

Learning and Memory in Virtual Spaces

by

Seth Tee

B.A., University of British Columbia, 2015

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Arts

in the

Educational Technology and Learning Design
Faculty of Education

© Seth Tee 2015

SIMON FRASER UNIVERSITY

Spring 2015

All rights reserved.

However, in accordance with the *Copyright Act of Canada*, this work may be reproduced, without authorization, under the conditions for "Fair Dealing." Therefore, limited reproduction of this work for the purposes of private study, research, criticism, review and news reporting is likely to be in accordance with the law, particularly if cited appropriately.

Approval

Name: Seth Tee
Degree: Master of Arts
Title: *Learning and Memory in Virtual Spaces*
Examining Committee: Chair: Allan MacKinnon
Associate Professor

Kevin O'Neill
Senior Supervisor
Associate Professor

John Nesbit
Supervisor
Professor

David Kaufman
Internal/External Examiner
Professor

Date Defended/Approved: November 20, 2015

Ethics Statement



The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

- a. human research ethics approval from the Simon Fraser University Office of Research Ethics,

or

- b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University;

or has conducted the research

- c. as a co-investigator, collaborator or research assistant in a research project approved in advance,

or

- d. as a member of a course approved in advance for minimal risk human research, by the Office of Research Ethics.

A copy of the approval letter has been filed at the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Burnaby, British Columbia, Canada

update Spring 2010

Abstract

Space and context are fundamental factors in cognition which have powerful effects on learning, memory, and recall. Previous studies have shown that changes in the physical context between learning and assessment tasks can degrade recall performance. The research on virtual context effects, however, is scant, especially in the area of learning. Virtual environments are increasingly utilized in educational technology research and application without a careful understanding of space and context. This study investigated the effect of context in a virtual space on learning and memory using a between groups experiment that controlled the use of context changes and the level of immersion in the environment (2D or 3D). It contrasted two existing hypotheses explaining these effects: context-dependence and situational model updating. The results suggest an interaction between the level of immersion in the environment and whether or not a context change occurred.

Keywords: virtual spaces; context effects; learning; memory;

Acknowledgements

I would like to thank Kevin O'Neill for the supervision and guidance that has allowed me to reach beyond what I thought I could achieve in this MA program. With his help, not only did I learn the theories and concepts of educational technology, but also gained a keen interest in discourse and research. I also thank John Nesbit who was a great help in putting together this project in the design and analysis stages.

I am also grateful to Yuting Zhao, Adam Zollo, and Ken MacAllister who volunteered their time to help facilitate the experiments. Without their help, this study would not have been possible.

Thanks,

Seth Tee

Table of Contents

Approval	ii
Ethics Statement	iii
Abstract	iv
Acknowledgements	v
Table of Contents	vi
List of Tables	vii
List of Figures	viii
List of Acronyms	ix
Glossary	x
Chapter 1. Introduction	1
Chapter 2. Literature Review.....	4
Chapter 3. Research Questions and Methods	20
Chapter 4. Results	32
Chapter 5. Discussion	42
References	57
Appendix A. Updated Learning and Assessment Tasks.....	59

List of Tables

Table 4.1.	Group Demographic Information	33
Table 4.2.	Assessment Descriptive Statistics	34
Table 4.3.	Group Assessment Results	36
Table 4.4.	Comparison of Assessment Times for 3D versus 2D training environments and Change versus No-Change conditions	40

List of Figures

Figure 3.1.	Learning Task Rules Outline.....	26
Figure 3.2.	2D and 3D Virtual Environments	27
Figure 3.3.	Experimental Groups	29
Figure 4.1.	Assessment Score Frequency Distribution	35
Figure 4.2.	Assessment Time Frequency Distribution	35
Figure 4.3.	Assessment Score Group Means Comparison	38
Figure 4.4.	Assessment Time Group Means Comparison	39
Figure 4.5.	Assessment Score Group Means Comparison for Female Participants	41
Figure 4.6.	Assessment Time Group Means Comparison for Female Participants	41

List of Acronyms

SFU Simon Fraser University

PFL Preparation for Learning

Glossary

Context Effects	A measurable effect on learning and memory performance produced by the environmental context of the task
Immersion	The degree to which a subject experiences a virtual world as if it were a physical one
Presence	The experience of being in a virtual environment as if it were a physical one
Virtual Space/Environment	A computer generated environment that can displayed visually and interacted with by the subject

Chapter 1.

Introduction

Virtual spaces and environments have evolved dramatically since the first video games and simulators were created. Today, the technology is available to produce complex visual worlds that feel closer and closer to real life. The last few decades have seen waves of new games, applications, and simulations that provide users with increasingly powerful experiences of "presence", a sense of "being there", in a virtual environment (Barfield and Hendrix, 1996). In the field of educational technology, the development of new environments is outpacing the development of scholarly understanding about how virtual environments are processed by learners, and how they influence learners' performance.

The importance of physical context is well established in the literature on context and memory (Smith et al., 1978, Eich, 1985) and by studies in knowledge transfer (Barnett and Ceci, 2002) and situated learning (Greeno, 1998). These studies have demonstrated significant memory degradation when participants are asked to recall memorized information after changing physical contexts by moving to a different environment or changing details within the same environment. These effects have also been found on memory tests in virtual environments (Radvansky and Copeland, 2006). As virtual environments become more realistic, does the mind process them in ways that are increasingly similar to physical environments, and do these same effects emerge? Studies on the efficacy of immersive games for learning have overall been inconclusive (Schrader & Bastiaens, 2011); could changes in the learning and assessment environments in these studies act as an extraneous variable, interfering with the performance results? Is learning affected the same way that memory is affected? What aspects of the environment produce these effects? These open questions reflect the fact that there is a general lack of understanding of how virtual contexts affect peoples' learning and memory.

Previous research has produced at least two competing hypotheses: context dependence and situational model updating. The *context dependence hypothesis* states that context effects are caused by the integration of learned knowledge with cues in the environment and the subsequent loss of cues with an environment change (Smith et al., 1978). The *situational model updating hypothesis* states that context effects arise from the increased cognitive load produced when entering a new environment (Radvansky and Copeland, 2006). Both hypotheses have been tested and confirmed to produce effects on memory in physical and virtual environments. However, they have not been directly compared with regards to immersive virtual environments and learning.

This study attempted to clarify the influence of virtual context on learning and assessment through a controlled experiment in which the immersiveness of a virtual environment was manipulated during learning and assessment. Separate contexts were designed which controlled for the level of immersion in the virtual environment. A questionnaire was used to confirm different subjective experiences of immersion in each environment. This experiment did not demonstrate a statistically detectable effect of context change and immersion on assessment scores; however, it justifies further study in that it demonstrated an effect on subjects' time to complete the assessment.

These findings can be analyzed and discussed to describe context effects on learning in virtual environments, expand the knowledge on the existing hypotheses explaining the nature of context effects, and prescribe best practices for training and assessment in virtual spaces. Firstly, this data can help establish immersion or virtual presence as a contributing factor to context effects; a finding which could imply that technologies enabling virtual presence are subject to context effects. Evidence for either the *context dependence* or *situational model updating* hypothesis would help to explain the cognitive processes behind how the mind processes knowledge and the environment. A further impact would be felt in the research and application of learning in virtual spaces. If performance can be influenced by context in virtual spaces, then the characteristics of that space must be controlled in any experiment. Lastly, the use of virtual spaces in training and assessment, such as in games and simulators, can be better informed to account for context effects.

What is meant by a “virtual space” needs to be clearly defined to avoid conflicting meanings. The concept of space as a collaborative environment (Wahlstedt et al., 2008) will not be discussed in this paper. For the purposes of this research, a virtual space will be defined as a simulated, computer-generated context that allows the user to experience and interact. Immersion, or virtual presence, is defined as the feeling of actually being present in the computer generated environment (Sheridan, 1992, Barfield and Hendrix, 1995).

To differentiate between learning and memory as measured in previous studies (Radvansky and Copeland, 2006, Murnane et al. 1999), what is meant by learning within the scope of this study also needs to be defined. In what follows, I will consider “learning” as referring to measurable learning outcomes defined by Barnett and Ceci as “the knowledge gained as a result of a learning activity that leads to the construction of new – or changes in existing – knowledge structures” (2002). Learning is thus differentiated from memory alone by the requirement that the learner must exhibit an application of a knowledge structure, or interconnected pieces of knowledge, rather than just recognition or recall of simple information.

Chapter 2.

Literature Review

Context and Memory

Physical Space Effects

The spaces and contexts in which a learner is situated is documented to influence their performance on memory tasks. This phenomenon has been well explored through many experiments comparing context and performance in physical spaces. Previous studies have investigated recall and recognition memory tasks while changing the physical context between learning and assessment activities to elicit an effect on performance (Smith et al., 1978). Smith et al. used this experiment structure to document significant effects in an early study in this area testing memorization of words (1978). Subjects were asked to recall and recite lists of words while participating in groups that either switched spaces between a cubicle and a classroom, or remained in the same space. Time between memorization and recall was kept constant across conditions, but the difference between remaining in a single context and relocating was found to have observable results on performance, with relocated subjects recalling fewer words overall. In a similarly structured follow-up study, the task was designed to observe word recognition rather than recall (Smith, 1986). Again the results showed significant performance degradation in the groups that changed context. Later studies have found similar effects from changing contexts across a range of different memory tasks. A few explanations have been proposed to account for this effect.

Context Dependence

The context dependence hypothesis involves a connection between acquired memories and the contextual details in which the individual experienced them. It proposes that people use cues in the environment to aid their memories, so therefore the removal of those cues during an assessment produces a decrease in performance ability. This explanation was briefly hypothesized as a possibility in the Smith et al. study (1978) but a

further exploration of the idea is found in Eich's research into integrated imagery (1985). Using a similar experiment structure and a list of words to memorize, the context effects were replicated, with same-context participants recalling higher numbers of words in a shorter period of time than those who switched contexts. However in Eich's study physical context was controlled not by changing rooms, but by changing features within the room, such as decorations. Eich further increased the subject's dependence or non-dependence on the context by asking them to either integrate their memorized objects with features of the room by visualizing them together, or by visualizing the target item as an isolated image. While the subjects in the changed-context group displayed slower and less accurate recall overall, the isolated imagery group displayed a minimal decrease in performance. These findings suggest that context effects could be caused by the level of mental integration made by the observer between the physical environment and the training material.

A subsequent investigation into what factors mediate memory and context was performed by Murnane et al. (1999) who posited that memory is structured according to items, context, and ensembles. Building from previous ideas of object and context integration, ensembles are defined as a distinct third element that is composed of a combination of items and context. Murnane et al. showed that rich visual contexts were much more likely than simple visual contexts to show degradation effects in memorization of word pairs. A computer screen was used to present the word pairs, while contexts were varied between photo-based backgrounds (rich visual context) or plain colour ones (simple visual context). Word matches in the rich visual context showed a significantly lower hit rate when changing environments than in the simple visual context. These results demonstrated that not all contexts are processed in the same way, and that context effects can be addressed by controlling the environment, such as combining words with photographic backgrounds to create ensembles. This research also supports the idea of context dependence, where removal of contextual cues is the cause of the resulting decrease in performance. However, context dependence isn't the only hypothesis that has been proposed to account for the effects of space and context on learning and memory.

Situational model updating

While the context dependence hypothesis attributes memory degradation to the loss of cues in an existing space, the situational model updating hypothesis emphasizes the effect of introducing a new environment whose features have to be processed by the mind (Radvansky and Copeland, 2006). The situational model updating hypothesis, also described as an event horizon model, combines several existing concepts in the research on spaces, and builds on research in several different areas. It suggests that foregrounding, event segmentation and, retrieval interference (all of which will be discussed later) work together to cause this phenomenon. The hypothesis states that cumulative interference caused by the process of updating one's understanding of the environment (known as the situational model) effectively reduces the memory capacity for information obtained from a previous environment.

Experiments by Radvansky and Copeland (2006) showed memory degradation produced by having subjects simply walk through doorways between two similar spaces. This study was performed in both a physical and a virtual environment. Subjects were instructed to pick up a colour-coded object, move to a different location either within the same room or another room, and then were asked to recall the objects they had picked up after they had moved. Context change was applied by dividing one large room into two smaller rooms with a doorway providing access between the two spaces. No image integration or emphasis on features of the environment was enforced, and both spaces were roughly similar. The doorway between the two rooms acted as a context cue, triggering a set of cognitive processes to adjust to the new environment. This adjustment, in turn, reduced the subject's memory performance. The results of the experiment confirmed the occurrence of context effects in both the physical and virtual spaces: subjects were slower and less accurate at recalling their selected objects when passing through the doorway than when they remained in one space.

Event segmentation

Event segmentation can be described as a memory strategy that breaks stored experiences into segmented sets rather than a continuous whole (Kurby and Zacks, 2008, Swallow et al., 2009). Special queues in the environment, such as pauses in action or

changes of the environment, are used to divide these experiences into segments. Several studies have found support for event segmentation in memory. In Swallow et al.'s (2009) experiment, object recognition was tested on movie clips with an object recognition task. Each movie clip included one or more event boundaries. These boundaries existed as scene changes or situations where the action of the film moves into a different environment (e.g. from an interior to an exterior setting). Participants were not explicitly told to pay attention to event boundaries. After viewing, subjects were tested on their recognition of objects occurring anywhere within the film. The test performance showed better recognition accuracy for objects that appeared close to event boundaries. Further data to support the event segmentation hypothesis was gathered in experiments in which segmentation was produced using spatial changes within verbal, narrative instructions (Sahakyan and Kelley, 2002). Sahakyan and Kelley found that explicitly asking subjects to imagine a new spatial context, without ever altering their physical context, was enough to produce an observable effect on a word recall task similar to those used in previous experiments.

Foregrounding

The process of foregrounding is described as an increased access to memory for relevant objects and decreased access to less relevant objects. In this hypothesis, objects in an environment are constantly assessed and prioritized by the mind to be foregrounded or not. This effect was documented in a study using narrative text events (Glenberg et al., 1987) in which subjects were tested at periodic intervals on their ability to remember objects from the narrative. Specific objects were foregrounded based on their relationship with the protagonist (i.e. objects that the protagonist possessed at a given moment of the story were considered to be foregrounded). Objects that were foregrounded (that is, in the possession of the protagonist at the time of assessment), had higher rates of recognition than objects left behind by the protagonist, even if the objects left behind had been mentioned more recently.

Retrieval Interference

One further addition to the situational model updating hypothesis discovered in a later study by Radvansky et al. is retrieval interference. This effect occurs when objects

become associated with two or more event models. Having multiple competing event models and slower retrieval of the relevant event model, results in slower retrieval of object information (Radvansky et al., 2011). In the study, which also controlled context changes, a three room situation was included in which colour-coded objects were carried across three distinct spaces before the recall task. The results affirmed the idea of retrieval interference, as the three-room group showed greater degradation than the two-room change group, while the no-change group showed the least. By entering three distinct rooms with an object, that object itself became associated in participants' minds with three competing situational models, producing the highest interference within the experiment.

These effects cannot be explained by context dependence alone, because the possible environments do not differ in any discernible details. Furthermore, in the retrieval interference study using three rooms, context dependence does not explain the greater memory degradation when entering a third room. Instead, the results are best predicted by the situational model updating hypothesis. In the situational model updating interpretation of these results, event segmentation occurs when the subject encounters a visual cue, such as a doorway, that predicts a new space. The previous memory set is terminated, and a new one is started in preparation for a new environment. The foregrounding process attempts to prioritize certain memories; it removes irrelevant information from memory, reducing access, while increasing performance for relevant, situational information. Then, when participants are asked to recall previous knowledge, retrieval interference hinders the retrieval of information that is not distinctly linked to the current situational model.

Thus, the situational model updating hypothesis offers an alternative, multiple-stepped explanation to context effects, with supporting evidence. Rather than conceiving the integration of memories with the context, situational model updating posits that the cognitive processes triggered when encountering novel spaces are the cause of context effects.

In the context of the present study, it is important to note that none of the hypotheses discussed above limits memory effects to physical environments alone, and some experiments have successfully used virtual environments and multimedia to

produce significant effects (Murnane et al., 1999, Radvansky & Copeland, 2006). It is also important to note that the literature has heretofore primarily focused on simple memory tasks (mainly recall and recognition of words or items) since the early studies in the field, rarely deviating from these kinds of tasks.

Knowledge Transfer & Situated Learning

Knowledge Transfer

The research areas of knowledge transfer and situated learning look beyond simple memory tasks and consider more complex learning tasks involving a deeper level of understanding and problem solving. Although the effects on memory described in the literature (Eich, 1985; Smith, 1986; Radvansky and Copeland, 2006) very likely influence learning as well, they cannot be generalized to explain or predict all learning effects when considering learning outcomes. Does performance degradation occur across contexts in learning tasks? To posit possible answers, a good starting point is the field of knowledge transfer, where for many years researchers have been studying the ability of students to generalize learned knowledge from one situation (physical or otherwise) to another. Knowledge transfer is a contentious topic, with some claiming that transfer and generalization of knowledge is the ultimate goal of education (Spence & Weisberg, 1986, Barnett & Ceci, 2002) while others (who will be discussed later) have debated whether knowledge transfer, as it is traditionally understood, even exists (Greeno, 1998; Anderson et al. 1996; Bransford & Schwartz, 1999). Regardless, context effects have played a significant role in the concept of knowledge transfer, and the literature on knowledge transfer can inform further study in both physical and virtual environments.

Near and far transfer

In their survey of the literature on transfer in 2002, Barnett and Ceci declared knowledge transfer a “salvageable concept” and attributed inconsistencies in the existing research to a lack of a well-defined taxonomy within knowledge transfer research. Going back to early studies that first explored ways to help students apply lessons learned in one situation to another situation, they organized the existing data and categorized the wide array of phenomena classified as “transfer.” The lack of definition of the idea of “transfer”,

they argued, obscured the import of the research findings in this area. Barnett and Ceci's work provided a framework for understanding knowledge transfer, broken down into categories of transfer: knowledge domain, physical context, temporal context, functional context, social context, and modality. Each dimension within their framework can be further split into near or far transfer, the degree of difference between two contexts. This framework allows researchers to isolate key dimensions of transfer and, in doing so, addresses some of the inconclusive results in the literature on knowledge transfer. One of Barnett and Ceci's key takeaways is that the far transfer of learning (and not only memory) can be measured and quantified. Furthermore, they propose physical context as a key dimension within transfer that warrants more study, due to many studies training and assessing in the same environment (usually a school or lab), and not considering the effects of context (Barnett and Ceci, 2002). However, that does not mean that no studies have tried to address the transfer of learning from one physical context to another.

Learning transfer between contexts

One of the earliest instances of knowledge transfer research was an informal activity and observation documented by Charles Judd (1908). In this experiment, Judd taught students the concept of refraction in the classroom, and then tested their ability to apply that knowledge by trying to hit a target in a nearby lake. Some students received classroom training while others did not, and the students with classroom training were able to perform better at the activity. Although not exactly scientific in its methods, this study proved to be influential to many researchers in education (Barnett & Ceci, 2002).

A much later study tested students' knowledge of statistics with a pre-test and post-test experiment that took place in two different contexts (Fong et al., 1986). The researchers taught lessons on statistics in the classroom but phoned students at home to perform a survey testing their knowledge. They were able to confirm that students performed better on the post test, showing that they were able to apply their knowledge in a different context from where they learned. Although these studies begin to address learning and context, it is important to note that they merely looked at whether learning can transfer across different physical contexts, and did not examine the effect of the context change itself.

Spencer and Weisberg (1986) showed that knowledge transfer can be affected by context changes. This study on the use of analogy to facilitate the application of one concept over two different problems, divided context change and non-change groups to examine the effect of context. It implemented a transfer task across distant knowledge domains (ie. the topics of the problems) and spaces both near and far physically. Subjects were posed an open ended problem in one room, and provided a sample solution. A second open-ended problem was presented following the first one, with a shift in knowledge domains; in this case from a military general trying to reach a fortress to a medical situation in which high-intensity rays must treat a tumor. The second problem would either be presented in the same room for the non-change group, or a novel room for the change group. Performance measures were attained by coding the subject's solutions and looking for use of a previously provided concept from the first problem in the new knowledge domain. The researchers' findings displayed degraded performance in the change groups, similar to documented effects of context change on memory -- participants were less likely to use their learned concepts after they had switched physical contexts. These findings provide evidence that the effects of space not only affect learning, but do so in a manner similar to effects on memory by degrading performance scores.

Situated Learning

The concept of situated learning emerged in opposition to the concept of knowledge transfer. Proponents of situated learning have claimed that knowledge is always situated in a concrete situation, and does not transfer between tasks (Greeno, 1998). In practice, they recommend that learners should learn in the same specific situation where they will perform, in order to avoid the need for transfer altogether. The literature on knowledge transfer and the consistent result of performance degradation was a strong impetus for researchers to adopt this new perspective (Greeno, 1998, Bransford and Schwartz, 1999). It was argued that transfer is simply too difficult to find, and cannot be depended upon to facilitate learning (Bransford and Schwartz, 1999). Situated learning theorists provide three propositions that outline the situated learning hypothesis, which are relevant to the investigation in this paper:

1. Concepts are not stored in isolation, but remain closely coupled with their background situations
2. Concepts do not take the same form across different situations
3. Situation effects occur across diverse tasks (Anderson et al., 1996, Yeh and Barsalou 2006, Cobb and Bowers, 1999).

If knowledge cannot be “transferred” from one context to another, how could learning take place? Bransford and Schwartz summarize the situated learning perspective on transfer with the concept of preparation for future learning (PFL) (1999). PFL rejects the idea that knowledge is transferred along with the learner; instead, knowledge is relearned in a new context. In this conception teaching, instead of providing transferrable knowledge, serves to prepare learners to more effectively acquire new knowledge when in a new environment.

While situated learning endorses a strategy for learning that occurs within a complete environment, with learners performing an authentic task or simulation, some proponents such as Anderson et al. (1996) still acknowledge that a better hypothetical model is needed to understand why transfer-like effects happen at all. It is not enough to have a strategy to avoid transfer degradation; the problem itself needs to be understood.

In both the ideas of knowledge transfer and situated learning, spatial context is seen as a critical facet to learning. The effects of knowledge degradation when transferring between multiple dimensions, including context, have been well documented in the literature, though researchers differ on how to proceed. Some have pursued methods and strategies to maximize far transfer and reduce the negative effect of context changes on learning. On the other hand, proponents of situated learning also accept the need for further study on learning and context, but prescribe learning and performing in as similar situations as possible to minimize context effects. In the theoretical and experimental papers reviewed here, the general opinion seems to support the existence of context effects on learning in physical environments, though solid explanations as to why these effects occur are lacking (Barnett and Ceci, 2002, Anderson et al., 1996).

Virtual Space Effects

As discussed earlier, context effects on learning and memory in the physical world are now well documented in the literature. These same kinds of effects have not been fully studied in virtual environments, even though the technology for creating virtual environments and the use of virtual environments in training become increasingly ubiquitous. However, previous studies on how humans experience virtual spaces and how memory is affected by them do provide valuable insights into the effects of virtual contexts on learning.

Barfield and Hendrix demonstrated how spatial features of a virtual environment affect a user's sense of place (1996). Their study investigated the idea of "presence" (a rough measure of how closely a person's experience in a virtual environment matches that of a physical one), by asking participants to report on their experiences in high-fidelity and low-fidelity 3D virtual environments. Each subject was presented a virtual environment controlled with a high quality and low quality rendering setting (affecting how the image looks), as well as varying the frames per second of the display (affecting the smoothness of animations). Experience of the environment was reported by each participant on a "presence" questionnaire measuring the degree to which users felt they were immersed in the environment. The results showed that more highly detailed environments rendered at a faster frame rate produced a much greater feeling of presence in virtual environments.

Further research in 3D environments helped to identify other specific and adjustable aspects of a virtual world that increased a sense of immersion (Bystrom et al., 1999). Tasks and activities that required users to interact with the environment, giving users control of movement and perspective, and the use of head tracking devices to control the direction of gaze, were all found to create a greater sense of presence. This finding was also replicated in studies that tested not just interactions, but the illusion of interactive elements in the environment by giving the user control of devices (a mouse and keyboard) that did not produce any effect in the virtual world (Regenbrecht and Schubert, 2002).

These results show that certain kinds of virtual environments, or certain aspects of the environment, can increase or decrease a person's sense of being within a virtual

space. It does not necessarily follow that because they feel immersed in a virtual environment (as if it were a physical one) that their minds are processing the environment in the same ways that people have been shown to process a physical one. It also doesn't follow that context effects on learning or memory would degrade performance in virtual spaces. However, there are studies delving into this specific topic.

As noted above, Radvansky and Copeland's study on situational model updating used an interactive 3D environment that included several of the key features shown to produce an immersive experience proposed by Barfield and Hendrix and Bystrom et al. - a high fidelity 3D environment, a high framerate, user control of perspective, and interaction with the environment (Radvansky and Copeland, 2006). It should be noted that Murnane et al. also used virtual environments (2D contexts) in their study by using photographs in the background of their rich visual contexts. Their results showed that static 2D images, not just 3D environments, are enough to illicit an effect on memory, while plain colour 2D environments were not. These findings show that context in a virtual space, without any changes in physical space, can produce significant context effects on memory in a variety of virtual environments. What of more sophisticated virtual environments and more complex tasks?

Virtual Learning Environments

One of the areas in education where virtual environments might make a profound impact is the creation of digital learning environments such as games and simulators. The technology available has the potential to create a deep sense of immersion, interactivity, and enjoyment for the user, and many have argued for the benefits of such immersion for learning. Psotka (1995) proposed many advantages that virtual environments (particularly those that produce an experience of virtual presence) may offer for learning, and studies are indeed finding that learners report stronger motivation to learn in immersive games rather than in more traditional learning environments (Verhagen et al., 2011, Barfield and Hendrix, 1999). It seems natural, to some, to predict that highly interactive environments would afford better learning. In addition to Psotka, other researchers have expressed their belief that more immersive and authentic learning environments will provide better

assessment results (Thomas, 2010, Kolb, 2002). These assumptions, however, have not been supported by currently available studies, which so far have found mixed results.

Several attempts were able to find beneficial results using immersive video games for education. In an experiment by Slater et al., subjects were taught to play a novel game of 3D chess within two possible virtual environments: a low-immersion 2D interface and a high-immersion 3D one (1999). Learners were then assessed by how well they could replicate the position of a set of pieces on a physical chess board, following moves permitted by the rules of the game. The group that received the immersive learning environment showed better performance overall.

Beale et al. (2007) studied the benefits of Re-Mission, a video game created to teach young cancer patients about cancer treatment while interacting with a 3D environment. In this experiment, the treatment group received PCs allowing them to play Re-Mission, while the control group could only play other, unrelated games. Two knowledge tests were given to the subjects in the form of an 18-item questionnaire (The method of distribution for the questionnaire (paper or digital) was not addressed in Beale et al.'s publications, but it is clear that the questionnaire was not part of the Re-Mission game.) The results showed that the group that played the Re-Mission game scored significantly higher on the knowledge tests than the group that didn't.

Despite these positive findings, other experiments have found contradictory results. Parchman et al. (2000) looked at learning outcomes in classes for navy electronic technicians. They created four conditions for different instructional methods: computer-based drill and practice, computer based instruction, classroom instruction, and a digital learning game. All participants' learning was tested using an in-class pencil and paper test. The results of the experiment showed that computer-based drill and practice and computer-based instruction produced the highest test scores, while the game based condition scored roughly as well as the classroom instructed group.

Further negative effects of learning using video games were documented by Schrader & Bastiaens (2010). The researchers delivered a learning task on the physics of light to eighth graders. Each participant was randomly assigned to one of two groups: one learning with an immersive game, and another an HTML website. The learning task was

followed by a questionnaire on virtual presence and a questionnaire assessing the learned knowledge. It should be noted that both groups were given the same questionnaires, and though the method of delivery was not reported, it was not the game environment. Data analysis indicated that while subjects reported a higher sense of presence within the immersive environment, test scores showed significantly better average performance by the web page viewers. Schrader and Bastiaens attributed the difference in test scores to the increased cognitive load involved in interacting with the virtual environment (2010).

Such mixed outcomes have cast doubt on the effectiveness of virtual environments as a teaching tool. There are no conclusive results on whether virtual environments or immersion can improve learning. However, a brief look at the literature also reveals inconsistencies in how learning and assessment environments are handled. None of the studies reviewed above carried out the assessment of learning via the virtual environment or measured the subject's experience of immersion in the environment. Several studies did not even report on the medium used to deliver the post-test questionnaire (Bealt et al., 2007, Schrader and Bastiaens, 2010). Slater et al. switched from virtual environments to a physical chess board (1995) and Parchman et al. switched from virtual environments to a classroom test (2000). These details relate back to the idea of context effects, where a change in context was enough to significantly affect memory performance in both physical and virtual environments (Smith 1976, Eich 19865, Murnane et al.. 1999, Radvansky & Copeland, 2006). This potential experimental confound raises a number of questions regarding how experiments on virtual learning environments should address context, space, and facilitation of knowledge transfer.

Areas for Investigation

The literature on context effects has long shown that space and context have a significant and reliable effects on learners' performance in memory tasks. Performance on both recall and recognition tasks have consistently been shown to degrade after changes in context between training and assessment tasks, when the content of the training and assessment are held constant. As discussed in prior chapters, researchers have proposed several explanations for this phenomenon: context dependence (the integration of knowledge with features in the environment), event segmentation (the tendency to segment experiences

into categories according to time and context), situational model updating (the process of realigning one's mental model with a new environment), and foregrounding (increased resources dedicated to processing certain stimuli in the environment) (Eich, 1985, Radvansky and Copeland, 2006, Kurby and Zacks, 2008, Glenberg et al., 1987). These mechanisms have been studied and reported in many different environments and tasks. The extensive research on physical environments has used name recognition, reading comprehension, analogous knowledge, and various other tasks to produce significant results. Recent research also moved in the direction of virtual spaces, using shape recognition tasks that uncovered significant effects of context change (Radvansky and Copeland, 2006). These findings provided ample evidence for the effect of context on learning and memory, and suggests many new research areas to pursue.

Context Knowledge Transfer Research

Going back to early educational psychology research, the literature on knowledge transfer provides a suitable basis for further research on learning contexts. While knowledge transfer studies began to struggle with questions regarding the generalizability of learning and the nature of knowledge, it is a field that has yielded valuable results into how humans attain and use knowledge. Barnett and Ceci (2002) argued that, when broken down into better-defined parts, knowledge transfer studies have provided many insights into learning. They specifically singled out the study of transfer between physical contexts as a major area lacking further attention, given its reliable effects and implications for the practice of education. The field of knowledge transfer research looks at the question of context as a transfer issue: how can knowledge be learned in one context and generalized to another context? How can knowledge be learned to better transfer to another context? How can training methods and environments be controlled to produce the best performance?

Contexts Conceived as Situated Environments

The theory of situated learning is heavily invested in the idea that knowledge is situated in the environment, and largely dismisses the idea of transfer as misguided. Context plays a significant role in situated learning, because it places a large emphasis on the environment and the need to train and assess learning in the same social and

physical environment in which performance is expected (Greeno, 1998). Interest in context effects has been building in the literature (Anderson et al., 1999), and the emergence of immersive virtual environments introduces many questions to the discussion. For example, how do humans perceive situatedness in virtual environments? Does the fidelity of a virtual learning environment to a real (physical) environment affect learning and memory? What characteristics of an environment affect situatedness in learning?

Learning Tasks and Virtual Environments

In previous studies, context effects on learning and memory have primarily been explored in physical contexts, using physical assessments like paper and pencil tests. More recently, studies have used electronic interfaces and even virtual environments to test memory through recognition and recall-- but this is only the beginning. The question still examined in the existing literature is, do learners within virtual environments experience the same context effects as they would in a physical environment? To delve further into this question, two factors can be explored: the learning task and the learning environment.

Research on virtual environments and learning is still in its early stages, and the variety of learning tasks studied so far do not compare to those studied in physical environments. The effects on memory and learning have been confirmed in physical spaces using many different learning tasks: some general memory tests and some specific skills or forms of learning (Smith et al., 1978, Eich 1987, Glenberg et al., 1987). This cannot be said for virtual environments, where the tasks examined in formal studies have been almost exclusively focused on memory (Radvansky and Copeland, 2010, Murnane et al., 1999). Large areas of learning, beyond rote memory, are left unexamined, while learning in virtual environments continues to expand, in the use of online learning resources, games, and simulations.

Furthermore, there are few explanations for what characteristics of a virtual environment are likely to produce context effects. Unlike physical environments, which are limited by the physical world (and most often the confines of a research lab), virtual environments have the potential to present vastly different scenarios; 2D or 3D, abstract or realistic rendering, claustrophobic spaces or large-scale landscapes, and so on. As

technology development progresses, virtual worlds will offer increasing flexibility to create novel environments. One compelling point of interest is the user's level of immersion in the virtual world. Virtual spaces have been shown to produce strong experiences of immersion, which can be adjusted by controlling aspects of the environment (Barfield & Hendrix, 1996, Bystrom et al., 1999). What role does virtual immersion play mediating context effects on learning and memory?

Discrepancies in Existing Results

The use of virtual environments to support learning, particularly in research on gaming, has been found to produce conflicting results with regard to in performance differences. Immersive environments, or even the users' experience of presence, have not reliably been shown to affect learning outcomes, despite greater user-reported motivation and enjoyment (Parchman et al., 2000, Schrader and Bastiaens, 2011). However, the effect of spaces on learning is not well understood, and the influence of context changes have not been fully addressed in many of the studies that involve learning in virtual environments. Many experiments use a variety of interfaces and environments for learning and assessment, switching between them for convenience but not necessarily controlled for the experiment (Slater et al., 1995, Beale et al., 2007). The use of virtual environments is a timely topic as technology is enabling educators to use high-fidelity 3D spaces for learning, and as researchers continue to study the applicability of 3d spaces to a wide array of educational problems.

How Practice can be Improved

One remaining area of knowledge to be investigated is how studies on the influence of context changes in virtual spaces can impact real-world practice, where virtual environments are already being used in the form of web applications, games, and simulators. While learning and memory have been studied in virtual environments, the findings are still limited and not generalizable to current uses of virtual environments both in terms of the learning tasks tested and the virtual environments used. The conclusions stop short of providing findings that can support strong recommendations with regard to future development.

Chapter 3.

Research Questions and Methods

The purpose of this study was to examine context effects on learning and memory in virtual environments, and to test the relative strength of two leading explanations for context effects. The study further examines immersion (a sense of being within a virtual space) as a possible moderating factor in context effects. The following questions are put forward:

1) Does context affect learning in virtual environments?

Research has repeatedly shown that effects of context on memory are prevalent in many tasks when switching spaces (Smith et al., 1978, Eich 1987, Glenberg et al., 1987, Radvansky and Copeland, 2006). But virtual environments may differ in effect from the physical spaces in which the majority of educational research has been conducted. Virtual environments offer a greater range of worlds and settings that can be created and delivered; yet they are also limited by a lack of physical presence, graphical fidelity, user interaction, and more. It is uncertain how the differences between physical and virtual environments affect how people process the world. Further, more complex learning tasks need to be observed instead of basic recognition or recall activities. Memory tests are not representative of much of the learning happening in vocational, educational, or social settings.

2) Does performance degrade when switching from one virtual environment to another?

Building on the first question, the nature of the context effect needs to be explored; whether performance improves or degrades when context is switched. The effects found in previous research have included significant degradation in memory performance when switching spaces (indicated by slower and less accurate recall) or a lesser ability to apply a given concept (Smith et al., 1978, Eich 1987, Glenberg et al., 1987). Existing hypotheses, such as context-dependence and situational model updating, predict performance degradation in virtual environments, but they do so in different ways (Smith

et al., 1978, Eich 1987, Radvansky and Copeland, 2006). If a significant effect is found in virtual environments, the condition, direction, and magnitude would need to be measured to support or refute these hypothesis.

3) Are performance effects dependent on the level of immersion in the virtual environment?

If a significant effect can be measured, it raises the further question of how context influences learning functions. What characteristics of the environment produce context effects? This study proposes immersion, or the level of “presence” in the environment, as a possible factor. Immersion is defined here as the degree to which the learner feels that they are “inside” a virtual environment as if it were a real environment. It is reasonable to suspect that users who feel they are in a real environment would process it similarly, and the context effects found in physical environments would also be present. Immersion can be controlled within a virtual environment, and current research on virtual spaces has identified several key features (such as 3D rendering, consistent frame rates, animations, user interaction, and control of 3D perspective) that can greatly increase the subject's experience of virtual presence (Barfield & Hendrix, 1996, Bystrom et al., 1999). Determining how immersion alters context effects will provide insight into what factors affect learning most and how developers can control these factors to create increasingly immersive experiences.

4) How does the direction of movement between high-immersion and low-immersion environments alter performance?

It is important to know how the switch between low-immersion and high-immersion environments affects learning and memory performance, beyond whether an effect exists. There are two benefits to comparing both low and high immersion environments. First, the results will help to prescribe the best use of immersion to optimize training, or how to best control it in an experimental context. Second, the change in directions from low-to-high and high-to-low can generate data that supports or refutes the context dependence hypothesis and the situational model updating hypothesis in virtual environments.

Context dependence proposes mental integration of memories and details in the environment. It requires highly detailed and immersive environments to produce significant effects. On the other hand, the situational model updating hypothesis suggests that effects are caused by a spike in processing load that occurs when entering a new space. It depends on novel environments to elicit effects, and would predict high effects when moving into more highly immersive environments. The two explanations predict differing results in the direction of context change where an effect would occur. Context-dependence predicts the largest effect when changing from a high-immersion environment to a low-immersion one, due to the loss of detail in the environment that would interrupt the integration between the environment and memory. The situational model updating hypothesis, conversely, predicts a large effect when moving from a low-immersion environment into a high-immersion one, because of the cognitive load required to process the new environment.

The answers to these questions would offer insights into how learning and memory function in virtual spaces, and how virtual spaces in general are processed. The findings may help to explain the inconclusive findings in the area of learning in virtual environments, and influence future studies. Tools and programs can also use this knowledge to design for context effects in training or assessment.

Methods

This study uses an experimental, between-groups design to answer the key research questions put forward. The group conditions are created by controlling two factors: the level of immersion in the assessment environment, and whether or not subjects switch environments (high or low immersion) between the learning and assessment tasks. The four resulting groups consist of change groups who switch either from a low-immersion to a high-immersion environment or from a high-immersion to a low-immersion environment, and non-change groups who are trained and assessed in either a low-immersion or high-immersion environment. The experimental environments are designed to isolate the effects of context change and immersion levels, while keeping all other factors, such as training content and assessment difficulty, equivalent.

The use of a between-groups experimental design aligns with most of the work done in this area of research, and enables this study to use existing methods of data collection (namely participant recruitment in universities, and the training and assessment task procedure). This allows for informed comparisons with the observations and results found in prior studies. It is important for the purposes of this study that the results and conclusions obtained can be generalized to the greater problem of space and context already established by the literature. Establishing a cause-and-effect relationship between memory effects and change in spaces is a key part of addressing the research question. Using a controlled environment and removing intervening variables is an essential part of these results, because environment effects are the phenomenon being measured.

The performance results of the assessment within these controlled environments is the dependent variable, and is used to identify effects under each condition. Differences in performance will be analyzed between the assessment scores and completion times of the groups to identify any significant effects. There are two independent variables being controlled for this experiment. The first independent variable is the level of immersion in each environment. For both the learning and assessment tasks, the environment can be either a high-immersion or low-immersion environment. The second independent variable is the occurrence of a context change between the learning and performance tasks. Groups in the change condition will perform their assessment task in either a high-immersion or low-immersion virtual environment than their learning task, while the control condition will remain in the same environment (either low or high immersion). How the change group's scores compare with the non-change group's scores can show an effect potentially caused by the changing of environments. Comparisons between these two groups will help to answer the first two research questions, regarding whether context change produces the effects studied. An additional moderating variable proposed is the type of virtual environments for each group: high immersion or low immersion. The effect of the moderating variable is directed at answering the third research question, regarding the influence of immersion in the environment. Overall, the experiment will consist of four groups (high to low, high to high, low to high, low to low) to which subjects are randomly assigned.

The procedure used in this experiment is similar to several other context studies that have successfully found context effects using two environments (Eich, 1985, Murnane et al., 1999, Radvansky and Copeland, 2006). The control group will remain in the same environment for both learning and assessment tasks, although which environment they interact with (low-immersion or high-immersion) will be randomly assigned. The treatment group will be automatically switched to a different environment for the assessment task, depending on their assigned group. Changes may occur from a low-immersion environment to a high-immersion environment, or vice versa.

The null hypothesis for this experiment is that changes in virtual environments will have no effect on learning and performance (as opposed to physical environments). This can be confirmed by finding no significant differences between the average scores or times of the experimental (change) and control (no-change) groups. Without a significant difference in performance when changing virtual environments, the question of immersion becomes irrelevant.

The context-dependence hypothesis predicts that the largest degradation of assessment scores will occur in a context change from a high-immersion to a low-immersion environment, losing the environmental markers linked with the subject's memory. This hypothesis can be rejected by finding a greater context effect size in the low-to-high-immersion than the high-to-low-immersion conditions, or finding no effect at all. The situational model updating hypothesis predicts the greatest performance degradation in conditions where subjects switch from a low to a high immersion environment, increasing the cognitive load to process the environment. This hypothesis can be rejected by finding smaller or context effect sizes in the low-to-high-immersion conditions than the high-to-low conditions. It can also be rejected by finding no effect of context on the dependent variable at all.

Subjects

The primary target subjects were university students at the SFU Surrey or SFU Burnaby campuses. As an incentive to participate, all participants were offered \$5.00 for completing the experiment. Additionally, 3 of the top scorers on the learning task were entered into a random draw for a chance to win a \$50.00 cash prize. The intent of the

prize draw was to encourage participants to take the task seriously, so as to score as high as possible. University students have been used subjects in several related experiments, so they should provide results comparable to previous studies (Radvansky and Copeland, 2006, Smith, 1986, Eich, 1985). Students also offer a suitable study group because their results can be generalized to many situations in which virtual learning environments may be applied, such as formal education or workplace training. Each experiment session was advertised on the SFU campus, including times that potential subjects could sign up or drop in.

Learning and Assessment Task

Through a web site, subjects were required to complete both a learning task introducing novel knowledge, and an assessment task testing their performance based on the knowledge learned. Times and assessments scores were tracked by the web site. Both stages were performed at a personal computer and monitored by the investigator so that subjects are not able to cheat, or get stuck on the interface.





The learning task involved understanding the behaviour of four “machines” depicted in the software (see figure 3.1). Each machine was represented by a simple shape (e.g. circle, triangle, square, diamond) and performed a unique mathematical or logical function (e.g. add all numbers together, rearrange numbers from first to last). The subjects were asked to study (via text and images) the rules by which each machine operates. The materials displayed and named each machine, and described its function. For each machine, one example of its use was provided. The last page of the training materials included a table outlining the function of all the machines. Users were encouraged to read over the materials until they were confident in their understanding before proceeding to the assessment. A minimum time study time of three minutes was required before the user was permitted to continue to the next stage.

Summary

Let's go over the key points of our number machines.

- Number machines take an input number, perform certain steps, and produce an output
- All number machines have a name, a symbol, and a unique set of steps

The table below outlines the key elements you should remember for each machine.

Name	Symbol	Steps
Adder		1. Add all the digits together
Exchanger		1. Order the digits in the number from lowest to highest 2. Swap the first and last number
Dicer		1. Multiply the input by 10 2. Divide the number by 2
Jumbler		1. Multiply by 2 2. Reverse the order of the digits

You've now completed the training portion of this exercise. You are free to go back and review any information before continuing to the assessment task. When you are ready, press continue to begin the assessment.

Figure 3.1: Learning Task Rules Outline

The assessment task required the participant to complete a 12-question quiz based on the properties of the machines. Subjects were not able to return and review the rules, or have the rules available to them for reference; the quiz had to be completed entirely from memory. All questions were presented in multiple-choice form. The first four questions simply required subjects to recall the names of the machines presented. The goal of this task was to test the recognition and recall performance of the subject, in a way similar to previous experiments. The following eight questions required subjects to calculate an output number given a starting number and a machine. There were two questions for each machine. This set of questions required the participants to be able to

recognize the machines, recall the rules, and apply the rules to a novel situation. All subjects were given a maximum time of ten minutes to complete the assessment task, after which the application automatically submitted their responses.

The design of the learning task included simple recall and recognition of names based on the shape and colour of the machine. However, it further required subjects to memorize and be able to apply logical and mathematical functions as they have been described. This task aimed to test problem solving and knowledge application beyond basic recall and recognition, unlike previous studies in virtual environments. The learning defined and measured in this study focused on the participants' ability to internalize a set of generalizable logical rules and relate it to abstract symbols. This kind of task, it is hoped, will better represent the learning students and staff experience in the real world, and fulfill this study's requirement for a complex learning task.

Virtual Environments

Two kinds of virtual environments were created, each with a learning (delivering the educational content) and assessment (testing application of learned knowledge) portion. Both environments were rendered in a web browser, and presented the exact same learning and assessment materials. The learning materials were presented using text and images, and navigation was handled using clickable buttons or arrows to move to the next or previous slide, or answer a question.



Figure 3.2: 2D and 3D Virtual Environments

The low-immersion environment consisted of 2D web pages created using Javascript, HTML, and CSS. The 2D interface allowed users to view content comprised of text and images on a flat page. Participants could navigate the content using navigational

buttons labeled next and previous. Pages were animated to slide left or right and the transitions took 2 seconds to complete.

A parallel high-immersion environment was created by introducing 3D rendering, animations, and user control of the camera to the interface; 3 factors that have been shown to increase immersion (Barfield & Hendrix, 1996). This application was created using Javascript, HTML 5, CSS, and WebGL technology and Blender to create and render 3D content. On the computers used in this experiment, the 3D environment ran at roughly 50fps consistently, resulting a smooth visual experience. Navigation was handled using the mouse, allowing users to click on next or previous arrows, or on a multiple-choice answer. Transitions were animated in the 3D space so that users maintained a sense of position within the environment and, as in the 2D environment, completed in 2 seconds. The machine symbols were also represented in 3D, and were animated to rotate. Users could use the mouse to click and drag to change the perspective of the camera, including the ability to look back at previous content. (Users could return to a previous slide in the 2D space as well as in the 3D space, but the ability to look around was a unique characteristic of the 3D space.) All content, including the assessment questions, were embedded in the 3D world. Besides the presentation of the context, the content was exactly the same as in the 2D environment, using the same text and symbols for training.

The design of these virtual environments offered several benefits. Because they run on any modern browser such as Firefox or Chrome, they could be used on any current computer, including most computers available on the SFU campus. They also provided a consistent experience for all users regardless of the computer used. Running the application in a browser also provided an interface that most participants are familiar with, so there was no need to train them on how to use the interface itself, outside of a few on-screen cues. The web application was designed to lead participants through the process without involving the investigator in interacting with participants. This allowed the participants to stay focused on their own screens and reduced variability in how the investigator interacted with each participant. The 2D and 3D environments differed in a few subtle details. Colour, font, and basic images have been shown to be suitable methods of changing context without affecting the utility of a 2D website (Murnane et al., 1999). Since these differences in the 2D and 3D environment can act as context markers (colour

of the environment, green arrows, font arrangement, etc.) that set them apart to test the context-dependence hypothesis. Most importantly, the 2D and 3D environments are differentiated by the level of immersion allowing the study to test the isolated effects of immersion on context change.

Experimental Groups

The two main conditions in this experiment were the 2D assessment and 3D assessment groups. They were distinguished by the virtual environments used in the assessment stage. The second set of conditions were the no-change control group where participants were trained and assessed in the same environment, and a change group, where the environment was changed between the learning and assessment tasks.

	2D Assessment	3D Assessment
No-Change	2D → 2D	3D → 3D
Change	3D → 2D	2D → 3D

Figure 3.3: Experimental Groups

All participants were automatically assigned a condition and group by the web application. The system selected groups by prioritizing the group with the fewest participants at the moment, and randomly selecting between tied groups. Users were not made aware of their group until after the experiment was completed.

Procedure

As each subject began their session, a facilitator would introduce them to the experiment using a standard script for all subjects. This script touched on the general steps of the experiment, including the learning and assessment task; however details of the conditions and different virtual environments were not mentioned.

Once properly oriented, the subjects were assigned a computer to work on. The first page included a form for collecting basic participant information, including familiarity with computers, familiarity with 3D applications, age, gender, etc. Once this form has been

submitted, the program automatically assigned the participant to a specific experimental group, and loaded the learning task environment. The learning task required a minimum study time of three minutes for participants to learn about the four machines. When ready, participants could continue to the assessment task after a brief loading screen. The assessment task was presented as a series of multiple-choice questions, and participants had 10 minutes to complete the 12-question assessment. The participants' responses and final scores in the performance session were recorded when users submitted their assessment, or as time ran out.

After submitting their results, subjects were asked to fill out a brief questionnaire comprised of 8 to 16 fields on presence, based on their experiences (see Appendix B). The questionnaire each participant received was contingent on the environments experienced by the participant, and is used to verify that participants' sense of immersion was different between the 2D and 3D environments. At the completion of a session, subjects could opt to write down their email address to be entered for the 3 cash prizes drawn from the top scoring participants based on assessment scores. Finally, subjects were debriefed by the facilitator and provided an opportunity to ask any questions they had about the study.

Data Collection and Analysis

All research data was stored on a secure server database. The web application saved all user data, completion times, assessment scores, and questionnaire responses automatically and anonymously. At the end of the experiment, the data were retrieved for analysis.

User data were used to determine if the condition group composition had any abnormalities that might explain away an effect. The user data were also used to compare results between user groups (e.g. if gender makes a difference in assessment scores). Two variables were used to measure performance: assessment scores and completion times. The completion times were measured from when the user first entered the assessment to when they submitted their responses. Users had a maximum time limit of 10 minutes to complete the assessment but no minimum time so they were likely to spend as little time as possible to complete. The assessment scores were totaled out of 12 for

the 12 assessment questions; 4 questions measured recall of names and 8 questions required application of learned rules. Users were incentivized to score as high as possible with top scorers having an opportunity to gain an additional \$50 prize. Finally, the questionnaire responses would help verify that users experience a different level of immersion between the high and low immersion environments.

Chapter 4.

Results

Participants

Data collection took place over a period of 6 months and in that time, 139 data points were collected over 12 sessions. All participants were SFU undergraduate or Masters students, who were recruited on two SFU campuses. Eighty-nine participants were female and 51 were male. Their average age was 22.47 years ($SD = 5.63$). Their average self-reported familiarity with computers was 4.02 on a scale from 1 to 5 ($SD = 0.89$) while their average familiarity with 3D applications was 2.24 on a scale from 1 to 5 ($SD = 1.15$) with 1 indicating “not familiar” and 5 indicating “very familiar”. These numbers show that most participants had a high familiarity with computers; though their familiarity with 3D applications (e.g. playing 3D games or simulators) was lower than anticipated.

In the low immersion assessment groups, the change group consisted of 35 participants (20 females and 15 males) while the no-change experimental group consisted of 36 participants (26 females and 10 males). Between the two groups, the average age, GPA, familiarity with computers, and familiarity with 3D applications did not show significant differences (Table 4.1). The gender ratio was identified as a possible source of bias in the results, and was considered when analyzing the group results.

In the high immersion assessment group, the change group consisted of 35 participants (24 females and 11 males) while the no-change experimental had 34 participants (19 females and 15 males). Once again, the reported ages, GPA, familiarity with computers, and familiarity with 3D applications do not show any noteworthy differences between the 2 randomly-selected groups, though the gender ratio could bias the results (Table 4.1).

Table 4.1

Demographic Information by Groups

	2D no-change	2D Change	3D no-change	3D Change
Subjects	35	35	34	35
Female:Male Ratio	26 : 9	20 : 15	19 : 15	24 : 11
Average Age	21.00 (3.89)	22.43 (5.61)	23.12 (5.91)	23.46 (6.53)
Average GPA	2.91 (0.52)	3.05 (0.47)	3.20 (0.43)	3.01 (0.75)
Average Familiarity with Computers	3.91 (0.97)	4.09 (0.74)	4.12 (0.82)	3.97 (1.03)
Average Familiarity with 3D apps	2.06 (0.93)	2.22 (1.13)	2.38 (1.21)	2.30 (1.31)

Immersion Questionnaire

The immersion questionnaire consisted of a series of 8 rating scale questions, where a lower response indicated a deeper level of immersion experienced by the user in the environment. Of the eight questions, one was reverse-scaled (i.e., a higher response represented a higher sense of immersion). Users were assigned questionnaires based on the virtual environment they experienced (3D versus 2D applications); users who experienced both environments would receive both questionnaires, for a total of 16 questions. The possible extremes of the overall immersion score were therefore 40 (not immersive) or 10 (very immersive).

The immersion questionnaire scores were confirmed to be acceptably reliable, with Cronbach's alpha calculated at 0.78. On average the 3D environment was rated 21.06 for immersiveness ($SD = 4.95$), while the 2D environment was rated 23.35 ($SD = 5.07$). A significant difference was found between the ratings of the two environments $t(202) = 3.26$, $p < .01$. Of all respondents, 66% found the 3D environment more immersive, 28% found the 2D environment more immersive, and 6% found the two environments equally immersive. These results suggest that the 3D environment was perceived as more immersive for users, although they did not report as large a difference as expected by the researcher. Possible explanations for this finding will be discussed below. Ultimately, the

basic requirements of the study were met, as to the level of immersion in the learning and assessment tasks was controlled.

Overall Assessment Results

Table 4.2

Assessment Descriptive Statistics

	Mean	Median
Scores (out of 12)	9.86 (2.59)	11
Times (seconds)	224.94 (83.04)	214

The two results relevant to the hypotheses being tested in the study are the participants' assessment scores and their times to complete the assessment. Each constitutes a measure of performance. Students' scores on the assessment were found to be reliable, with Cronbach's alpha calculated at 0.83. The overall mean score on the final assessment was 9.86 out of 12 ($SD = 2.59$), and the median score was 11 out of 12 (see Table 4.2). Plotting the frequency of the scores reveals a substantial ceiling effect, which places limitations on the findings of this study (see Figure 4.1). The ceiling effect on assessment scores will be discussed further below.

The mean time to complete the assessment was 224.93 seconds ($SD = 83.04$), while the median was 214 and the standard deviation was 83.04. When plotted, the assessment times reveal a slight skew to the right, but are more normally distributed than the assessment scores (Figure 4.2).

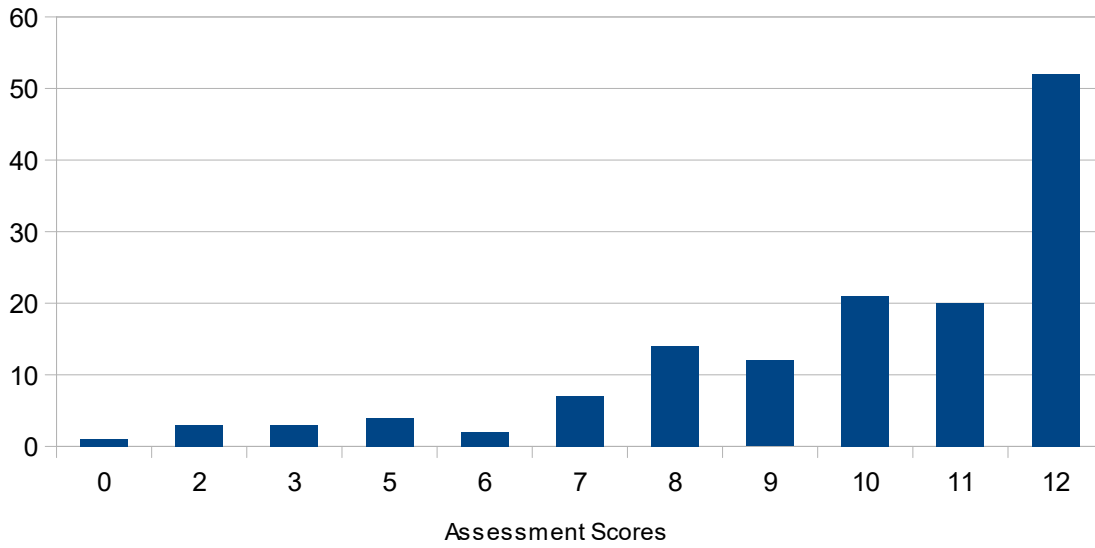


Figure 4.1: Assessment Score Frequency Distribution

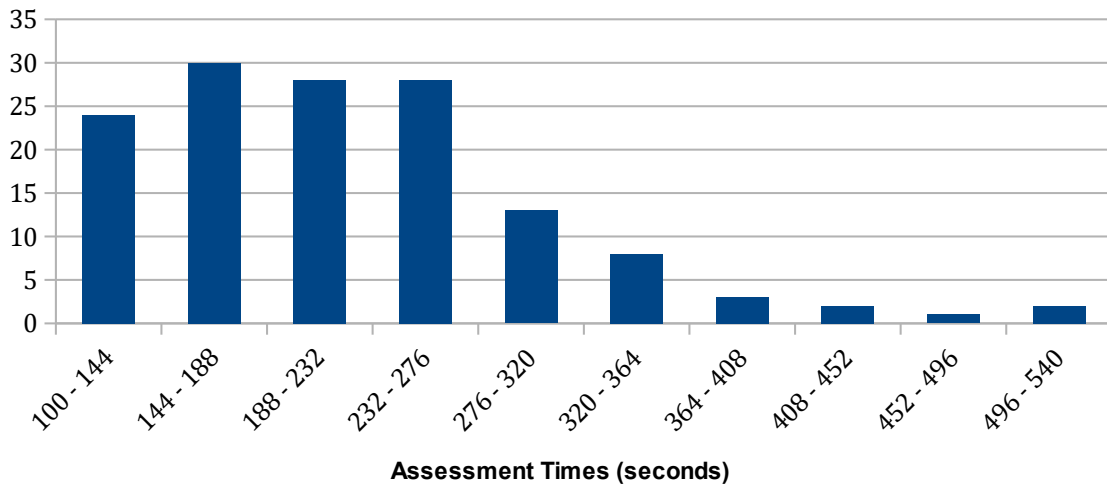


Figure 4.2: Assessment Times Frequency Distribution

Main Effect

The no-change group scored slightly higher on the recognition questions, while the change group scored slightly higher on the problem solving questions. Looking at the

average times to complete the assessment, the no-change group appeared to take slightly longer at 231.51 seconds ($SD = 83.40$) while the change group completed it in 204.23 seconds ($SD = 82.28$ sec) on average. In both the recognition and problem solving questions, the no-change group scored higher. The average time to complete the assessment was found to be higher on average in the change group ($M = 246.56$ seconds, $SD = 74.30$) as opposed to the no-change group ($M = 218.06$ seconds, $SD = 89.88$), indicating that performance degraded in the change group. It should be noted that the high immersion assessment no-change group performed the best of all groups in terms of both scores and completion times, and the high immersion assessment change group performed the worst.

Table 4.3

Assessment Results

	2D No-Change	2D Change	3D No-Change	3D Change
Average Score (out of 12)	_9.69 (2.28)	_9.80 (3.14)	_10.46 (1.77)	_9.47 (3.08)
Average Recognition Score (out of 4)	_3.74 (0.61)	_3.57 (1.01)	_3.94 (0.34)	_3.65 (0.88)
Average Problem Solving Score (out of 8)	_5.94 (2.32)	_6.23 (2.34)	_6.51 (1.72)	_5.82 (2.32)
Average Time (seconds)	231.51 (83.4)	204.23 (82.28)	218.06 (74.30)	246.56 (89.88)

Assessment Scores 2-Way ANOVA

Assessment scores were the main measure of an effect on cognitive load, and three effects are relevant to this study. The change variable determines whether the environment change had an effect on scores. The assessment immersion variable describes whether the immersion level of the assessment environment had an effect on scores. Finally, the interaction between the environment change and the assessment immersion determines if there is an effect of the immersion level of the training environment.

When plotted, the data appears to display an interaction: in the low immersion assessment, scores stay roughly the same between the no-change to change conditions but the high-immersion assessment scores are shown to decrease (Figure 4.3). The

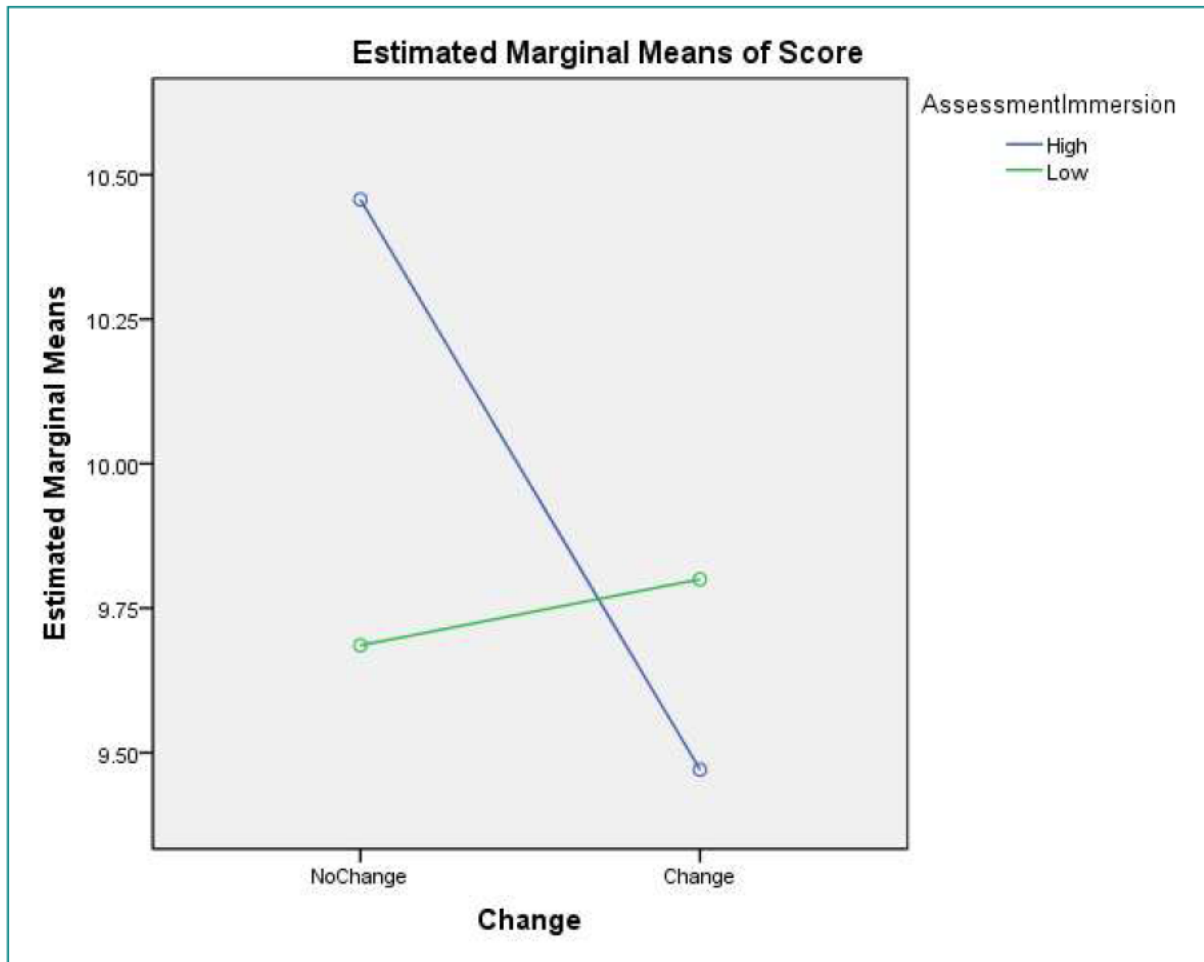
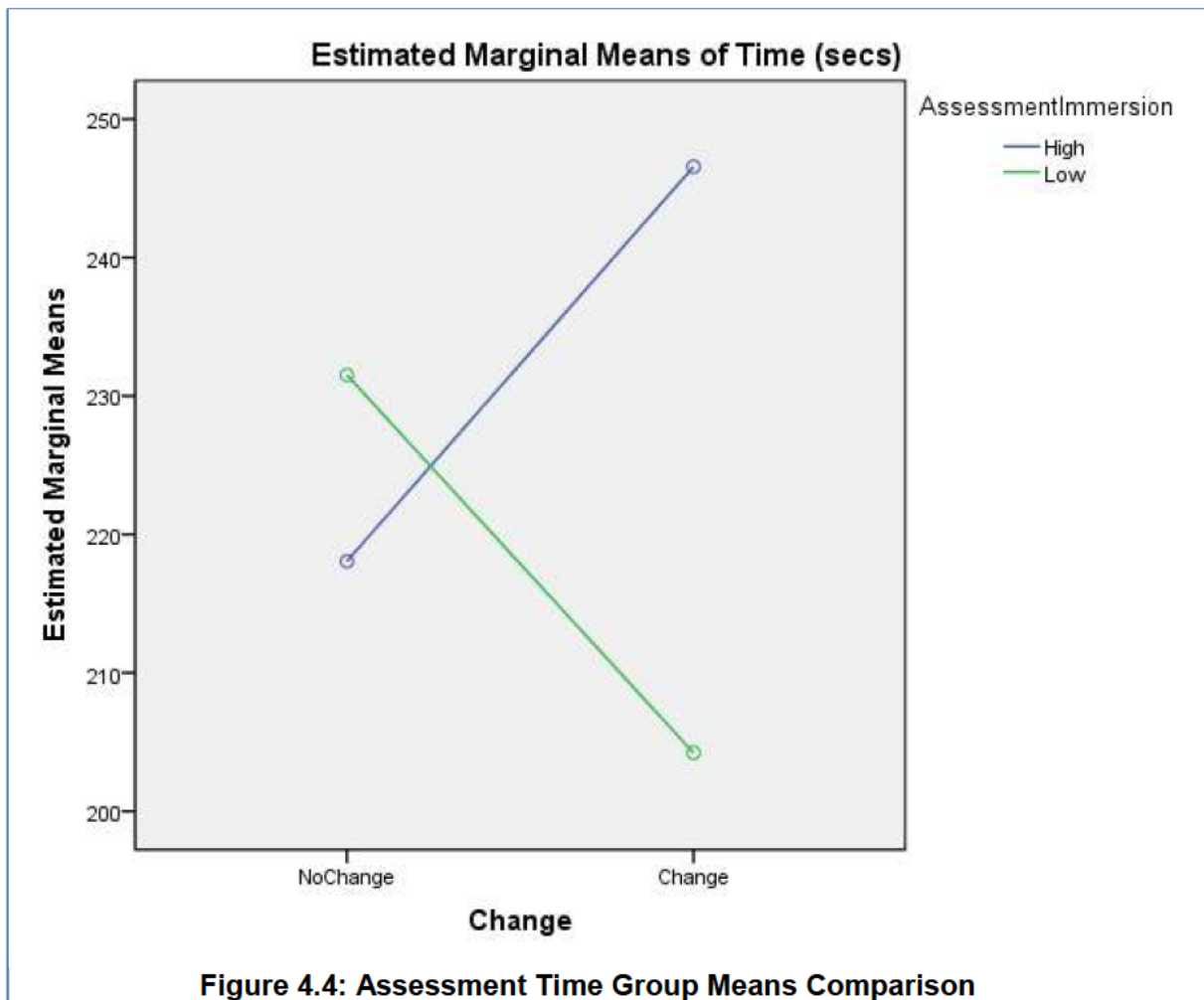


Figure 4.3: Assessment Score Group Means Comparison

ANOVA results, however, show no significant effects or interactions. Neither the change treatment $F(1, 135) = 0.98, p = 0.32$ nor the assessment immersion level $F(1, 135) = 0.25, p = 0.62$ of the assessment environment showed a statistically significant effect. Similarly, the interaction between the change and assessment immersion variables (as seen in the graph) was not found to be significant $F(1, 135) = 1.57, p = 0.21$.

Assessment Times 2-Way ANOVA

The time taken to complete the assessment was the second measure of performance in this study. Again, we are interested in a change effect, assessment immersion effect, or an interaction between the two. When plotted, the data show an interaction between the assessment immersion variable and environment change: the high-immersion assessment group's performance degraded (they took more time) when changing environments; but the low-immersion assessment group improved (took less time) (see Figure 4.4).



The ANOVA results show that neither the assessment immersion, $F(1,135) = 1.07$, $p = 0.30$ or environment change, $F(1,135) = 0.002$, $p = 0.97$, variables had a significant effect on assessment times. We can conclude that testing in a 2D or 3D environment or changing the environment by themselves do not have an effect on assessment times.

However, a significant interaction was detected between the change and assessment immersion treatments $F(1,135) = 3.98, p = 0.05$.

To further investigate the influence of this interaction on assessment times, I calculated the effect size for the assessment times of the groups that trained in the 2D environment versus the groups that trained in the 3D environment. Although effect size was small ($d = 0.32$), the significant interaction showed that the level of immersion in the subject's training environment has an effect on their time to complete; subjects that trained in the 2D environment took longer to complete the task than those in the 3D environment.

To better understand how the training environment affected assessment times, I compared the effect of the training environment immersion when a change occurs and when it doesn't. The results of this analysis showed that there was a significant effect on assessment times when a change occurs ($t(66.07) = 2.04, p = .048$) with a moderate effect size ($d = 0.49$); but there was no effect when a change did not occur ($t(66.94) = 0.72, p = 0.48$) (Table 4.4).

Table 4.4

Comparison of Assessment Times for 3D versus 2D training environments and Change versus No-Change conditions

	<u>3D Training</u>		<u>2D Training</u>		t-test
	M	SD	M	SD	
Change	204.23	82.28	246.56	89.88	0.05
No-Change	218.06	74.30	231.51	83.40	0.48

Gender Differences

As mentioned earlier, the participants recruited for this study were not gender balanced. To account for any effects on the data due to the differences in assessment performance between male and female subjects, we analyzed the score and time data again using an ANOVA with male participants removed. An ANOVA on the assessment scores of the female students, like the ANOVA on the overall scores, did not detect any significant effect from the change treatment, $F(1,85) = 0.59, p = 0.45$; assessment

environment, $F(1,85) = 0.004$, $p = 0.95$; or an interaction, $F(1,85) = 2.77$, $p = 0.10$. However, a plot of the means shown in Figure 4.5 shows a similar pattern to that found for all subjects, with a change from low to high immersion resulting in degraded performance. The assessment times for female participants (Figure 4.6) also show a similar shape to the overall results; times appeared increased when switching to a high immersion environment and decreased when switching to a low immersion one although no significant effects were discovered with the change treatment, $F(1,85) = 0.14$, $p = 0.71$; assessment environment, $F(1,85) = 0.38$, $p = 0.54$; or an interaction, $F(1,85) = 2.47$, $p = 0.12$. Overall, the data for female participants resembles the data from all participants but with no significant effects. This should show that the distribution of male participants' scores are not the only factor in determining the shape of the interactions between the assessment immersion and change variables.

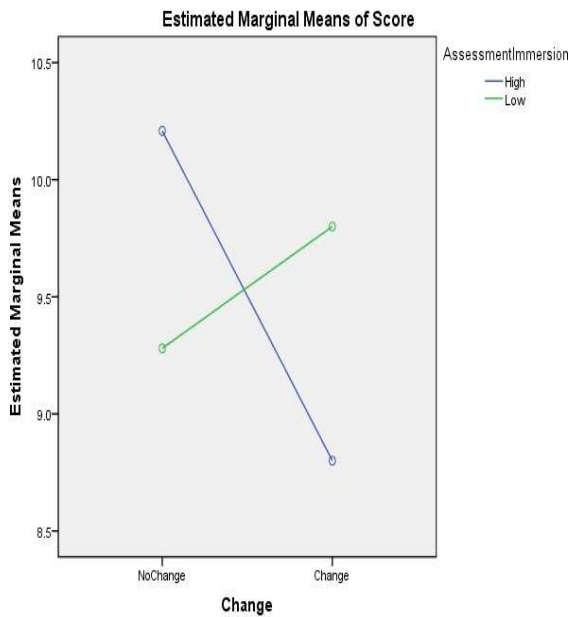


Figure 4.5: Assessment Score Group Means Comparison for Female

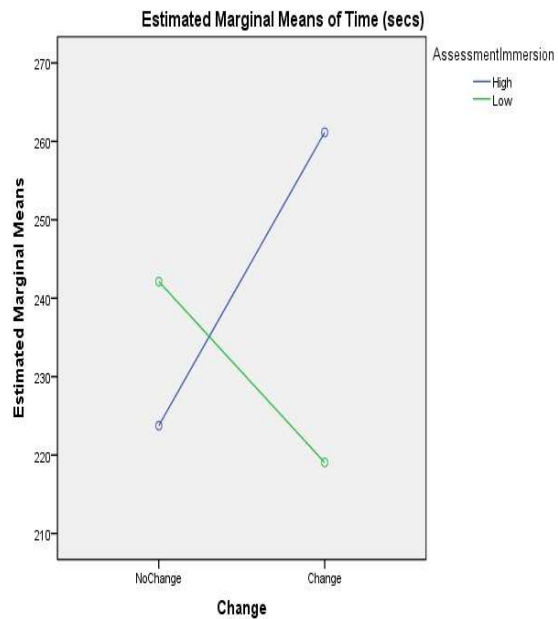


Figure 4.6: Assessment Time Group Means Comparison for Female

Chapter 5.

Discussion

Participants

The composition of the participant pool allows the study to be generalized to university-aged students with a relatively high familiarity with computers. This happens to be the target demographic for many educational tools in schools and in the workplace. It also aligns with the research demographics in most academic research on learning and memory. For the experimental tasks, users were randomly assigned to their groups and the composition of the experimental groups overall does not suggest any differences that would have interfered with the results. The factors considered from the data collected were: age, gender, GPA, familiarity with computers, and familiarity with 3D applications. The largest area of difference was the distribution of males and females between the groups with male participants scoring slightly higher on average than female participants. However, results from male subjects did not appear to affect the shape of the data or the interaction found.

Virtual Environment Immersion

The immersion questionnaire results support a key requirement of the study: that the 3D environment needs to be measurably more immersive than the 2D environment. Users reported a noticeable difference in the environments, and rated the 3D environment as more immersive by about two points on a 30-point scale. This finding confirms that the two environments present different levels of immersion, and reduces the possibility that a lack of environmental immersion affected the results.

Although the immersion questionnaire displayed a basic difference, the relatively weak immersion scores for the 3D environment were surprising. This might be attributed to a few factors. The questionnaire emphasizes the resemblance of the virtual world to the real world, while the experimental virtual environment used flat, coloured and abstract images instead of photo-realistic textures and models found in other 3D applications.

Immersion might also have been limited by reduced movement in the virtual world, where user movement was limited to camera control and the “next” and “previous” buttons as opposed to full movement (allowing the user to “walk” in the world). However, the immersive characteristics of the environment also had to be balanced with practical aspects of the study such as having a web application that could run on any lab computer and presenting a simple user interface. Possibilities for reducing or removing these immersion limitations will be discussed later.

Learning & Assessment Task

Subjects were able to complete the learning and assessment tasks with few issues. The participants' relatively high performance on average demonstrates that participants did not struggle with the interface or other aspects of the experiment. In addition, given that the average familiarity rating with 3D applications was relatively low, the comparable 3D environment scores and high 3D no-change group average score indicates that the novelty of navigating the 3D environment did not affect the score or assessment time outcomes. We can conclude that the web applications and user interfaces used in this study did not interfere with the results of the assessment.

Beyond participants being able to complete the assessments, the design of this study requires that the 2D and 3D assessments be equivalent in terms of difficulty. This is necessary to ensure that the results are not affected by the nature of the assessments themselves. For this reason, both assessments contained the exact same questions, format, and animation timing between questions. To confirm this requirement, the assessment scores and times were analyzed for the effect of assessment immersion. The ANOVA on both assessment scores and times provided evidence that assessment immersion alone did not have an effect on participants' performance. This shows that although the 3D assessment participants appeared to score slightly higher on average and users spent a little bit more time on it, there is not a significant difference in terms of performance between their performance and those of the 2D assessment participants. We can assume from these results that immersion in the assessment environment alone does not have a significant effect on performance as measured by assessment scores and time to complete.

The difficulty of the experimental task was a problematic aspect of the study, as the assessment scores and times indicate that the task was too easy for most participants. This is substantiated by a strong ceiling effect on the scores, with a high frequency of subjects scoring 12/12 and an average score of 9.86. The ceiling effect made identifying any effect more difficult. Steps to reduce the ceiling effect in future experiments will be discussed later. Despite the strong ceiling effect, it is consistent throughout all groups (they all completed the same training and assessment) and did not affect one group in particular. The effect was also less pronounced on assessment times, which were also used as a measure of performance.

Assessment Scores

When plotting the data, a possible interaction can be seen as the high-immersion assessment group scored higher on average in the no-change condition while the low-immersion assessment group scored higher in the change condition. This result would fall in line with the situational model updating hypothesis, which predicts that the switch from a lower to higher immersion environment would cause a greater increase in cognitive load. The lack of a significant difference between the 3D to 2D and 2D to 2D assessment groups also affirms this hypothesis, as the low cognitive load caused by processing the low immersion environment would predict little to no effect. These findings dispute the context dependence hypothesis in virtual environments, which predict the opposite: that the switch from a high-immersion to a low-immersion environment would result in the greatest performance degradation, due to the loss of environmental markers.

An ANOVA on the assessment scores showed that the interaction between the environment change and the direction of the change towards a low or high immersion environment, was not significant. The problems with the data can likely be attributed to the evident ceiling effect. Ultimately, the results from the assessment scores are inconclusive but hint at increased cognitive load depending on the direction of environment change (i.e., the situational model updating hypothesis).

Assessment Times

Although participants' assessment scores did not show a significant interaction, the times to complete the assessment provide an additional measure of performance. Neither the change nor the assessment immersion variables showed a significant effect on completion times, as found by an ANOVA. However, an interaction was discovered between whether a change in environment occurred and the level of immersion in the training and assessment environments. Subjects that trained in the 3D environment took less time to complete the assessment than subjects trained in the 2D environment. Furthermore, this effect appeared primarily when a change in environment occurred between the learning and assessment stages, as there was no significant effect of the 3D training environment found when no change in the immersiveness of the environment took place.

The degradation of assessment completion times found when changing from a low-immersion to a high-immersion environment accords with the situational model updating hypothesis. Situational model updating would expect an increase in cognitive load from the high-immersion environment, and a consequent increase in time to complete. The remaining results, particularly the finding of performance improvement when changing from a high-immersion to a low-immersion environment over staying in a low-immersion environment, does not provide evidence either for or against the situational model updating hypothesis.

Both results reported here are in conflict with the context dependence hypothesis. The low-immersion assessment change group improved where context dependence would have predicted degradation, and the high-immersion change group degraded where context dependence would have predicted no effect.

Research Questions

1. Does context affect learning in virtual environments?

The data suggests that like physical environments, virtual spaces can affect learning and memory performance. Learning, defined as the ability to memorize, recall, and apply internalized knowledge, was found to be affected by factors controlled in the environment, although the effect was not demonstrated as conclusively as hoped. A

significant interaction on assessment scores was not found, but the data hints at a possible effect as evidenced by the shape of the plotted results. Additionally, in relation to participants' times to complete the assessment, a significant interaction was found between the environment change and the immersion level of the assessment environment. This effect showed that users that trained in high-immersion environments tended to achieve faster completion times. The existence of any effect or interaction suggests that context has an effect, though neither the assessment environment or the environment change alone had any effect on the scores or times.

2. Does performance degrade when switching from one virtual environment to another?

This question cannot be conclusively answered with the data gathered in my study. Environment switching alone did not produce a significant effect on performance in this experiment. In both the assessment scores and times, the environment change variable was not found to have a significant effect. However, one experimental group was found to display the greatest degradation in both the assessment scores and completion times. The group that switched from a low-immersion environment to a high-immersion environment showed degraded scores and times when compared to the group that was trained and assessed in high-immersion environments. This interaction was confirmed statistically on assessment times, though it could be confirmed statistically on assessment scores. On the other hand, switching to a low-immersion environment was actually found to improve performance on assessment times, while showing no effect on assessment scores. This suggests that simply switching environments in virtual environments in between the learning and assessment tasks does not produce an effect. Rather the effect is dependent on the direction of the change from the learning environment.

3. Are performance effects dependent on the level of immersion in the environment?

Level of immersion was varied between the two environments, with every other aspect held constant. The differences in experience of immersion in the two environments were confirmed using questionnaire results. I was able to determine that the level of immersion alone (whether the assessment was delivered in 2D or 3D) did not have a significant effect on assessment scores or times. However, when switching from a low to high-immersion environment, subjects scored lower on average and required the most

time to complete. When switching from a high to low-immersion environment, subjects scored higher and required less time. Although the assessment scores did not show a significant effect, the assessment times did. We can therefore conclude that immersion was an important characteristic of the environment, and had an effect on performance in certain situations.

Theoretical Implications

Based on the findings of this study, I am not able to conclusively reject the null hypothesis. The ANOVA on assessment scores revealed no significant effect for the change groups, and no significant interaction between the change and assessment immersion variables. The ANOVA on assessment times revealed a significant interaction between the change and assessment immersion variables, indicating that whether subjects switched environments, together with the direction of the switch (towards a high or low immersion environment) had an effect on the scores. However, assessment times are a secondary measure of performance, and the nature of the interaction does not match the prediction of the situational model updating hypothesis. Although a significant effect was not attained, the findings of this experiment do present a case for further research.

The first reason a follow up study is needed is that the assessment scores generally accord with the situational model updating hypothesis. The collected scores show an observable performance drop in the high immersion assessment from the no-change to the change group, while the low immersion assessment did not show a significant difference between the no-change and change groups. This finding was partially confirmed with a t-test that detected a significant effect in the high immersion assessment groups, but not the low immersion assessment groups. Furthermore, the results suggest that immersion in virtual environments is not influenced by context dependence. The loss of an immersive learning environment did not result in performance degradation in any of the experimental groups. The potential impact on two well-established hypotheses in the learning and memory literature helps to justify further inquiry.

A likely factor in the lack of more significant results in this study was the prominent ceiling effect on the assessment scores. The large number participants scoring the

maximum possible grade on the quiz made identifying any experimental effects difficult. Luckily, this can be addressed in a future study by increasing the difficulty of the assessment task or the learning task. This can be done by increasing the complexity of the machine rules to be memorized, the number of assessment questions, and the complexity of the questions. A suggested strategy for removing the ceiling effect in future research will be presented in a later section. The fact that useful data was obtained despite interference from the ceiling effect suggests that an effect is likely to be found with better instrumentation.

Finally, the assessment times displayed a significant interaction between the assessment immersion variable and whether users switched environments. While the change variable alone did not show a significant effect, the interaction shown in the data suggests that the direction of the environment switch affected the completion times. Moving from low immersion to high immersion environments significantly increased the time required for participants to complete the assessment, and moving from high to low immersion decreased the time required. This interaction contradicts the context dependence hypothesis in because that hypothesis predicts the opposite results in each group: decreased performance in the high to low immersion group and no difference in the low to high immersion group. Conversely, situational model updating predicts the degraded performance when switching from the low immersion to high immersion environment. It does not, however, account for the improved performance in the low to high group. These results further support the proposition that immersion in virtual environments affects learning, and lends plausibility to the situational model updating hypothesis without decisively confirming it.

Although the results were not as conclusive as hoped, they shed some light on the underlying cognitive processes studied. They show very little support for context dependence in virtual environments that vary in immersion. With respect to both the assessment scores and times, the results appeared to contradict the context dependence predictions. This might be explained by the differences in immersion of the 2D and 3D elements not being sufficient environmental markers. It also suggests that the aspects of immersion used (3D space, animation, colour, and shapes) are not processed as markers and integrated with learned knowledge. Another possibility is that context dependence is

not triggered by changes in virtual environments and the mind does not process the change in the virtual world as an actual change of environment.

Limitations

Participant Group

The participants recruited for the study were university-aged students with relatively high familiarity with computers. This is one of the prime demographics targeted for training through 2D or 3D computer-based environments; but the results and conclusions cannot be generalized to younger children or older users with less exposure to computer-based learning environments.

Since a 64% majority of respondents being female, male subjects were slightly underrepresented in the data. Analysis of the results showed that gender distribution alone likely does not account for the results found in the assessment scores, but it also shows that male and female individuals performed differently on the learning and assessment tasks.

Language competence is another potential concern, since all the instructions provided for the training and assessment tasks was given in English only, and participants' primary language was not captured on the participant survey form. It is known that several students who participated in the study spoke English as a second language, however, because the students were recruited from a primarily English-speaking university, it is unlikely language competence alone could have made a large impact on the results.

Assessment Ceiling Effect

The most important limitation of this study was the ceiling effect or left skew found in the assessment score data. This result limited my ability to identify effects and draw decisive conclusions from the study. The cause of the ceiling effect was likely too low a level of difficulty of the learning task and assessment. The task was deliberately designed to be simple, and allow users to complete the experiment in a relatively short period of time. There was also early concern that the test could be too difficult, and a floor effect

would occur. Before data collection began, a pilot test was performed to test the tools and assessment questions. The results from the pilot, however, did not demonstrate the ceiling effect uncovered in the experiment results. Participants scored lower overall and even reported that the questions were very difficult. However, only eight subjects were recruited for the pilot and two rules of the final experiment had not yet been added: a cash prize for users who scored well, and a minimum training time to prevent users from progressing to the assessment without spending time training.

Tools and Measurements

The 3D web application satisfied the need for a 3D environment that was more immersive for the user than a 2D web page; but the experience falls short of the applications available to consumers today. Current technologies allow for nearly photo realistic graphics and even full virtual reality through the use of virtual reality glasses and natural gesture controls. In comparison, the tools used in this experiment represent only the basic requirements for an immersive environment (3D visuals, animations, and a user controlled camera). This is reflected in the immersion questionnaire results, which showed that participants found the 3D application only slightly more immersive than the 2D site (though the difference was statistically significant). It might be assumed that as environments become more immersive through the use of new technologies, effects on learning may increase in size.

Environment Changes

One concern with the interaction between the assessment immersion and change variables is that two of the four experimental groups shared same training environment; the high immersion assessment change group shared the same training environment as the low immersion assessment no-change group. With this arrangement, it can be difficult to identify if the performance change is due the training environment or due to the change itself, when the both the higher valued groups share the same training environment. This is apparent in the assessment times, as the high immersion assessment change group and low immersion assessment no-change group shared the same training environment.

To consider this issue, we can look at the assessment scores that show only degradation of the high immersion assessment change group and little change in the low immersion assessment group. This suggests that, on assessment scores, the learning environment did not appear to make a difference without a change of environment occurring as well, and thus that cognitive load can be affected by environmental changes and not the training environment alone. However, the effect of the learning environment on assessment times also raises interesting questions with regard to how the training environment affects memory during an assessment task. Because the assessment environment has been shown to have no effect through analysis on both assessment scores and times, the learning environment can be assumed to have an effect. This does not contradict the concept of situational model updating. The cognitive processes involved when reacting to a new level of immersion, along with the direction of the change, remains the key variable affecting performance.

Learning Task

The learning task chosen challenged participants with recognition, recall, and problem solving tasks. However the findings cannot be generalized to all types of learning tasks. For instance, the findings of this experiment may not generalize to context effects on the learning of physical actions, playing an instrument, or social interactions.

Experiment Environment and Situation

Further, the situation presented by the experimental environment is different from the experience many learners would have in a real-world learning situation. The experiment presented the assessment task immediately after the learning task, without any delay in between. These conditions do not resemble much real world training, which takes place over longer periods of time. It does, however, provide an idea of how short-term storage of information is affected by changes in immersiveness.

In this experiment, participants were paid \$5 to complete the experiment, and promised that three of the top scoring participants would receive \$50. This reward system

differs from most learning situations, in that rewards are provided simply for participating, and participants who feel they are not likely to score in the top three might not put forward much effort. Ultimately, the effects of this reward scheme on the findings of the experiment cannot be known for certain.

Context Dependence Hypothesis

Although the findings do not demonstrate the effects predicted by the context dependence hypothesis, it cannot be concluded that context dependence does not occur in virtual spaces. It is possible that both the 3D and 2D environments may have lacked the necessary environmental markers to produce a context dependence effect. It can only be concluded from this study that the factors that were present within the 3D environment (e.g. 3D space and colours, and shading on the walls and floor) were not sufficient to produce a context dependence effect.

Implications for Practice

Although the null hypothesis was not decisively rejected in this study, the results most closely accord with the predictions of the situational model updating hypothesis. A follow up study is highly recommended to better understand this effect, which could substantially change how current research and practice are performed with virtual learning environments.

Learning and memory experiments in virtual environments

This study sought to examine the effects of the immersiveness of virtual environments on learners' short-term memory performance. If not controlled, the effect of processing a new virtual environment may be enough to interfere with task performance. The literature suggests that the use of mixed environment training and assessment (ie. training in immersive 3D, followed by paper-and-pencil assessment) is still common, particularly in the study of learning and video games. It may not be a coincidence that existing studies on the effectiveness of virtual reality games for learning have seen mixed results. Just as physical spaces are carefully controlled in an experiment, virtual spaces must also be carefully controlled.

Another finding in this experiment was that the direction of change in immersiveness can alter memory performance. Moving from a low to high environment produced the greatest degradation, while switching to (or staying on) a low immersion environment provided the most consistent results. So, if consistent results are important across several different training methods, a low immersion application or tool would provide the most stable assessment environment.

Training and assessment in the field

Teachers, trainers, designers, and developers in educational fields should also be aware of the effects of differences in immersiveness in virtual contexts. My findings suggest that when designing an assessment, the nature of the learning and assessment environments can influence the outcomes. Highly immersive assessment environments must be carefully considered, because they can impose a greater cognitive load on the learner and result in lower scores. Simple, low immersion, 2D assessments can provide more consistent scores. Users switching from a high immersion environment show no more degradation than users trained in the same environment. However, high immersion environments may sometimes offer a more robust assessment because it better represents the performance environment, in which case the user's training environment becomes much more important. When training or learning new knowledge, context still matters. My findings show that learners trained in a high immersion environment do not suffer from additional cognitive load when assessed in another high immersion environment. They also do not experience increased cognitive load when moving to a less immersive environment for assessment, and may even benefit from this change.

Future Research

Learning Tasks

The first issue that should be addressed in future studies is the ceiling effect found in this experiment as a result of the low assessment difficulty. As a follow up, the number of functions to memorize, the complexity of the functions, or the number of questions in

the assessment can be modified to increase the difficulty of the assessment and the variability of the data. This simple change is likely to reduce the ceiling effect, and provide a better chance of significant results. An updated learning and assessment task is provided in the Appendix that increases the complexity of the machine rules and the number of questions in the assessment task. The new rules and assessment can be implemented while reusing the basic structure of this experiment.

Further study could expand on the types of tasks examined in virtual learning environments, and represent more variations in human learning. Names, shapes, colours, and mathematical and logical functions, such as those examined in this study, cover only a small portion of human learning. Other areas of interest might include scientific concepts, reading comprehension, or social interactions. These and other forms of learning can be incorporated into virtual learning tasks, and many are already integrated in modern video games and simulations.

Environment Change Timing

In this study, the environment switch between training and assessment occurred immediately, with roughly a 2 second. How the switch between environments happens in the physical world has been shown to change how learning and memory is affected (Radvansky and Copeland, 2000) and this is likely true in virtual environments as well. Modifications to the experimental procedure can be introduced by including a time delay between the learning and assessment tasks, requiring participants to perform an unrelated task before starting the assessment (ie. a memory wash), or even adding an additional environment switch in the learning task, before the assessment (e.g. learning in 2D and 3D and assessing in 2D). These variations on the experiment may afford a better understanding of context effects on both long term and short term memory.

Environment Changes

One aspect not examined in this study is the change of visible features within two environments with equivalent levels of immersion. The data provide evidence that controlling immersion has an effect on learning and memory, however other characteristics of the environment have yet to be examined. For example, an effect

supporting the context dependence was not found in the 2D condition. Immersion alone might not be sufficient to produce an effect or some other aspect altogether might be required. Environments can be controlled for other factors such as colour, environmental markers in the form of decoration, or layout (e.g. navigating right to left instead of left to right).

Another environment change that could have a notable effect in practice is the switch from a virtual environment to a physical one. Such a study could have a substantial impact on how current electronic training applications are designed and administered. Many electronic training programs are specifically aimed at providing real world advantages for learners.

Immersive Virtual Environments

Lastly, the emergence of technologies to create more highly immersive environments at the time of this report opens up many needs and opportunities for research. Products like Oculus Rift and Google Glass offer rich, interactive virtual environments far more immersive than what was used in this experiment, for lower costs than ever before. The use of these kinds of virtual environments for training are inevitable, though their effects on human learning and memory are not well understood. If the environment used in this study was able to produce a small effect, the greater sense of immersion created by an Oculus Rift might have a markedly larger effect. Future experiments can study these technologies and their impact on the field of education and understanding of human learning.

References

- Anderson, J. R., Reder, L. M., Simon, H. A. (1996). *Situated learning and education. Educational Researcher*, 25(4), 5-11.
- Barfield, W., Hendrix, C. (1995) *The effect of update rate on the sense of presence within virtual environments. Virtual Reality*, 1(1), 3-16.
- Barnett, S. M., Ceci, S. J. (2002) *When and where do we apply what we learn? a taxonomy for far transfer. Psychological Bulletin*, 128(4), 612-637.
- Beale, I., Kato, P., Marin-Bowling, V., Guthrie, N., Cole, S. (2007) *Improvement in Cancer-Related Knowledge Following Use of a Psychoeducational Video Game for Adolescents and Young Adults with Cancer. Journal of Adolescent Health*, 41(3), 263-270.
- Bransford, J. D., Schwartz, D. L., (1999) *Rethinking transfer: a simple proposal with multiple implications. Review of Research in Education*, 24(61), 61-100.
- Bystrom, K., Barfield, W. (1999) *Collaborative task performance for learning using a virtual environment. Presence*, 8(4), 435-448.
- Cobb, P., Bowers, J. (1999). *Cognitive and situated learning perspectives in theory and practice. Educational Researcher*, 28(2), 4-15.
- Eich, E. (1985). *Context, memory, and integrated item/context imagery. Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(4), 764-770.
- Glenberg, A. M., Meyer, M., Lindem, K. (1987). *Mental models contribute to foregrounding during text comprehension. Journal of Memory and Language*, 26(1), 69-83.
- Greeno, J. G. (1998). *The situativity of knowing, learning, and research. American Psychologist*, 53(1), 5-26.
- Kolb, D. (2000). *Learning places: building dwelling thinking online. Journal of Philosophy Education*, 34(1), 121-133.
- Kurby, C. A., Zacks, J. M. (2008) *Segmentation in the perception and memory of events. Trends in Cognitive Science*, 12(2), 72-79.
- Murnane, K., Malmberg, K., Phelps, M. P., (1999) *Context-dependent recognition memory: the ice theory. Journal of Experimental Psychology*, 128(4), 403-415.

- Pspotka, J. (1995) *Immersive training systems: virtual reality and education and training*. Instructional Science, 23(5-6), 405-431.
- Radvansky, G. A. (2006) *Walking through doorways causes forgetting: situation models and experienced space*. Memory & Cognition, 34(5), 1150-1156.
- Radvansky, G. A., Krawietz, S. A., Tamplin, A. K. (2011) *Walking through doorways causes forgetting: further explorations*. The Quarterly Journal of Experimental Psychology, 64(8), 1632-1645.
- Regenbrecht, H., Schubert, T. (2002) *Real and illusory interactions enhance presence in virtual environments*. Presence, 11(4), 425-434.
- Sahakyan, L., Kelley, C. M. (2002) *A contextual change account of the directed forgetting effect*. Journal of Experimental Psychology: Learning, Memory, Cognition, 28(6), 1064-1072.
- Schrader, C., Bastiaens, T. J. (2011) *The influence of virtual presence: effects on experienced cognitive load and learning outcomes in educational computer games*. Computers in Human Behavior, 28(2), 648-658.
- Slater, M., Linakis, V., Usoh, M., Kooper, R. (1996) *Immersion, presence, and performance in virtual environments: an experiment with tri-dimensional chess*. ACM Virtual Reality Software and Technology, 163-172.
- Smith, S. (1986) *Environmental context-dependent recognition memory using a short-term memory task for input*. Memory & Cognition, 14(4), 347-354.
- Smith, S., Glenberg, A., Bjork, R. A. (1978) *Environmental context and human memory*. Memory & Cognition, 6(4), 342-353.
- Spencer, M. R., Weisberg, R. W. (1986) *Context-dependent effects on analogical transfer*. Memory & Cognition, 14(5), 442-449.
- Swallow, K. M., Zacks, J. M., Abrams, R. A. (2009) *Event boundaries in perception affect memory encoding and uploading*. Journal of Experimental Psychology, 138(2), 236-257.
- Thomas, H. (2010) *Learning spaces, learning environments and the dis'placement' of learning*. British Journal of Educational Technology, 41(3), 502-511.
- Verhagen, T., Feldberg, F., Van den Hooff, B., Meents, S., Merikivi, J. (2012) *Understanding users' motivations to engage in virtual worlds: a multipurpose model and empirical testing*. Computers in Human Behavior, 28(2), 484-495.
- Wahlstedt, A., Pekkola, S., Niemela, M. (2008) *From e-learning space to e-learning place*. British Journal of Educational Technology, 39(6), 1020-1030.

Yeh, W., Barsalou, L. W. (2006) *The situated nature of concepts*. American Journal of Psychology, 119(3), 349-384.




Appendix A.


Updated Learning and Assessment Tasks

Due to the ease of the learning and assessment task, the data gathered for this study displayed a ceiling effect on assessment scores that hindered our ability to identify significant effects. The following changes are proposed to increase the difficulty of the assessment task for any future experiments.

Machine Rules

The trained rules have been updated to increase the number of steps for three of the four machines. The new steps have been bolded below.

Name	Symbol	Steps
Adder		<ol style="list-style-type: none">1. Add all the digits together2. Add 100
Exchanger		<ol style="list-style-type: none">1. Order the digits in the number from lowest to highest2. Swap the first and last digits
Dicer		<ol style="list-style-type: none">1. Multiply the input by 102. Divide the number by 23. Drop the last digit

Jumbler		<ol style="list-style-type: none"> 1. Multiply by 2 2. Swap the first and second digits 3. Reverse the order of the digits
---------	---	--

Quiz Questions

To increase the difficulty of the quiz, several changes have been made. The new rules are considered that have added more steps to recall. An additional section has also been added to for questions that require users to combine the rules of two machines together. The total score of the assessment is now raised to 16. Finally, each multiple choice question now has 6 options instead of 5, reducing the effectiveness of guessing at the answers.

These changes aim to make perfect scores occur less frequently and lower the central tendency of the assessment scores. The additional questions should also increase the times to complete.

1) Recall the machines

For each of the following symbols, select the correct name of the machine

1.1)



Adder, Exchanger, Dicer, Jumbler

1.2)



Adder, Exchanger, Dicer, Jumbler

1.3)



Adder, Exchanger, Dicer, Jumbler

1.4)



Adder, Exchanger, Dicer, Jumbler

2) Recall the functions

For each of the following questions, look at the given scenario and select the best answer.

2.1) 49 > dicer > ____

245, 13, 94, 24, 130, none of the above

2.2) 82 > adder > ____

410, 10, 110, 20, 164, none of the above

2.3) 421 > jumbler > ____

642, 124, 842, 284, 14 none of the above

2.4) 618 > exchanger > ____

903, 309, 861, 168, 1236, none of the above

2.5) 54 > jumbler > ____

81, 45, 801, 246, 108, none of the above

2.6) 127 > dicer > ____

635, 721, 63, 254, 452, none of the above

2.7) 379 > adder > ____

1895, 180, 973, 758, 18, none of the above

2.8) 294 > exchanger > ____

249, 15, 147, 942, none of the above

3) Combine the functions

For each of the following questions, look at the given scenario and select the best answer.

3.1) 253 > dicer > exchanger > ____

126, 621, 1100, 609, 352, none of the above

3.2) 830 > jumbler > adder > ____

616, 166, 415, 113, 210, none of the above

3.3) 29 > adder > dicer > _____

111, 1110, 55, 44, 33, none of the above

3.4) 687 > jumbler > exchanger > _____

1347, 137, 7430, 876, 1365, none of the above

Appendix B.

Immersion Questionnaire

1. How strong was your sense of presence, in the virtual environment?
2. How strong was your sense of "being there", in the virtual environment?
3. How strong was your sense of inclusion, in the virtual environment?
4. How aware were you of the real world surroundings while navigating in the virtual world (i.e. sounds, room temperature, other people, etc.)?
5. With what degree of ease were you able to navigate within the virtual environment?
6. Do you feel that you could have reached into the virtual world and grasped an object?
7. How interactive was the virtual environment?
8. What was your overall comfort level in navigating throughout this environment?