in the simulator. If you experience any of the symptoms mentioned, please tell the researcher and remain seated until the symptoms disappear.

Will being in this study help me in any way?

We cannot promise any benefits to you or others from your taking part in this research. However, possible benefits include gaining insight into how psychological research studies are conducted.

What happens to the information collected for the research?

Efforts will be made to limit the use and disclosure of your personal information. The only personally identifiable information collected during this study will be your SONA ID, which will be used only to schedule and verify your appointment. All other data will be assigned a participant number (different from your SONA ID) for this study and will be and stored separately from the SONA system. We cannot promise complete secrecy. Organizations that may inspect and copy your information include the IRB and other representatives of this organization. All data collected for this study will be kept for 5 years before being deleted.

What else do I need to know?

The Psychology Department at UCF greatly appreciates your participation in research conducted within the department. If you wish to provide feedback on this experiment to the department, you will be provided with a Research Experience Feedback Form at the conclusion of this visit. This form is completely optional, but if you choose to complete it you may return it to the staff at the Faculty Suite, located on the third floor of the Psychology Building.

REFERENCES

Arvind Pala, P., N'Kaoua, B., Mazaux, J. M., Simion, A., Lozes, S., Sorita, E., & Sauzeon, H. (2014). Everyday-like memory and its cognitive correlates in healthy older adults and in young patients with traumatic brain injury: a pilot study based on virtual reality. *Disability and Rehabilitation: Assistive Technology*, *9*(6), 463-473.

Bäckman, L., Small, B. J., & Fratiglioni, L. (2001). Stability of the preclinical episodic memory deficit in Alzheimer's disease. *Brain*, *124*(1), 96-102.

Bakdash, J. Z., Augustyn, J. S., & Proffitt, D. R. (2006, July). Large displays enhance spatial knowledge of a virtual environment. In *Proceedings of the 3rd symposium on Applied perception in graphics and visualization* (pp. 59-62). ACM.

Bakdash, J. Z., Linkenauger, S. A., & Proffitt, D. (2008, September). Comparing decision-making and control for learning a virtual environment: Backseat drivers learn where they are going. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 52, No. 27, pp. 2117-2121). Sage CA: Los Angeles, CA: SAGE Publications.

Burgess, N., Maguire, E. A., Spiers, H. J., & O'Keefe, J. (2001). A temporoparietal and prefrontal network for retrieving the spatial context of lifelike events. *Neuroimage*, *14*(2), 439-453.

Bustamante, E. A., & Spain, R. D. (2008, September). Measurement invariance of the Nasa TLX. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 52, No. 19, pp. 1522-1526). Sage CA: Los Angeles, CA: SAGE Publications.

Carassa, A., Geminiani, G., Morganti, F., & Varotto, D. (2002). Active and passive spatial learning in a complex virtual environment: The effect of efficient exploration. *Cognitive processing*, *3*(4), 65-81.

Cheng, M. T., She, H. C., & Annetta, L. A. (2015). Game immersion experience: its hierarchical

    structure and impact on game-based science learning. Journal of Computer Assisted

    Learning, 31(3), 232-253.

Cheng, M. T., She, H. C., & Annetta, L. A. (2015). Game immersion experience: its hierarchical

    structure and impact on game-based science learning. *Journal of Computer Assisted*

    Learning, *31*(3), 232-253.

Chrastil, E. R., & Warren, W. H. (2012). Active and passive contributions to spatial

    learning. *Psychonomic Bulletin & Review*, *19*(1), 1-23.

Crook, T. H., Youngjohn, J. R., & Larrabee, G. J. (1993). The influence of age, gender, and

    cues on computer-simulated topographic memory. *Developmental*

    Neuropsychology, *9*(1), 41-53.

Gaunet, F., Vidal, M., Kemeny, A., & Berthoz, A. (2001). Active, passive and snapshot

    exploration in a virtual environment: influence on scene memory, reorientation and path

    memory. *Cognitive Brain Research*, *11*(3), 409-420.

Golledge, R. G., Dougherty, V., & Bell, S. (1995). Acquiring Spatial Knowledge: Survey Versus

    Route-Based Knowledge in Unfamiliar Environments. *Annals of the association of*

    American geographers, *85*(1), 134-158.

Hart, S. G. (2006, October). NASA-task load index (NASA-TLX); 20 years later. In Proceedings

    of the human factors and ergonomics society annual meeting (Vol. 50, No. 9, pp. 904-

    908). Sage CA: Los Angeles, CA: Sage Publications.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of

    empirical and theoretical research. In Advances in psychology (Vol. 52, pp. 139-183).

    North-Holland.

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach.* Guilford Publications.

Herlitz, A., & Rehnman, J. (2008). Sex differences in episodic memory. *Current Directions in Psychological Science*, 17(1), 52-56.

Jebara, N., Orriols, E., Zaoui, M., Berthoz, A., & Piolino, P. (2014). Effects of enactment in episodic memory: a pilot virtual reality study with young and elderly adults. *Frontiers in aging neuroscience*, *6*, 338.

Kit, D., Katz, L., Sullivan, B., Snyder, K., Ballard, D., & Hayhoe, M. (2014). Eye movements, visual search and scene memory, in an immersive virtual environment. *PLoS One*, *9*(4), e94362.

Kosslyn, S. M., Chabris, C. F., Marsolek, C. J., & Koenig, O. (1992). Categorical versus coordinate spatial relations: computational analyses and computer simulations. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(2), 562.

Landau, B., & Jackendoff, R. (1993). "What" and "where" in spatial language and spatial cognition. *Behavioral and brain sciences*, *16*(2), 217-238.

Lemogne, C., Piolino, P., Friszer, S., Claret, A., Girault, N., Jouvent, R., Allilaire, J.F., & Fossati, P. (2006). Episodic autobiographical memory in depression: Specificity, autonoetic consciousness, and self-perspective. *Consciousness and cognition*, *15*(2), 258-268.

Lewin, C., Wolgers, G., & Herlitz, A. (2001). Sex differences favoring women in verbal but not in visuospatial episodic memory. *Neuropsychology*, *15*(2), 165-173.

Long, C. J. (1996). Neuropsychological tests: A look at our past and the impact that ecological issues may have on our future. In Sbordone, R.J., Long, C.J. (eds.) Ecological Validity of Neuropsychological Testing, pp 1-14. GR Press: Delray Beach, FL.

Ma, R., & Kaber, D. B. (2006). Presence, workload and performance effects of synthetic

    environment design factors. International Journal of Human-Computer Studies, 64(6),

    541-552.

Meier, M. J. (1997). The establishment of clinical neuropsychology as a psychological specialty.

    In Maruish, M.E. and Moses, J.A. (eds). *Clinical Neuropsychology: Theoretical*

    Foundations for Practitioners, 1-31. Mahwah, NJ: Lawrence Erlbaum and Associates.

Morganti, F., Stefanini, S., & Riva, G. (2013). From allo-to egocentric spatial ability in early

    Alzheimer's disease: a study with virtual reality spatial tasks. Cognitive neuroscience,

    4(3-4), 171-180.

Nash, E. B., Edwards, G. W., Thompson, J. A., & Barfield, W. (2000). A review of presence and

    performance in virtual environments. *International Journal of human-computer*

    Interaction, *12*(1), 1-41.

Oertelt-Prigione, S., & Regitz-Zagrosek, V. (Eds.). (2011). *Sex and gender aspects in clinical*

    medicine. Springer Science & Business Media.

Paas, F. G., & Van Merriënboer, J. J. (1993). The efficiency of instructional conditions: An

    approach to combine mental effort and performance measures. Human factors, 35(4),

    737-743.

Paas, F., Renkl, A., & Sweller, J. (2004). Cognitive load theory: Instructional implications of the

    interaction between information structures and cognitive architecture. *Instructional*

    science, *32*(1), 1-8.

Parker, E. S., Eaton, E. M., Whipple, S. C., Heseltine, P. N., & Bridge, T. P. (1995). University of

    Southern California repeatable episodic memory test. *Journal of Clinical and*

    Experimental Neuropsychology, *17*(6), 926-936.

Parsons, T. D., & Rizzo, A. A. (2008). Initial validation of a virtual environment for assessment of

    memory functioning: virtual reality cognitive performance assessment

    test. *CyberPsychology & Behavior*, *11*(1), 17-25.

Parsons, T.D. (2011) Neuropsychological Assessment using Virtual Environments: Enhanced

    Assessment Technology for Improved Ecological Validity. In S. Brahnam (Ed.),

    Advanced Computational Intelligence Paradigms in Healthcare: Virtual Reality in

    Psychotherapy, Rehabilitation, and Assessment (pp. 271- 289). Germany: Springer-

    Verlag.

Parsons, T.D. (2016). Virtual environments for the assessment of memory. In *Clinical*

    Neuropsychology and Technology: What's New and How We Can Use It. Springer

    International, New York.

Parsons, T. D., & McMahan, T. (2017). An initial validation of the Virtual Environment Grocery

    Store. *Journal of Neuroscience Methods*, *291*, 13-19.

Pengas, G., Patterson, K., Arnold, R. J., Bird, C. M., Burgess, N., & Nestor, P. J. (2010). Lost

    and found: bespoke memory testing for Alzheimer's disease and semantic

    dementia. *Journal of Alzheimer's Disease*, *21*(4), 1347-1365.

Plancher, G., Barra, J., Orriols, E., & Piolino, P. (2013). The influence of action on episodic

    memory: a virtual reality study. Quarterly Journal of Experimental Psychology, 66(5),

    895-909.

Plancher, G., Gyselinck, V., Nicolas, S., & Piolino, P. (2010). Age effect on components of

    episodic memory and feature binding: A virtual reality study. *Neuropsychology*, 24(3),

    379.

Plancher, G., Piolino, P.(2017). Virtual reality for assessment of episodic memory in normal and

      pathological aging. In R.L. Kane and T.D. Parsons (Eds) *The Role of Technology in*

      Clinical Neuropsychology, 237-260.

Plancher, G., Tirard, A., Gyselinck, V., Nicolas, S., & Piolino, P. (2012). Using virtual reality to

      characterize episodic memory profiles in amnestic mild cognitive impairment and

      Alzheimer's disease: influence of active and passive encoding. *Neuropsychologia*, 50(5),

      592-602.

Ruff, R. M. (2003). A friendly critique of neuropsychology: Facing the challenges of our future.

      Archives of Clinical Neuropsychology, 18(8), 847-864.

Sauzéon, H., Arvind Pala, P., Larrue, F., Wallet, G., Déjos, M., Zheng, X., Guitton, P., &

      N'kaoua, B. (2012). The use of virtual reality for episodic memory assessment: effects of

      active navigation. *Experimental psychology*, *59*(2), 99.

Sauzeon, H., N'kaoua, B., Arvind Pala, P., Taillade, M., & Guitton, P. (2016). Age and active

      navigation effects on episodic memory: a virtual reality study. *British Journal of*

      Psychology, *107*(1), 72-94.

Sauzéon, H., N'kaoua, B., Pala, P. A., Taillade, M., Auriacombe, S., & Guitton, P. (2016).

      Everyday-like memory for objects in ageing and Alzheimer's disease assessed in a

      visually complex environment: The role of executive functioning and episodic

      memory. *Journal of neuropsychology*, *10*(1), 33-58.

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

Stanney, K. M., Salvendy, G., Deisinger, J., DiZio, P., Ellis, S., Ellison, E., Fogleman, G.,

      Gallimore, J., Hettinger, L., Kennedy, R., Lackner, J., Lawson, B., Maida, J., Mead, A.,

      Mon-Williams, M., Newman, D., Piantanida, T., Reeves, L., Riedel, O., Singer, M.,

      Stoffregen, T., Wann, J., Welch, R., Wilson, J., & Witmer, B. (1998). Aftereffects and

sense of presence in virtual environments: Formulation of a research and development

agenda (Report sponsored by the Life Sciences Division at NASA Headquarters).

International Journal of Human–Computer Interaction, 10, 135–187.

Stone, J. M., & Towse, J. (2015). A working memory test battery: Java-based collection of

seven working memory tasks. *Journal of Open Research Software*, *3*.

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive

load. *Educational psychology review*, *22*(2), 123-138.

Teigen, K. H. (1994). Yerkes-Dodson: A law for all seasons. *Theory & Psychology*, *4*(4), 525-54

Astur, R. S., Ortiz, M. L., & Sutherland, R. J. (1998). A characterization of performance

by men and women in a virtual Morris water task:: A large and reliable sex

difference. *Behavioural brain research*, *93*(1-2), 185-190. 7.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving and W. Donaldson (Eds)

Organization of Memory, *1*, 381-403. Oxford, UK: Academic Press.

Tulving, E. (1983). *Elements of Episodic Memory*. Oxford, UK: OUP Oxford Press.

Wallet, G., Sauzéon, H., Pala, P. A., Larrue, F., Zheng, X., & N'Kaoua, B. (2011). Virtual/real

transfer of spatial knowledge: Benefit from visual fidelity provided in a virtual

environment and impact of active navigation. Cyberpsychology, Behavior, and Social

Networking, 14(7-8), 417-423.

Wallet, G., Sauzéon, H., Pala, P. A., Larrue, F., Zheng, X., & N'Kaoua, B. (2011). Virtual/real

transfer of spatial knowledge: Benefit from visual fidelity provided in a virtual

environment and impact of active navigation. *Cyberpsychology, Behavior, and Social

Networking*, *14*(7-8), 417-423.

Welch, R. B., Blackmon, T. T., Liu, A., Mellers, B. A., & Stark, L. W. (1996). The effects of pictorial realism, delay of visual feedback, and observer interactivity on the subjective sense of presence. *Presence: Teleoperators & Virtual Environments*, *5*(3), 263-273.

Wilson, P. N., & Péruch, P. (2002). The influence of interactivity and attention on spatial learning in a desk-top virtual environment. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*.

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

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