THE EFFECTS OF COMMUNICATION MODALITY ON PRESENCE, COGNITIVE LOAD AND RETENTION IN SECOND LIFE

BY

STEPHANY FILIMON WILKES

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Technical Communication in the Graduate College of the Illinois Institute of Technology

Approved ________________________
Advisor

Chicago, Illinois
December 2009
© Copyright by

STEPHANY FILIMON WILKES

December 2009
ACKNOWLEDGMENT

This dissertation is dedicated to the memory of Maryanne Wilkes (August 27, 1944–May 20, 2009) and to Ian Wilkes: There is no better reason to have written this dissertation in absentia.

I am grateful to my adviser, Professor Susan Feinberg, and to my committee members, Professors Roya Ayman, Kathryn Riley, Jack Snapper and Karl Stolley. They provided exceptional patience, support, and a willingness to work with me, a student who not only worked full time during graduate school but lived more than 3,000 miles from Chicago during the majority of my dissertation work. Each of these individuals made themselves available both online and off, and I believe my education would not have been accessible or attainable without them.

Many individuals recruited study participants on my behalf. If not for James Braman, Vivian Caputo, Strata Rose Chalup, Jeska Dzwigalski, Susan Feinberg, Stephen Filimon, Sarah Hutchinson, Gayathri Kamath, Noel and Kathryn Kelly, Robert and Stephany Luszczak, Kathryn Riley, and Karl Stolley I would still be reading screening questionnaires.

I would have given up this solitary pursuit long ago if not for the encouraging spirit of the Shady Ladies of Fire Island Pines, namely Dan Delany, Patrick Higgs, Cory Krebsbach, and Dan Michel. Brad Nem er entertained a great many dissertation ideas and offered valuable critical input on every single one of them. Aleks Krotoski has been a true sister in the struggle and known sympathizer, and every Thursday, Sandi Esquivel, Kate Hoag, Dana Lamb, Farah Madhat, Radha Sharma, and Kate Tillery got me out of the house to think about something else.

I am indebted (literally and figuratively) to the Illinois Institute of Technology for making my education accessible. Their very practical benefits—the IIT shuttle
service, the addition of Global Grounds coffee, the Galvin Library EZProxy service, and a generous supply of evening classes—enabled me to complete my Ph.D. while working fulltime. After a long day at work, a cup of coffee before a three-hour evening class makes more difference than anything else. I had the good fortune to work for employers who offered tuition assistance or reimbursement and flexible work schedules: Thank you, Feeding America and Linden Lab, and especially Gabi Fitz and Bethanye McKinney Blount. Likewise, I thank the Technical Communication program at IIT for reimbursing a portion of my study participant compensation costs through early 2009.

Professors Michael Barry, Elaine Carey, Hugh Culik, and Nick Rombes first planted the graduate school seed and, more importantly, convinced me such a thing was possible.

None of this would have happened if not for my family, most especially Robert and Stephany Lusczak, Noel and Kathryn Kelly, Stephen Filimon, and Nicholas Filimon, each of whom has instilled and cultivated in me curiosity, a love of learning, and the importance of formal education. I happily acquired new family members during this journey: I am grateful to Tim Wilkes, for allowing me to hole up in Calistoga and with my laptop open, and to Michael Jon Watt for support and encouragement.

Rebecca Przybycien constantly reminds me of what is (and is not) important by virtue of her existence. Jeska Dzwigalski has been with me for every day of my higher education journey since 1995, nearly half my life. This pursuit was a little less solitary for having her along the way. Ian Wilkes brought me great joy and perspective, and has demonstrated exceptional patience and support. I am very fortunate.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENT</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>vii</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1. The Research Problem</td>
<td>4</td>
</tr>
<tr>
<td>2. REVIEW OF THE LITERATURE</td>
<td>10</td>
</tr>
<tr>
<td>2.1. A Brief History of Virtual Environments</td>
<td>10</td>
</tr>
<tr>
<td>2.2. Theories of Working Memory</td>
<td>14</td>
</tr>
<tr>
<td>2.3. Modality and Learning</td>
<td>33</td>
</tr>
<tr>
<td>2.4. Presence and Co-Presence</td>
<td>51</td>
</tr>
<tr>
<td>3. METHODS</td>
<td>60</td>
</tr>
<tr>
<td>3.1. Hypotheses</td>
<td>60</td>
</tr>
<tr>
<td>3.2. Research Design</td>
<td>61</td>
</tr>
<tr>
<td>3.3. Study Procedures</td>
<td>63</td>
</tr>
<tr>
<td>3.4. Instrumentation and Materials</td>
<td>85</td>
</tr>
<tr>
<td>4. RESULTS</td>
<td>90</td>
</tr>
<tr>
<td>4.1. Proposed Analyses</td>
<td>93</td>
</tr>
<tr>
<td>5. DISCUSSION</td>
<td>119</td>
</tr>
<tr>
<td>5.1. Contributions</td>
<td>119</td>
</tr>
<tr>
<td>5.2. Review of Results</td>
<td>120</td>
</tr>
<tr>
<td>5.3. Limitations</td>
<td>128</td>
</tr>
<tr>
<td>6. GLOSSARY OF KEY TERMS</td>
<td>132</td>
</tr>
<tr>
<td>APPENDIX</td>
<td>133</td>
</tr>
<tr>
<td>A. STUDY INFORMATION WEB PAGE</td>
<td>134</td>
</tr>
<tr>
<td>B. CONSENT FORM</td>
<td>139</td>
</tr>
<tr>
<td>C. SCREENING QUESTIONNAIRE</td>
<td>143</td>
</tr>
</tbody>
</table>
D. STUDY ACCEPTANCE AND REMINDER EMAIL MESSAGES 148
E. INSTRUCTOR SCRIPT FOR INTRODUCTION TO BUILDING CLASS 152
F. POST-ACTIVITY PRESENCE QUESTIONNAIRE 160
G. ADDITIONAL PRESENCE ITEMS FROM GERHARD 164
H. COGNITIVE LOAD MEASURES 166
I. RETENTION MEASURES 168
BIBLIOGRAPHY 182
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Comparison of Penney’s Two Models of Short-term Memory: Traditional and Proceduralist</td>
<td>27</td>
</tr>
<tr>
<td>2.2</td>
<td>Characteristics of Auditory and Visual Streams</td>
<td>28</td>
</tr>
<tr>
<td>2.3</td>
<td>Group Conditions and Assignments from Experiment #1 by Mousavi et al.</td>
<td>35</td>
</tr>
<tr>
<td>2.4</td>
<td>Group Conditions and Numbers of Participants from Experiment #2 by Mayer and Moreno</td>
<td>46</td>
</tr>
<tr>
<td>2.5</td>
<td>Group Conditions and Numbers of Participants from Experiment #3A by Moreno and Mayer</td>
<td>47</td>
</tr>
<tr>
<td>2.6</td>
<td>Group Conditions and Numbers of Participants from Experiment #3B by Moreno and Mayer</td>
<td>48</td>
</tr>
<tr>
<td>3.1</td>
<td>Reasons for Elimination of Prospective Study Participants to and Final Sample Size (N=60)</td>
<td>73</td>
</tr>
<tr>
<td>3.2</td>
<td>Chi-square Scores for Assignment to Test Condition based on Age, Gender and Race</td>
<td>75</td>
</tr>
<tr>
<td>3.3</td>
<td>Study Participants by Age Group (N = 60)</td>
<td>80</td>
</tr>
<tr>
<td>3.4</td>
<td>Study Participants by Compressed Age Group (N = 60)</td>
<td>81</td>
</tr>
<tr>
<td>3.5</td>
<td>Age of Second Life user account (N = 60) (Experience)</td>
<td>82</td>
</tr>
<tr>
<td>3.6</td>
<td>Recoded Age of Second Life user account (N = 60) (Experience)</td>
<td>82</td>
</tr>
<tr>
<td>3.7</td>
<td>Total Time Spent in Second Life Prior to Study Completion (N = 60) (Time)</td>
<td>82</td>
</tr>
<tr>
<td>3.8</td>
<td>Recoded Total Time Spent in Second Life Prior to Study Completion (N = 60) (Time)</td>
<td>83</td>
</tr>
<tr>
<td>3.9</td>
<td>Frequencies of Primary Interest in Second Life (N = 60)</td>
<td>83</td>
</tr>
<tr>
<td>3.10</td>
<td>Cronbach’s alphas for Presence, Cognitive Load and Retention Scale Reliability</td>
<td>89</td>
</tr>
<tr>
<td>4.1</td>
<td>Correlation matrix for Independent, Dependent and Demographic Variables (N=60)</td>
<td>92</td>
</tr>
</tbody>
</table>
ABSTRACT

This thesis reports findings from a study (N = 60) of the impact of three communication modalities (voice only, text only, and voice and text simultaneously) on cognitive load, as measured by subjective reports of mental effort; on learning, as measured by tests of recall and retention; and on perceptions of presence as measured by a Presence Questionnaire (Witmer & Singer, 2005). Based on the results of prior empirical research, it is hypothesized that retention scores will be higher for voice participants and voice-and-text participants than for text-only participants; that cognitive load will be lower for voice participants and higher for text conditions; that voice will contribute to greater perceptions to presence; and that higher perceptions of presence will not correlate with deeper learning. Study results indicate that communication modality significantly effected cognitive load ($F(2, 54) = 4.58, p = .01$) and retention ($F(2, 54) = 3.53, p = .04$), and that experience with and time spent in the virtual environment had significant effects on measures of cognitive load, retention, and presence: Significant between-subjects effects were found for cognitive load and time ($p = .23$), for retention and time ($p = .21$), and for retention and experience ($p = .03$).
CHAPTER 1
INTRODUCTION

This dissertation describes how communication modality affected cognitive load, learning (recall and retention), and perceptions of presence (one’s sense of being there) and co-presence (one’s sense of being there with others) in Second Life, a social virtual world. To this end, this dissertation empirically evaluated the impact of three communication modalities (voice only, text only, and voice and text simultaneously) on cognitive load, as measured by subjective reports of mental effort; on learning, as measured by tests of recall and retention; and on participants’ perceptions of presence, as measured by a Presence Questionnaire (Witmer & Singer, 2005). While based on a small sample (N = 60) and limited context, the study described here resulted in unexpected findings about both cognitive load and immersion in a social virtual world.

Today, social virtual worlds are online media used to provide distance education, academic and professional conferences, art events, and more. According to the Chronicle of Higher Education, professors use Second Life to hold distance-education classes, anthropologists and sociologists use it as a laboratory for studying human behavior, and architects use it as a canvas on which to explore design (Foster, 2007). The New Media Consortium (NMC), for example, is a nonprofit organization comprised of approximately 250 learning-focused organizations. The NMC Campus in Second Life includes NMC member organizations (primarily universities) and their faculty. Universities on campus in Second Life include the Massachusetts Institute of Technology (MIT), Princeton University, the University of Michigan, and many others (NMC, 2009).

The idea that instruction in social virtual worlds is inherently more immersive and engaging, and thus better, than web-based or other multimedia delivery methods
can be an easy one to fall into. This attitude is known as the media-affects-learning hypothesis, used during the 20th century to integrate newer technologies in education and based on nothing other than the idea that new technologies must be better (Moreno, 2006). Claims like this have been made despite empirical work that supports the method-affects-learning hypothesis, which indicates that it is the method rather than the media that impacts learning. Some of the more compelling reasons for developing highly immersive environments are based on cognitive load theory: By making interactions with technology more natural, immersive learning environments may result in the reduction of cognitive load. More immersive environments will supposedly drive students’ limited attention to learning content rather than the interface (Mayer & Moreno, 2002).

In regard to virtual environments specifically, claims have been made that presence is causally related to task performance and learning (Schuermie et al., 2001) and that many of the factors that appear to affect presence are known to enhance learning and performance (Witmer & Singer, 1998). Biocca et al. (2003) note that communication and human-computer interaction (HCI) researchers are typically interested in social presence because it may mediate the effects of other variables of central concern to the researcher, including interface features, learning and memory (Biocca et al., 2003). There is, however, no conclusive empirical evidence to support these claims. Whether or not presence contributes to better task performance remains controversial based on reported findings (Schuermie et al., 2001).

In contrast to the lack of conclusive evidence on presence and learning, a wealth of empirical research exists on communication modality and learning. In addition, some relationships between modality and learning are well established. Moreno and Mayer’s modality principle describes one such empirically established relationship between communication modality and learning and was a guiding one for the study
presented here. The modality principle holds that words should be presented audito-
ri ly (spoken or narrated) rather than visually (written) when presenting a multimedia 
 explanation, i.e. a combination of visual information like pictures with corresponding 
 verbal information (Moreno & Mayer, 2002). The work of Moreno and Mayer can 
 and should be extended in ways that can improve our understanding of teaching and 
 learning in social virtual worlds, in cooperative contexts, and with voice and text 
 modalities simultaneously (Moreno, 2006).

First, although a great amount of empirically valid research exists on modality 
 and learning, there is scant research on communication modality as a contributing 
 factor to presence and co-presence. Much research has been devoted to identifying 
 specific factors that contribute to presence (Schuemie et al., 2001) but, in a list of 
 25 interaction factors empirically evaluated for their affect on presence, modality was 
 not among them. The omission of modality is noteworthy because co-presence has 
 been shown to correlate with presence (Schuemie et al., 2001), and because communi-
 cation modality has been shown to affect co-presence and thus, potentially, presence. 
 Sallnas found that, when measuring co-presence with and without voice, audio co-
 presence overshadowed the visual co-presence of other participants, an indication that 
 immersive desktop social virtual worlds with voice may provide more immersive co-
 presence than those without (Sallnas, 2004). As Schroeder notes, ”We know little 
 apart from Sallnas’ work on the weight of different modalities in conveying the co-
 presence of the other person” (Schroeder, 2006). The study presented here compared 
 three communication modalities and evaluated measures of presence and co-presence 
 from participants assigned to each modality.

Second, presence and co-presence are both factors present in social virtual 
 worlds. The more lifelike a virtual world is the more successful it is assumed to be. 
 This is a debatable point, however, and the impact of presence and co-presence on
learning is questionable and not well understood. Social virtual worlds are actively used for distance education and other forms of training: Educators should have an interest in improving the understanding of presence, co-presence and any relationship between these and learning.

In summary, communication modality alone has been shown to affect learning (Moreno, 2006; Moreno & Mayer, 2003; Penney, 1989), and some research has shown that it also affects measures of co-presence (Sallnas, 2004). The primary purpose of this dissertation was to extend research on communication modality, cognitive load, learning and presence within a publicly available, social virtual world already in common use; to determine if communication modality affects cognitive load, perceptions of presence and/or learning measured via recall and retention; and to better understand the relationships between each of these factors and their impact, if any, on learning. The secondary purpose of this dissertation was to evaluate simultaneous text and voice communication (a previously untested condition in the work of Moreno and Mayer until 2003) and to collect data from a social environment rather than one comprised of a single multimedia instructional tool and one student, as tested in the work of Moreno and Mayer (2002).

1.1 The Research Problem

This thesis seeks to extend empirically established claims and to test weak and unproven claims in order to fill in some of the knowledge gaps about learning in social virtual worlds. The review of the literature, which follows this section, outlines the specific claims and theories that provide the foundation for the study described later in this document, and elaborates on knowledge gaps and points of departure noted in this Introduction.

Moreno and Mayer’s (2002) modality principal, an empirically established and
sound theory, provided the foundation for the study presented here. The modality principle holds that words should be presented auditorily (spoken or narrated) rather than visually (written) when presenting a “multimedia explanation,” i.e. a combination of visual information like pictures with corresponding verbal information. In the study described here, participants were presented with a multimedia explanation in the form of an interactive class called Introduction to Building in Second Life, conducted entirely in Second Life. Participants learned how to build basic 3D objects (culminating with each participant building their own chair) by stretching, turning, moving and linking together more primitive 3D shapes. This activity served as the “visual information” part of the modality principle. Each class was conducted in one of three possible modalities (voice only, text only, or voice and text simultaneously), which provided the “corresponding verbal information” and “words” part of the modality principle.

Moreno and Mayer’s (2002) modality principle owes much to two preceding theories: Sweller’s theory of cognitive load and Penney’s separate streams hypothesis. Both are discussed later in this dissertation but, briefly, Sweller’s cognitive load theory describes a working memory of limited capacity, with partially independent subcomponents that deal with 1) auditory or verbal material and 2) visual 2- or 3-dimensional information (Sweller et al., 1998).

Penney’s separate streams hypothesis is based on modality effects observed during research on verbal learning and short-term (working) memory, which indicated that auditory presentation almost always results in higher recall than does visual presentation (Penney, 1989). Penney’s hypothesis describes how written and spoken words are encoded differently, and how acoustic (A) code persists and boosts recall of recent auditory (spoken or heard) items relative to visual items in short-term memory tasks, thus producing the modality effect. Moreno and Mayer (2002) acknowledge that
their modality principle is comparable with the one offered by Penney, which they have included in their cognitive-affective theory of learning with media (CATLM).

The CATLM is a theoretical summary in that it is comprised of a list of seven assumptions compiled from cognitive research by investigators other than Moreno and Mayer (2002). Despite this summary quality, the CATLM is both useful and convenient because it has done the hard work of combining empirically sound cognitive principles into a single, cohesive theory that can be tested whole.

For the purposes of the study presented here, which sought to understand differences between communication modalities and their impact, the modality principle that Moreno and Mayer (2002) derived from the broader CATLM is more relevant and interesting. The modality principle is derived from three of the seven total CATLM assumptions: 1) independent information processing channels, 2) limited working memory capacity and 3) dual coding. The rationale underlying the modality principle is that:

By using the auditory channel to process the words, effective working memory capacity is expanded because students are not forced to split their limited visual working memories between the on-screen text and the pictorial information. Pictures are processed through the visual information channel, while spoken words are processed through the auditory channel. Processing of words is off-loaded onto the auditory channel, which is otherwise underused (Moreno, 2006).

In contrast to the modality principle, the information-equivalency hypothesis predicts no differences between groups that have concurrent auditory narration presented with corresponding visual diagrams, and groups that have printed text presented concurrently with corresponding visual diagrams (Mayer & Moreno, 1998). The hypotheses and results presented later in this document oppose the information equivalency hypothesis, which is acknowledged here to ensure contrasting theories and perspectives have been identified and considered.
The study presented here aimed to extend and test the modality principle by:

- Examining the modality principle in the context of a social virtual world, a co-operative learning context with multiple learners rather than the more common context of a single instructor with a single learner using a website. Social communication between two or more learners was possible in the study presented here, and social communication may compete with or supplement communication from the instructor. In addition, the availability of multiple communicators may impact cognitive load and overload one working memory channel more than another. Further, information from multiple communicators could overload either or both channels (visual or auditory), depending on the modality to which participants are assigned: The visual presence of avatars in the surrounding 3D space and a visual (text) IM could increase the load of the visual channel, while hearing the voices that belong to multiple avatars could increase the load of the auditory channel.

- Testing two modalities simultaneously by testing voice and text communication together. Previously, voice narration and printed text have been tested in isolation but not often at the same time. Moreno and Mayer (2002) tested a concurrent narration and text condition (Group NT) but did not do so in a social context. Voice and text together, and simultaneously, was a condition tested here because it is an increasingly common scenario in social virtual worlds that deserves to be tested as a third or new modality.

- Testing the modality principal and learning in a publicly available (i.e. not a laboratory) environment. Although this study, by virtue of its being a research study, necessarily created a test environment (i.e. comprised of screened and approved test participants, in a private area of Second Life, undergoing an experimental activity), it still took place in Second Life, a popular and commonly
available social virtual world, not a virtual reality lab on a single campus. In addition, participants logged in to Second Life from private locations using their usual computers (either desktops or laptops). This meant the modality principle was tested in a new type of multimedia environment, a widely available social virtual world.

This study also tested some more tenuous and less conclusive claims, namely that presence affects learning and task performance (Scheunie et al., 2001). After the Introduction to Building class, participants from each of the three test groups (Group V, Group T, and Group VT) reported their sense of presence and co-presence using a Likert-scale style questionnaire. The following is one example question from the Presence Questionnaire: “How natural did your interactions with the environment seem?” The question was followed with a range of seven Likert-style options, beginning with “Not Natural”, having a midpoint of “Moderately Natural,” and ending with “Extremely Natural.”

After each participant completed these questions and tasks, data were analyzed for evidence of any relationships (intercorrelations) between not only modality and learning outcomes but participants’ presence scores and learning outcomes. The Witmer and Singer Presence Questionnaire was used for part of the study because it is considered valid (Schuenie et al., 2001) and was the instrument used by Moreno in modality work, which may make these findings slightly more comparable than they may otherwise be. This method follows Welch’s (1999) suggestions for measuring presence and learning: Obtaining quantitative measures of each for the same subjects and then determining if, or to what extent, the two measures are correlated; carrying out experiments deliberately designed for the purpose of measuring presence and learning (i.e. rather than having outcomes related to presence and learning come out of studies focused on other things); and focusing on limited variables (the only in-
dependent variable in this study was communication modality, with three dependent variables of presence, learning and cognitive load).
CHAPTER 2
REVIEW OF THE LITERATURE

This literature review begins with a brief background and history of social virtual worlds generally and Second Life in particular. Readers familiar with this subject matter may wish to skim or bypass this initial section. The following section describes the theoretical background relevant to the proposed study, beginning with theories of working memory: Dual Coding Theory (DCT) from Paivio, followed by Baddeley’s theory of working memory, then Penney’s separate streams hypothesis. It is helpful to understand Paivio’s, Baddeley’s, and Penney’s models of working memory before moving on to Sweller’s cognitive load theory and, subsequently, on to specific aspects of learning that influence cognitive load, namely modality (Mousavi et al., 1995; Mayer & Moreno, 1998; Moreno, 2006) and presence (Heeter, 1992; Witmer & Singer, 2005).

2.1 A Brief History of Virtual Environments

If telephone, radio, film, and TV defined life in the 20th century, the virtual world is one true new medium of the 21st century (Damer, 2007). It is important to take a moment to define “social virtual world” clearly, a term used frequently in this dissertation. Though it can mean many things, for the purposes of this dissertation “social virtual world” will be used as shorthand for “a 3D, online, collaborative, multiplayer virtual world that is not a game.”

It is not surprising that electronically mediated communication (EMC) in virtual environments parallels the evolution of virtual environments from text-based chat forums to visual 3D environments that offer both text and voice chat. Text-based role-playing games, like the first Multi-User Dungeons (MUDs), were first seen in the 1970s and 1980s (Damer, 2007). Communication in MUDs and similar vir-
tual environments was necessarily text-based. Text-based environments are those in which messages are typed on a computer keyboard and read as text on a computer screen, typically by a person or persons at a different location from the message sender. Text-based EMC has taken a variety of forms since the 1970s (like email and chat), but something that all forms have in common is that the activity that takes place through them is constituted primarily—in many cases, exclusively—by visually presented language (Herring, 2001).

These text-based virtual environments evolved into visual ones as the age of affordable graphical computing dawned in the 1980s and 1990s and expanded to include other genres: First-person shooter games, fantasy role-playing games, simulators, shared board and game tables, and social virtual worlds (Damer, 2007). This last genre, the social virtual world, is the focus of the study presented here.

It is important to distinguish social virtual worlds from game play worlds. While the two may appear visually similar their terms are not interchangeable or synonymous. Put most simply, social virtual worlds are not games. Damer (2007) points out that social virtual worlds and game play worlds have different primary purposes: The primary purpose of social virtual worlds is the creation of meaning through the manipulation of the world and communication with others within the world, while structured play is the primary purpose of game play worlds (Damer, 2007).

Second Life is one such social virtual world. Second Life users, known as Residents, create content and by extension meaning. All content in Second Life is created and owned by Residents, not by Linden Lab, the creators of Second Life. The primary purpose of Second Life is communication and interaction with others. Though Residents may create games within Second Life, Second Life itself is a social virtual world and not a game. Anything that takes place in Second Life is referred
to as happening “in-world” or "IW.” Anything that happens in real life is referred to as happening “in real life” or “RL.”

Second Life is neither the first nor the only social virtual world. In the spring of 1995, a company called Worlds Incorporated launched Worlds Chat, a 3D space station where users “teleported,” could navigate through a rich environment, and could exchange text chat (Damer, 2007). Worlds Incorporated also launched Alphaworld three months later (Damer, 2007). Alphaworld is the first known social virtual world that allowed users to build objects and create content in-world using prefabricated objects. Alphaworld became and is known today as Active Worlds.

During the same period of time (i.e. in the 1990s) rich MMORPG environments became extraordinarily popular, much more so than social virtual worlds. While MMORPGs like Everquest (1999) and, more recently, World of Warcraft (WoW) saw their numbers of players grow from the thousands into the millions, most social virtual world companies proved too soon for their time. By the late 1990s, most social virtual world companies had lost financial backing, changed hands or vanished (Damer, 2007).

Second Life and There, another virtual world, both launched in beta in 2003. Both of these social virtual worlds allow users to create their own content (3D objects) using tools provided in the interface, and both worlds have their own economies. The currency of Second Life is the Linden Dollar (L$), also called Lindens, which at the time of this writing had an average exchange rate of $265-$272 L$:1$ USD. Residents use L$ to buy and sell virtual objects from other Residents in-world or outside of Second Life via XStreet SL, a website marketplace of Second Life goods also owned by Linden Lab. Linden Dollars can then be exchanged for U.S. dollars.

In addition to content creation and virtual economies, both Second Life and
There offer integrated real-time voice chat and continue to offer text chat. Voice chat means that Residents in Second Life, for example, can speak to one another in-world in real-time, with their real-world voices. Voice chat in Second Life is essentially voice-over IP (VOIP, like Skype and Ventrilo) in-world, with one’s real-world voice attached to one’s avatar. Though no two social virtual worlds are the same, they share some common characteristics:

- They are rich in visual appearance and media; 3D places that contain colorful sunsets and impressive textures (such as avatar skin that appears to reflect light and chairs that appear upholstered in fabric) as well as some combination of audio (environment or mood sounds like running water and blowing wind), streaming recorded music and/or live music performances, voice chat, video, web browsing, paintings and more.

- They are both inhabited and co-created by people participating from different physical locations (Damer, 2007).

- They are persistent. Social virtual worlds exist independently of their users, before a user logs on and after a user logs off (Yee, 2006). In contrast, the worlds of stand-alone games and local network games only exist when the game is started by the user and are thus dependent on the user activating it (Yee, 2006). More importantly, events and interactions occur in the world (driven by other users) even when any particular user is not logged on to the persistent world (Yee, 2006). This is true of Second Life, which exists and is very much active whether one’s avatar is logged in or not. Persistence is, of course, critical to keeping virtual economies alive and well.

- They offer text-based communication.

There is one more distinction worth making, which is between social virtual
worlds and virtual reality (VR). Social virtual worlds can be considered a very basic form of VR, which at its most basic is any technology that allows a user to interact with a computer-simulated environment. But social virtual worlds are not the same thing as VR, at least as VR is popularly defined and commonly described in the research literature. Systems that are considered true VR technology, for example, tend to offer different sensory input modalities to users, like head mounted displays (HUDs), wired gloves, and haptic (tactile) interfaces. Residents interact with Second Life using standard input devices, almost always a keyboard and mouse and sometimes a headset (to hear and speak). Please see the Glossary for a more detailed definition of VR.

2.2 Theories of Working Memory

This section describes the theories that provided the foundation for this study. Before moving on to cognitive load and ideas of how instructional methods impact working memory, it is helpful to understand working memory itself and, specifically, where the currently accepted model of the three-part model of working memory originated. To this end, Paivio’s Dual Coding Theory (DCT) is presented first, followed by Baddeley’s theory of working memory, and concludes with Penney’s separate streams hypothesis.

2.2.1 Dual Coding Theory (DCT). Dual Coding Theory (DCT), introduced by Allan Paivio in the late 1960s and early 1970s, is an empirically well-founded model of the mental processes that underlie human behavior and experience (Paivio, 1975). Dual Coding Theory uses the collective action of nonverbal and verbal mental systems, which are specialized for the processing of imagery and linguistic information respectively, to explain psychological phenomena (Clark & Paivio, 1991). Imagery and verbal associative processes jointly determine learning and memory performance. Direct and indirect associations between verbal codes influence not just processing,
but also the storage and retrieval of information (Clark & Paivio, 1991).

Paivio describes thinking as a process comprised of three primary tasks:

1. Taking in or encoding stimulus information;
2. Organizing and storing it in memory; and
3. Retrieving that information according to the requirements of a given task.

With DCT, Paivio showed that certain characteristics of imagery contribute to the efficiency of each of these three stages (encoding, storing and retrieving) of thinking. Although Paivio emphasizes the role of imagery in DCT, he also notes that imagery could not be studied or understood in isolation from language. This belief led Paivio to compare and contrast nonverbal imagery with verbal symbolic processes, on the assumption that human thinking involves a continuous interplay of both cognitive systems which, though interconnected, are functionally distinct (Paivio, 1975). This, in a nutshell, describes the dual coding part of DCT: Verbal and nonverbal systems are functionally distinct (two systems, differently coded) but continuously interface with each other. The following describe some points of DCT:

- It distinguishes between synchronic and serial processing. A person who enters a room can see and visually process all of the items in the room at once, but the constraints of language prevent the person from articulating the items all at once. The person has describe the items in the room to others in some kind of order (by type, placement, color and so on). This person can also change the order in which they relay information by request; left to right, then right to left, then all the red items first, for example. This example illustrates dual coding, in which sequential verbal structures constrain the order in which corresponding visually processed (imaginal) units are generated or retrieved by verbal cues,
but do not constrain the processing of the spatial information contained in the
integrated visual units themselves (Paivio, 1975).

- The more items there are in a set, the less likely it is that verbal representa-
tion will be used for them. Paivio described a series of experiments using
sets of two to four items and describes when either pictorial or verbal repre-
sentation was fastest. He concluded that “encoding into meaningful structures
is more efficient with pictures than words” because simultaneously available
pictures “have relatively direct access to the imagery processes that transform
them into synchronously-organized scenes.” This, in turn, activates the verbal
processes necessary for generating descriptive sentences. Words are at a dis-
advantage because images must be generated indirectly and sequentially from
the individual items in a set, retained in that form, and then reorganized into
imaginary scenes before sentences can even be generated to describe the images
(Paivio, 1975).

- People recall high imagery items faster.


Nearly two decades later, Clark and Paivio (1991) published a paper on the role
of DCT in education. Initially, they described the application of DCT to many aspects
of education—the science of educational psychology, motor skills development, and
teacher education to name just a few areas—but the application of DCT to effective
instruction is most relevant to the purposes of this study. A few of Clark and Paivio’s
guidelines for effective instruction are now briefly mentioned, most of which will be
familiar to technical communicators and instructional designers.

First, DCT has shown that lessons containing concrete information and vivid
images are easier to comprehend and remember than abstract lessons. Clark and
Paivio (1991) cite one study, for example, in which fourth graders were asked to
recall information from tape-recorded newspaper stories. Children who listened to
the stories while viewing relevant pictures recalled more information than children
who only heard the stories and did not view relevant pictures.

Second, DCT theory has led to instructional guidelines such as “use models,
examples and illustrations” (Clark & Paivio, 1991). Models, examples and illus-
trations create concrete information and vivid images.

Instructional guidelines like these have led some teachers and professors to
explore Second Life as a medium for creating vivid images, models, examples and
illustration. Professor Edward L. Lamoureux, of Bradley University in Peoria, IL,
teaches ethnography classes in Second Life, while the students of Beth L. Ritter-
Guth, English instructor at Lehigh Carbon Community College and an adjunct at
DeSales University, created Literature Alive in Second Life for her students, a project
that engages students in reading by guiding their avatars on tours of pixilated versions
of famous literary spots, like Dante’s Inferno (Foster, 2007). There’s no arguing that
virtual worlds like Second Life are visually rich and may—through tools that enable
3D object creation, manipulation, physics changes and texture creation—help online
instruction to be more concrete or effective than web sites comprised primarily of
text.

Paivio’s DCT shares a key assumption with Baddeley’s theory of working
memory (discussed next), which is that the coding processes underlying working
memory are neutral with respect to input modality, including the verbal-nonverbal
dichotomy. Penney’s separate streams hypothesis, discussed later, departs from both
Paivio and Baddeley on this point.

Paivio and Penney share a common finding, however, that simultaneous pre-
sentation seems preferable for the presentation of visual material. Paivio wrote that
“The defining property of spatially-parallel processing is simultaneity of functioning.” He saw this as evidence that “Simultaneously given information can be processed over a broad area of the retina” (Paivio, 1975). Later, Penney (1989) contended that the simultaneous spatial presentation of information (as opposed to serial/sequential presentation) aided recall and other tasks. This is because the dominant organization mode for the visual stream (one of the two streams in her separate streams hypothesis) is spatial, with items presented simultaneously in different spatial locations strongly associated. This strong association supports recall.

2.2.2 Baddeley’s Theory of Working Memory. The term “working memory” appears to have first been proposed by Miller et al. (1960) in the book titled Plans and the Structure of Behavior (Baddeley, 2002). In the field of cognitive psychology the term “working memory” describes the cognitive system or systems involved in the temporary maintenance and manipulation of information (Baddeley, 2002).

Baddeley and Hitch (1974) proposed a multi-component model of working memory as a theoretical framework whose function was, according to Baddeley, “to give an economical and coherent account of a relatively wide range of data” on how people process information (Baddeley, 2002). In Baddeley and Hitch’s original model, working (also known as short-term) memory is divided into three components:

1. The Central Executive, which controls attention and has limited capacity. The Central Executive is assisted by two subsystems:

2. The Phonological Loop, concerned with acoustic (auditory) and verbal (spoken) information, and

3. The Visuospatial Sketchpad, concerned with visual and spatial information (including written text).
2.2.2.1 The Phonological Loop. Some findings from early empirical tests of the phonological loop that have contributed to the field of instructional design, and that may by now be common knowledge to many in the fields of linguistics and technical communication, are:

- Things that sound different are easier to remember. Conrad and Hull (1964) showed, for example, that the recall of a series of letters similar in sound—B, V, G, T, C, D—was poorer than the recall of dissimilar sounding letters—F, K, Y, W, M, R. This study and others like it indicated that our ability to recall a sequence of items is greatly affected by the phonological characteristics of the items. This became known as the "phonological similarity effect" (Conrad & Hull, 1964), which indicates the importance of the role played by phonological coding in short-term memory.

- Shorter length words are remembered better than longer length words: People are more likely to correctly recall a series like sum, pay, wit, bar, hop than a series like helicopter, university, television, alligator, opportunity.

- The phonological loop was originally called the “articulatory loop” which more accurately described its scope: The phonological loop does not include speech one hears from others, but is instead concerned with the words an individual speaks (articulates) either aloud or silently to oneself, by oneself. Baddeley’s theory of working memory holds that auditory (spoken words) information has direct access to the phonological store, but visual information (written words) only has access through the articulatory control process, which allows visually presented material to be phonologically coded. Put more simply: Written words cannot reach the phonological store until they are not just read, but articulated either aloud or silently to oneself. The act of articulating words matters most.
2.2.2.2 The Visuospatial Sketchpad. This system, not nearly as tested as the phonological loop at the time of Baddeley’s writing, was assumed to be capable of temporarily maintaining and manipulating visuospatial information and to play an important role in spatial orientation and in one’s ability to solve visuospatial problems. The visuospatial sketchpad, theoretically, formed an interface between visual and spatial information, accessed either through the senses or from long-term memory (LTM).

2.2.2.3 The Central Executive. In retrospect, Baddeley speaks quite frankly about the initially vague, catchall conception of the central executive:

The third component of the working memory framework, the central executive, was initially conceived in the vaguest possible terms as a limited capacity pool of general processing resources. For the first decade, it served principally as a convenient ragbag into which could be thrust such awkward questions as what determined when the sketchpad or phonological loop was used and how they were combined. Implicitly, the central executive functioned as a homunculus, a little man who took the important decisions as to how the two slave systems should be used (Baddeley, 2002).

Baddeley did not feel, however, that the executive’s resemblance to a homunculus was a bad thing, “as long it is accepted that the role of a homunculus is to remind researchers of those functions that they have not yet explained.” The following are some roles (processes) thought to be performed by the central executive:

• Role #1: Focusing Attention: The central executive has the capacity to focus attention, which is important if we follow Baddeley’s related assumption that anything that limits attention capacity will impair performance (learning). Subsequent decades of work by Baddeley, his colleagues and others has shown that focusing attention is almost certainly one important process the central executive performs.
• **Role #2: Dividing Attention:** The central executive can divide attention, not just focus it. Baddeley and others designed studies to measure dual-task performance and thus attention divided between two different types of tasks, one task targeting the phonological loop and the other task targeting the visuospatial sketchpad. We’ll read more about divided attention when we discuss the split-attention effect later in this literature review.

• **Role #3: Switching Attention:** It’s unclear if the central executive can switch attention, due to mixed findings in the literature. Some studies suggest, for example, that the capacity to switch attention is not strongly dependent on the central executive but instead on the phonological loop, making it difficult to pinpoint the central executive as the source of this ability.

Baddeley’s original model of working memory did not address the need to integrate information from the subsidiary systems and from LTM in a way that allows active maintenance and manipulation of information. To solve this problem, Baddeley added a fourth component called the “episodic buffer” to his original three-part model of working memory. The episodic buffer represents a storage system using a multimodal code (phonological and visuospatial), episodic in the sense that it holds integrated episodes (or scenes) and “buffer” in the sense that it provides a limited capacity interface between systems that use different codes.

Today, Baddeley sees the role of the central executive as extending beyond memory function to managing attention, whereas the episodic buffer is considered purely mnemonic in character. Information retrieval from the episodic buffer is thought to happen through conscious awareness, allowing multiple sources of information to be considered simultaneously and using LTM to facilitate chunking.

Chunking, aggregating items into larger units (words into phrases, for exam-
ple, or numbers into sets), enhances recall by allowing more economical storage and alleviation of cognitive load. It is now a key principle in the field of document design and technical communication that resulted at least partially from Baddeley’s work, though Simon and Chase (1973) also used the term “chunking” to describe how people used and organized their short term memories. Baddeley’s revised model of working memory differs from the initial model in two important ways:

- He proposes an explicit link between the phonological loop and the visuospatial sketchpad, and between verbal and visual LTM. The flow of information is bidirectional: The subsidiary systems feed the relevant areas of LTM but are themselves assisted by implicit knowledge of language and of the visuospatial world, aiding recall.

- He considers the episodic buffer capable of combining information from LTM with that from the phonological loop and visuospatial sketchpad. The updated model of working memory does not directly link the two subsystems to the episodic buffer, which indicates that these transformations depend heavily on the central executive.

For the study described here, the idea that articulated speech and visual information are encoded and managed differently in working memory is a salient one. According to Baddeley (2002), any study participant who articulates speech (whether silently to him or herself or aloud to others) accesses the phonological loop. In the study presented here, this could conceivably include all participants in all groups (Group V, Group T and Group VT).

In addition, all test participants would access the visuospatial sketchpad when asked to create and manipulate 3D content. The goal was to see which channel, the audio or the visual, both, or neither became overloaded during the 3D building class.
in Second Life. A combination of printed text (for communication) and 3D content (for manipulation) may overload the visual channel due to the visual nature of both. Baddeley’s early notion of separate areas of working memory, for phonological and visuospatial information, went on to become common in other theories of working memory and an empirically established theory in and of itself.

In addition, the concept of chunking and knowledge of other conditions under which people better remember things have informed the fields of technical communication and instructional design.

Other theorists depart from Baddeley on two main points: That auditory information seen without being spoken aloud (silently articulated) is the same as auditory information that is heard, and that modality of information does not matter (other theorists think it does). Penney’s separate streams hypothesis reflects these two points of departure from Baddeley and is considered even more relevant and directly applicable to this study.

2.2.3 Penney’s Separate Streams Hypothesis. In a 1989 paper, Penney reviewed the literature on the effects of auditory and verbal presentation on short-term retention of verbal stimuli. She also proposed a model of the structure of short-term memory called the separate streams hypothesis. The separate streams hypothesis holds that processing of auditorily and visually presented verbal items is carried out separately in short-term memory.

Penney cited the work of Baddeley and his colleagues as both a foundation and point of departure for the separate streams hypothesis. Unlike Baddeley and Paivio, Penney’s hypothesis recognizes the importance of modality effects, and modality marks Penney’s main point of departure:

In contrast to this (i.e. Baddeley’s) view of modality effects as being inconsequential, I want to argue that modality effects reflect the inherent structure of
short-term memory, and that no theory of human memory will be adequate if it does not provide a complete account of these phenomena” (Penney, 1989).

Modality effects were a primary concern of the study described here, which tested the role modality played in perceptions of presence, learning and cognitive load. Following are five key summaries of findings from Penney’s (1989) paper, each of which provides evidence for the separate streams hypothesis. Because a large portion of Penney’s paper is a literature review, please be assured that Penney (1989) cited ample evidence in the literature for each of the summaries presented here in more condensed form.

1. Two modalities are better than one for dividing attention. People more easily divide attention between two concurrent tasks (both tasks requiring processing of verbal stimuli) when the stimuli are presented to two modalities (auditory and visual modalities) rather than to one modality. Processing capacity apparently increases when two modalities are used, when semantic processing is required. These findings led Penney to suspect that modality-specific processing resources exist at both the perceptual and semantic levels, thus dividing attention more easily between two modalities than between two messages in a single mode.

2. Presenting items to two sensory modes rather than one improves short-term memory retention. Because separate processing streams exist for auditorily and visually presented items in short-term memory, retention is greater when information is presented to two modalities rather than one. Here, Penney specifically means bisensory presentation, in which lists of different items are presented simultaneously to different sensory modalities.

3. Selective interference occurs for both auditory and visual streams. The presentation of auditory memory items concurrently with visual items; requiring
test participants to articulate irrelevant words during the presentation of visual items; and the simultaneous presentation of visual memory items all promote the use of a visual code for retention. This visual code, however, is susceptible to interference from subsequent visual presentations of a verbal distractor stimulus (i.e. written words).

4. Presentation modality provides a strong basis for organization. Not only can subjects organize their recall by modality when asked (for example, recall items in temporal order and then in alphabetical order), but it also appears that organization by modality is the preferred or optimal organization scheme. When subjects are presented with lists of different items presented in the auditory and visual modalities, the preferred recall strategy is to report one modality first and then recall items from the other mode. Presentation modality provides a stronger organization than language or category.

5. Modality-specific short-term memory deficits have been discovered. Case studies with disabled patients demonstrated severe impairments in retention of auditorily presented verbal material: In the case of auditory presentation, the A code that persists and is so useful for normal subjects was not available to impaired subjects.

As noted earlier in this section, Penney’s primary point of departure with Baddeley is that his theory of working memory did not account for any modality effects (like the five effects described by Penney and summarized above). Modality effects indicate that there are differences in the processing of auditory and visual input that Baddeley’s phonological loop (formerly called the articulatory loop, in which silently articulated visual items produce the same kind of code in short-term memory) does not allow.
In another point of departure from Baddeley, Penney argues that the short-term memory trace that is laid down when a subject silently articulates a visually presented item does not contain the same information as the memory trace that results from auditory presentation of the item. For instance, when someone sees a printed word (while reading silently to oneself for example, as in some of Baddeley’s studies) or imagines the sound of a word (seeing a printed word on a page and imagining someone else saying the word aloud), this act of silent articulation does not contain the same information as the memory trace that results from hearing someone else read the word aloud.

To elaborate on this fundamental difference between silent articulation and hearing, Penney discussed acoustic code (A code) and phonological code (P code). The P code is internally generated by the silent articulation of visually presented items. Reading silently to oneself from a book, for example, would internally generate the P code. Penney’s P code is similar to Baddeley’s phonological code produced by the articulatory loop, and P code is common to both auditory and visually presented stimuli.

The A code, however, is presented only for stimuli that are heard: The A code is a sensory-based code created as a result of auditory presentation. Penney believes, and cites ample evidence to support her belief, that A code boosts human recall of recent auditory items relative to visual items in short-term memory tasks and thereby produces the modality effect.

Penney also distinguishes between two different models of short-term memory: The traditional and the proceduralist views. Her separate streams hypothesis reflects the proceduralist view. These two models of short-term memory are described in Table 2.1.
### Table 2.1. Comparison of Penney’s Two Models of Short-term Memory: Traditional and Proceduralist

<table>
<thead>
<tr>
<th>Topic</th>
<th>Traditional View</th>
<th>Proceduralist View</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory Traces</td>
<td>Semantic aspects of the memory trace are primary and sensory aspects are secondary. Perceptual processing extracts the meaning of a stimulus. This meaning forms the core of the memory trace.</td>
<td>Sensory aspects of the memory trace are primary and the semantic aspects secondary.</td>
</tr>
<tr>
<td>Memory Storage</td>
<td>Information about the sensory properties of a stimulus is stored separately, as tags or attributes appended to the memory trace.</td>
<td>Perceptual processing provides the basis of a memory trace. Meaning is added later, not first.</td>
</tr>
<tr>
<td>Representation</td>
<td>Knowledge representation in memory is independent of the way in which knowledge is acquired. There is no difference between reading and listening, for example. Features of a message like cadence and pitch play no role in how the message is represented.</td>
<td>Knowledge representation in memory depends on the way in which the knowledge was acquired. There is a difference between reading and listening, for example. Features of a message like cadence and pitch play a prominent role in representation in memory.</td>
</tr>
</tbody>
</table>

If Penney’s separate streams hypothesis has one shortcoming, it is the almost exclusive focus on the processing and retention of verbal material. This focus has its advantages, though: With verbal material, it is possible to vary input modality while keeping the informational content of the memory material the same (as the study here did by testing the same learning script in three modalities: voice only, text only, and voice and text simultaneously). It is easier to compare retention of auditorily and visually presented words under parallel presentation conditions when verbal material
Penney’s separate streams hypothesis, however, is valuable because she elaborates on the nature of both auditory and visual streams and the differences between them, despite the primary focus on verbal material. Understanding the auditory and visual streams reinforces the design of the study proposed here (Group V, Group T, and Group VT) and why we might expect to see differences between these groups. The characteristics of the auditory stream and A code, and of the visual stream are summarized in Table 2.2. Please note that the visual stream has no corresponding or equivalent code to A code because there is little evidence for a sensory-based visual code that contributes to performance in typical short-term memory tasks (Penney, 1989). Instead, it seems that subjects generate phonological or P code for the visual stream, and that short-term memory performance is based primarily on this.

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Auditory Stream</th>
<th>Visual Stream</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perception</td>
<td>Based on perceptual analysis of acoustic information: A word must be spoken and heard in order to be presented.</td>
<td>Has a stronger spatial component and a relatively weaker temporal one than the Auditory.</td>
</tr>
<tr>
<td>Organization Mode</td>
<td>Sequential: Seems specialized to preserve the order of items in short-term memory. Temporal output order aids recall.</td>
<td>Spatial: Items presented simultaneously in different spatial locations strongly associated. Simultaneous spatial presentation aids recall.</td>
</tr>
<tr>
<td>Attention</td>
<td>Less: Items can be retained in short-term memory without continued attention and &quot;be remembered anyway.&quot;</td>
<td>More: Require attention to be retained in short-term memory.</td>
</tr>
<tr>
<td>Decline</td>
<td>Little with age</td>
<td>Greater with age</td>
</tr>
</tbody>
</table>

2.2.4 Cognitive Load Theory. For the most part, cognitive load theory has been
applied in the context of learning and instructional design but it has progressively broadened in disciplinary applications in recent years (Lawrence, 2006). Cognitive load is the amount of mental energy required to process a given amount of information. As the amount of information increases, so does the cognitive load on our mental resources (Feinberg, 2003). Cognitive load theory holds that learning will be inhibited when the amount of information and instruction exceed the capacity and limitations of our mental resources.

Sweller and Cooper (1998) described a cognitive modal model of learning that distinguishes between three distinct memory types (modes) described below: sensory memory, working memory and long-term memory (Feinberg, 2003):

1. Sensory memory deals with incoming stimuli from our senses, including sights, sounds, smells, tastes and touches. Sensory memories extinguish extremely quickly (approximately 30 seconds for visual information and three seconds for auditory information), and will be lost during this short period of time if we don’t assign meaning to them (Feinberg, 2003). The importance of assigning meaning is also described by Penney, who states that semantic processing of an item in short-term memory occurs continuously with perceptual processing. Sensory aspects of the memory trace are primary and semantic aspects are secondary, not the reverse (Penney, 1989).

2. Working (short-term) memory has limited capacity. Working memory, also referred to as short-term memory, is itself a three-part system of limited capacity. Modern theories of working memory place a greater emphasis on these partially independent processors, associated with different sensory modes (Sweller et al., 1998). This type of model was presented earlier in Baddeley’s model of working memory, which Feinberg reiterates. Sweller focuses less on models of working memory than on the implications of working memory limitations for learning
and instructional design. He emphasizes this point by stating that all conscious cognitive activity occurs in a structure (working memory) that seems to preclude all but the most basic processes: “Anything beyond the simplest cognitive activities appears to overwhelm working memory.” Because working memory is key to learning and yet so limited in capacity, learners are automatically placed in a difficult position.

3. Long-term memory is “permanent memory.” Long-term memory refers to the immense body of knowledge and skills that we hold in a more-or-less permanently accessible form (Feinberg, 2003). According to Sweller, people are not directly conscious of long-term memory because our awareness of its contents and functioning is filtered through working memory (Sweller & Cooper, 1988). The human cognitive system emphasizes the ability to store seemingly unlimited amounts of information in long-term memory, from small, isolated facts to large, complex interactions and procedures. Our intellectual prowess comes from the ability to recall stored knowledge, not from our ability to engage in long, complex chains of reasoning in working memory (Sweller & Cooper, 1988).

2.2.5 Intrinsic, Extrinsic and germane Cognitive Load. Cognitive load theory identifies three types of cognitive load: intrinsic, extraneous and germane. Each type is briefly described here:

1. Intrinsic cognitive load relates to the difficulty of the content to be learned (Sweller et. al. 1998). Intrinsic load cannot be modified by instructional design (Feinberg, 2003). In other words, difficulty is inherent to the content or subject matter itself and not to the way in which content is presented.

2. Extraneous cognitive load is defined as any cognitive activity engaged in because of the way the task is organized and presented, not because it is essential to
attaining relevant goals (Feinberg 2003) or inherent to the content to be learned. Unlike intrinsic cognitive load, extraneous cognitive load can be modified by instructional design. By changing materials presented to students, the level of extraneous cognitive load may be modified, thus facilitating learning (Feinberg 2003).

3. Germane cognitive load, like the term germane itself, means closely related; relevant; pertinent. In this case, germane refers to items and processes highly relevant to learning. Like extraneous cognitive load, germane cognitive load can be modified by instructional design. Unlike intrinsic and extraneous and cognitive load, germane cognitive load is best when it is high. According to Sweller, increasing germane cognitive load involves redirecting attention: Learners attention must be withdrawn from processes not relevant to learning and directed toward processes that are relevant to learning.

The total cognitive load associated with an instructional design is the sum of intrinsic cognitive load plus extraneous cognitive load plus germane cognitive load. In regard to this study, for example, intrinsic cognitive load was created by the inherent difficulty in learning how to use Second Life (software) to create and manipulate 3D objects, and with creating and manipulating 3D objects in general. Extraneous cognitive load may be increased by modality or by other aspects of Second Life: additional environment sounds (like wind), social interactions with other participants or interruptions. Finally, germane cognitive load could be increased by redirecting attention from other things in Second Life (side conversations or shopping, to name just two of many possibilities for distraction in Second Life) to the course content and demonstrations on building 3D objects.

2.2.6 Split-Attention, Redundancy and Modality. In addition to describing a three-part model of modal memory, cognitive load theory describes three learning
techniques that directly apply to instruction using multimedia technology: the split-attention effect, the redundancy effect, and the modality effect (Feinberg, 2003):

1. The split-attention effect holds that cognitive load generated by irrelevant activities, like text and graphics competing for a learner’s attention, can impede skill acquisition. Instruction that mitigates competition between elements of information (by integrating graphics and text, for example) has been found to be superior to split-sources of information (see Chandler & Sweller, 1991; Mayer & Moreno, 1998; Moreno & Mayer, 2002; Penney, 1989).

2. The redundancy effect states that redundant sources of information (such as voice narration that repeats selected portions of information provided in text) can increase cognitive processing that could otherwise be freed for other purposes. For this reason, cognitive load theory cautions against the use of simultaneous representations of redundant content in instruction (Feinberg, 2003).

3. The modality effect complements the split-attention effect, though at first glance it may seem to do just the opposite. Research indicates that more items will be recalled in a memory test if some of the items are presented in a visual modality and some in an auditory modality, rather than all in a single modality. This finding suggests that more memory capacity is available when dual modalities are used (Penney, 1989). In this sense, dual-modality involves a split-attention effect but in so doing supplies a larger working memory. Larger working memory enables learners to better assimilate instructional material. As Penney and others point out, however, if the auditory component is too long or highly complex, it will create excessive demand on working memory. Unconstrained use of multimedia (e.g. animation and streaming video) will also create excessive demand on working memory (Mayer et al., 2003).
There is little published research on the application of cognitive load theory to virtual environments (including MMORPGs and social virtual worlds) created for entertainment and leisure purposes. The literature does acknowledge, however, several compelling reasons for developing highly immersive environments, including reasons based on cognitive load theory (Mayer & Moreno, 2002). By making interactions with technology more natural, immersive learning environments might result in cognitive load reduction: More immersive environments will supposedly drive students limited attention to learning content rather than the interface (Mayer & Moreno, 2002).

There is a wealth of research on cognitive load theory as it applies to education and instructional design, some of which examines learning in virtual environments created for educational purposes. Though the virtual and multimedia educational environments described in the literature are not the same as social virtual worlds and none of them focus on Second Life, some of the literature on cognitive load and learning in multimedia environments is useful and reinforces the relevance of the study design and methodology proposed here. We now cover several experiments in this vein, by Mousavi et al. (1995) and then by Moreno and Mayer (1998; 2002).

2.3 Modality and Learning

This section begins by discussing the relationship between learning and modality. For the purposes of this literature review, studies by Mousavi et al. (1995) and a series of experiments by Moreno and Mayer are of greatest relevance. We begin with the work by Mousavi et al. (1995) since the subsequent work of Moreno and Mayer directly builds on that of their predecessors.

2.3.1 Mousavi, Low and Sweller. The work of Mousavi et al. (1995) provided an important empirical foundation for the work of Moreno and Mayer (described later in this section) and for the study presented here, which aimed to expand the work
of Moreno and Mayer. The primary goal of Mousavi et al.’s (1995) work was not to measure cognitive load theory directly but to use cognitive load theory to generate experiments to provide direct implications for instruction.

Mousavi et al. (1995) generated six experiments that are important to review because their methodology is strong and reliable: Six experiments were conducted one after the other (linearly), in the order presented here, and each experiment was designed to clarify and test possible alternative explanations for the results of the experiment that came before it. These six experiments are now reviewed in greater detail.

2.3.1.1 Experiment #1: Three Groups and Three Presentation Modes.

This is the first and primary experiment that Mousavi et al. (1995) conducted, not only in the sense that it was literally the first experiment conducted of a total six, but also that each subsequent experiment either clarifies or expands on the results from this experiment.

Study participants were divided into three groups, depending on the presentation mode (a combination of diagrams, written text, and audio statements from a tape recorder) of worked geometry problems. These three groups were:

**The simultaneous group:**

a. Studied worked examples with a diagram

b. With associated statements presented visually (text)

c. While hearing statements played from tape recorder (audio narration).

**The visual-visual group:**

a. Studied worked examples with a diagram

b. With associated statements presented visually (text)
c. Without statements playing from a tape recorder.

The visual-auditory group:

a. Studied worked examples with a diagram
b. Without statements presented visually (no text)
c. With an audio recording playing.

Each group and its attributes is also summarized in the table below:

<table>
<thead>
<tr>
<th>Group</th>
<th>Diagram</th>
<th>Text</th>
<th>Narration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simultaneous</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Visual-Visual</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Visual-Auditory</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

2.3.1.2 Experiment #1: Results Summary.

- The visual-auditory group required less time to solve the problems than the visual-visual group. The simultaneous group also required less time to solve problems than the visual-visual group.

- The authors found no significant difference between the simultaneous and visual-auditory groups.

- For combined (i.e. similar) test problems, the authors found a significant difference between the three groups. When examining data grouped by problem similarity, they found:

- The visual-auditory group spent less time solving the same problems than did the visual-visual and simultaneous groups.

- No significant difference in solving the same problems between the simultaneous and visual-visual groups.
Overall, these results demonstrate that the presentation of information in mixed auditory and visual mode, rather than a single mode, had beneficial effects on learning, presumably by expanding effective working memory capacity by not overloading a single working memory channel.

2.3.1.3 Experiment #2: Test Time Equalization for Experiment #1. In the second of Mousavi et al.’s (1995) experiments, times were equalized to test the possibility that presentation times may have skewed the results of Experiment #1. The authors thought that the visual-visual group (in Experiment #1) may have shown the worst performance because they spent less time studying the worked examples. In Experiment #1, participants in the visual-visual group determined their own presentation times (by looking at the example for as long as they needed), whereas the length of the tape recording determined the presentation times for the visual-auditory and simultaneous groups.

2.3.1.4 Experiment #2: Results Summary. Overall, the pattern of results in Experiment #2 stayed similar to those found in Experiment #1, so additional detail is not necessary here. These results indicate that differences between groups in Experiment #1 were not due to different presentation and thus acquisition times, thereby supporting the hypothesis that presenting worked examples in two modes expands effective working memory.

2.3.1.5 Experiment #3: Simultaneous vs. Successive Presentation. Experiment #3 also expanded on Experiment #1 by testing the order in which information was presented: simultaneously or successively. Since written statements cannot be viewed simultaneously, students need to read one statement, hold it in working memory and then attend to the diagram in order to integrate the diagram with the written statement. This test was designed to determine whether participants in the simultaneous presentation groups had had an advantage, thus affecting results.
Experiment #3 was designed to test the following cognitive load hypothesis: If working memory increases with a dual mode of presentation, then it should be easier for students to integrate the diagrams and statements regardless of whether they’re presented simultaneously or successively.

To this end, Experiment #3 compared the following four modes of presenting geometry worked examples:

1. Visual-visual simultaneous
2. Visual-visual sequential
3. Visual-auditory simultaneous
4. Visual-auditory sequential

Both of the simultaneous groups (items #1 and #3 in the list above) were identical to the presentation modes given to the visual-visual and visual-auditory groups in Experiment #1.

2.3.1.6 Experiment #3: Results Summary.

- Analysis showed a significant modality effect between the two visual-visual and the two visual-auditory groups. The two visual-auditory groups required less time to solve the two repeat problems (i.e. problems repeated from Experiment #1).
- Presentation sequence (simultaneous vs. sequential) produced no significant effect.
- The authors found no significant modality-by-presentation-sequence interaction.
These results suggest that presentation sequence, the order in which information is presented, does not matter. The advantage of visual-auditory presentation is what matters (that the presentation is visual-auditory), not the order in which information is presented. These results further reinforce the results from Experiments #1 and #2, which also showed the superiority of visual-auditory information to other presentation modes.

2.3.1.7 Experiment #4: Test Time Equalization for Experiment #3. In Experiment #4 Mousavi et al. (1995) equalized the study times for Experiment #3, just as Experiment #2 did for Experiment #1.

2.3.1.8 Experiment #4: Results Summary. The results of Experiment #4 were consistent with those of Experiment #3 so further detail is not provided here. It is worth noting, however, that the results of Experiments #3 and #4 are two reasons why the study presented here included the simultaneous presentation of information in its methodology, and why this study tested modality only and not simultaneous vs. sequential presentation conditions for each modality.

2.3.1.9 Experiment #5: Expansion of Test Materials. Experiment #5 expanded on the type of materials tested, based on the idea that test results “should not be restricted to instructional materials in which one source of information is a diagram and the other text” (Mousavi et al., 1995). In other words, the authors wanted to ensure that the results of the earlier experiments in the series weren’t based on the instructional materials themselves: If materials were varied would the results still hold?

In Experiment #5 both sources of information consisted of verbal statements, rather than a diagram and statements. This change tested the mixed-mode effect with mathematical problems in the following four modes:
• Verbal-written: The mathematical problem and solution are both presented in written text/sentence format.

• Verbal-auditory: The mathematical problem is presented in visual/written sentence form and the solution is played as spoken narration from a tape recorder.

• Diagram-written: The mathematical problem is presented as a diagram and the solution is presented as written text.

• Diagram-auditory: The mathematical problem is presented as a diagram and the solution is played as spoken narration from a tape recorder.

Based on the results of the previous four experiments, which favored the auditory condition, the authors predicted that the performance of students who learned under verbal-auditory and diagram-auditory conditions would be superior to that of students who learned under verbal-written and diagram-written conditions, respectively.

2.3.1.10 Experiment #5: Results Summary. The results demonstrated the superiority of auditory solution statements over written solution statements, regardless of whether the problem information is presented in written or diagrammatic form. The authors found no significant interaction between the problem and solution statements, which suggests that the auditory statements had equivalent impact whether paired with written or diagrammatic problems. These results reinforce cognitive load theory: If two sources of information are presented in different modalities, learning may be facilitated by an expansion of available working memory capacity.

2.3.1.11 Experiment #6: High Intrinsic Cognitive Load for Visual Processes. All five experiments described so far demonstrated the beneficial effects of presenting instructional material in a mixed auditory and visual mode. This finding supports the idea that working memory capacity can be increased when a dual-
presentation mode is used. Mousavi et al. (1995) felt this finding was still open to alternative explanation, so they designed Experiment #6 to test the possibility that visual processes (like reading) are inherently more cognitively demanding than the listening process. Visual processes being inherently more cognitively demanding would explain the results of the prior five experiments, all of which favored the auditory mode. In Experiment #6 the authors tested two groups:

- Visual-visual mode
- Auditory-auditory mode

If prior results were due to reading difficulties, the performance of the auditory-auditory group should be better than that of the visual-visual group. Please note that there is an age difference between the subjects in Experiments #5 and #6: The subjects in Experiments #1-#4 are older (eighth grade U.S. equivalents) so reading may have been easier for these more experienced readers than the subjects in experiments #5 and #6 (fourth grade equivalents).

### 2.3.1.12 Experiment #6: Results Summary.

Overall, the results of Experiment #6 indicated that reading was no more cognitively demanding than listening to the same information. A significant difference in solving repeat problems during acquisition favored the visual-visual group, indicating that the opposite may actually be true (i.e. that auditory-auditory may be more intrinsically demanding of cognitive load).

### 2.3.2 Moreno and Mayer’s Modality Principle and Cognitive-Affective Theory of Learning with Media.

As noted in the Introduction, the modality principle holds that words should be presented auditorily (spoken or narrated) rather than visually (written) when presenting a “multimedia explanation,” i.e. a combination of visual information like pictures with corresponding verbal information. In
the study presented here, participants were presented with a multimedia explanation in the form of an interactive class on building 3D objects using primitive 3D shapes that could be stretched, turned, moved, and linked together (the “visual information” part of the modality principle). Each class was conducted in one of three possible modalities (voice only, text only, or voice and text together), which provided the “corresponding verbal information” and “words” part of the modality principle.

The modality principle owes much to two preceding theories: Sweller’s theory of cognitive load and Penney’s separate streams hypothesis. Both were discussed earlier in this literature review but, briefly, Sweller’s cognitive load theory describes a working memory of limited capacity, with partially independent subcomponents that deal with auditory/verbal material or visual 2- or 3-dimensional information (Sweller et al., 1998). Penney’s separate streams hypothesis is based on modality effects observed during research on verbal learning and short-term (working) memory, which indicated that auditory presentation almost always resulted in higher recall than did visual presentation (Penney, 1989). Penney’s hypothesis describes how written and spoken words are encoded differently, and how acoustic (A) code persists and boosts recall of recent auditory (spoken/heard) items relative to visual items in short-term memory tasks, thus producing the modality effect. Moreno and Mayer (2002) acknowledge that the modality principle is comparable with the one offered by Penney, which they have included in their cognitive-affective theory of learning with media (CATLM).

The CATLM is a sort of theoretical summary in that it is comprised of a list of seven assumptions compiled from cognitive research by investigators other than Moreno and Mayer (2002). Despite this summary quality, the CATLM is both useful and convenient because it has done the hard work of combining empirically sound cognitive principles into a single, cohesive theory that can be tested whole. For the
purposes of this study, which studied differences between communication modalities, the modality principle that Moreno and Mayer derive from the broader CATLM is more relevant and interesting. The modality principle is derived from three of the seven total CATLM assumptions. These three assumptions are:

1. Independent information-processing channels
2. Limited working memory capacity
3. Dual coding

The rationale underlying the modality principle is that, by using the auditory channel to process the words, effective working memory capacity is expanded because students are not forced to split their limited visual working memories between the on-screen text and the pictorial information. Pictures are processed through the visual information channel, while spoken words are processed through the auditory channel. Processing of words is off-loaded onto the auditory channel, which is otherwise underused (Moreno, 2006).

The work of Moreno and Mayer (2002) has expanded on the findings of Mousavi et al. (1995) (described in the prior section). In a format similar to that of Mousavi et al. (1995), Moreno (2006) summarizes a set of three experiments conducted by either Mayer and Moreno together or by Moreno herself, described later in this literature review.

In all three experiments the modality principle was tested across different media and found to produce significant learning benefits. As described earlier in this literature review, the modality principle states that when presenting a multimedia explanation (visual and verbal information), words should be presented auditorily rather than visually (i.e. spoken rather than written). This was tested in three
different studies of three different multimedia environments:

1. A desktop multimedia environment
2. Two agent-based multimedia games
3. Virtual reality environments

2.3.2.1 Experiment #1: Desktop Multimedia Environment. Please note that Experiment #1 actually consisted of two experiments: The same experiment was conducted with two different kinds of subject matter. These two "sub-experiments" are referred to as Experiment 1A and 1B for clarity.

Mayer and Moreno (1998) designed Experiment 1A as a straightforward test of Baddeley’s dual-processing theory, described earlier. Among other things, Baddeley’s theory holds that working memory is comprised of auditory working memory (the phonological loop) and visual working memory (the visuospatial sketchpad), and that each working memory store has limited capacity. Connections between these memory stores can be made only if corresponding pictorial and verbal information is present in working memory at the same time.

To test the impact of modality learning, Mayer and Moreno (1998) compared learning outcomes between two groups of students: Group AT (animation and on-screen text) and Group AN (animation and concurrent auditory narration using the same words given to Group AT as on-screen text). Mayer and Moreno conducted two different experiments (1A and 1B), both of which employed Group AT and Group AN: In Experiment 1A students viewed animation depicting the formation of lightning, and in Experiment 1B students viewed animation depicting how a car's braking system worked. Based on dual-processing theory, Mayer and Moreno hypothesized that students in Group AT would perform more poorly than students in Group AN.
on retention (retaining the steps in a cause-effect chain), matching, and transfer tests.

2.3.2.2 Experiments 1A and 1B: Results Summary. Please note that differences between the groups were significant in all cases. For the three dependent measures (i.e. retention, matching and transfer) in both experiments Mayer and Moreno found:

- A split-attention effect on verbal recall: Group AN students tended to recall more relevant idea units than Group AT students did.

- A split-attention effect on visual-verbal/picture-name matching: Group AN students tended to correctly match more items than Group AT students did.

- A split-attention effect on problem-solving transfer: Group AN students generated more solutions than did Group AT students.

Mayer and Moreno’s 1998 study provides robust results because:

- It was based on three different dependent measures across two experiments.

- The two experiments controlled for any possible influence subject matter (lightning formation vs. car braking) may have had.

- Despite the subject matter differences between Experiments 1A and 1B (lightning formation vs. car braking), the experiments can be compared to one another because they both tested scientific, cause-and-effect material.

- The same words were used in narration and text: There was no difference in content, only in modality.

Mayer and Moreno extended Mousavi et al.’s paper-based results to a computer-based multimedia environment and provided empirically based evidence that learners
fare better when words and pictures are presented in separate modalities: When both are presented visually, working memory is overloaded. Mayer and Moreno’s (1998) study also provided important points for expansion in the study presented as part of this thesis, which:

- Used a longer learning episode (45 minutes on average), per Mayer and Moreno’s note that their learning episodes were brief;
- Allowed hands-on interaction (or rather, avatar hands-on interaction) with more active how-to content rather than more passive how-it-works content;
- Expanded narration in the form of a live human voice rather than a recorded narration;
- Studied modality in a social situation, with the opportunity for interaction between other students rather than one student viewing one animation alone;
- Primarily studied adults, not only college students: Though three college students were part of the final study sample, they were not the only participants.
- Similar to Mayer and Moreno’s Group AT and Group AN, the study presented here was comprised of three groups: Group V, Group T, and Group VT. Moreno noted that simultaneous modes had not been tested and this study does exactly that.

Experiment #1 was just one study by Mayer and Moreno that showed narration and pictures were more effective for learning than text and pictures. Mayer and Moreno went on to conduct other studies that reached beyond the desktop multimedia environment, discussed next.

2.3.2.3 Experiment #2: Two Agent-Based Multimedia Games. In Experiment #2, Mayer and Moreno (2001) investigated the effectiveness of a science
simulation game called Design-a-Plant, in which an on-screen animated agent interacts with learners by asking questions and providing guidance (Moreno, 2006). Mayer and Moreno (2001) wanted to see if the modality principle, which states that words should be presented auditorily rather than visually (i.e. spoken rather than written), would apply to the Design-a-Plant game and whether delivering the instructional message via an animated agent would help learning (Moreno & Mayer, 2002).

Mayer and Moreno varied whether the animated agents words were presented as auditory speech or as written on-screen text. Test participants were divided into the following four groups shown in Table 2.4:

<table>
<thead>
<tr>
<th>Image &amp; Text</th>
<th>Image &amp; Narration</th>
<th>No Image &amp; Text</th>
<th>No Image &amp; Narration</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>17</td>
<td>16</td>
<td>16</td>
</tr>
</tbody>
</table>

2.3.2.4 Experiment #2: Results Summary. Participants performed better on tests of retention and problem-solving transfer when words were presented as speech rather than on-screen text (i.e. the Image and Narration Group and the No Image and Narration Group performed better than participants in the Image and Text and the No Image and Text Groups). These results supported Mayer and Moreno’s modality principle.

2.3.2.5 Experiment #3: Virtual Reality Environments. Experiment #3 is particularly relevant to the study conducted for this thesis. Moreno and Mayer wrote that “An important research issue is whether more immersive media result in different learning outcomes than less immersive media” (Moreno, 2006). Higher levels of immersion may promote a higher sense of presence, which in turn may promote more engagement and deeper learning in students who learn by participating in the
learning task with a higher sense of being in the environment may learn more deeply than students who feel more like observers (Moreno & Mayer, 2002). To test this research issue, Moreno and Mayer (2002) conducted two experiments, subsequently referred to as Experiment 3A and 3B.

In Experiment 3A, the authors varied the degree of media immersion by delivering the Design-a-Plant game (the same used in Experiment #2, described above) in three different ways:

1. Desktop (low immersion)
2. Head-mounted display (medium immersion)
3. Head-mounted display with walking (high immersion)

Moreno and Mayer also tested the modality principle in Experiment #3, which created the test groups, immersion and modality conditions described in Table 2.5:

Table 2.5. Group Conditions and Numbers of Participants from Experiment #3A by Moreno and Mayer

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Low Immersion</th>
<th>Medium Immersion</th>
<th>High Immersion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narration</td>
<td>17</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Text</td>
<td>17</td>
<td>13</td>
<td>14</td>
</tr>
</tbody>
</table>

After the instructional program concluded, participants were asked to indicate their perceived sense of presence with a Presence Questionnaire (PQ) created by Witmer and Singer and to answer a set of retention and problem-solving transfer questions (Moreno & Mayer, 2002). Participants in the study described here completed highly similar activities: Following the 3D building class, participants completed Witmer and Singer’s Presence Questionnaire and answered questions to evaluate cognitive load and retention (see Appendices F-I).
2.3.2.6 Experiment 3A: Results Summary.

- Moreno and Mayer predicted that higher levels of immersion would create a higher sense of presence, and the results of Experiment 3A supported this: Participants in the Desktop groups (D) reported a significantly lower sense of presence than did the participants in both head-mounted (H and W) display groups. Media environment affects perceptions of presence.

- Groups presented with a higher level of immersion (Groups H and W) did not differ in the number of items they recalled from groups presented with a lower level of immersion (Group D). Immersion (media environment) did not affect responses to cognitive assessment questions.

- Receiving information as on-screen text rather than narration significantly hindered learning. Groups presented with verbal information as speech recalled significantly more information than those presented with verbal information in the form of text. Participants in one immersion condition did not outperform others in a different immersion condition in this regard: Presentation of information as on-screen text hindered learning in all groups.

- Participants in more immersive conditions (Groups H and W) reported higher levels of presence, but groups did not differ in their learning outcomes. Immersion and presence had no affect on learning outcomes.

In Experiment 3B, Moreno and Mayer made just two changes:

1. They removed one immersion condition (Group H, the head-mounted display without the walking condition), keeping only the most (W) and least (D) immersive media conditions used in the previous experiment (3A).
2. They added a redundant communication modality, narration and text together (Group NT).

Table 2.6. Group Conditions and Numbers of Participants from Experiment #3B by Moreno and Mayer

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Low Immersion</th>
<th>High Immersion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narration (N)</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>Text (T)</td>
<td>14</td>
<td>11</td>
</tr>
<tr>
<td>Narration + Text (NT)</td>
<td>14</td>
<td>12</td>
</tr>
</tbody>
</table>

The experimental procedure was identical to that used in Experiment 3A (described above).

2.3.2.7 Experiment 3B: Results Summary.

- As in Experiment 3A, participants who used a head-mounted display with walking (W) reported a significantly higher sense of presence than those using the less immersive desktop display (D). Media effects perceptions of presence.

- As in Experiment 3A, media did not affect learning. Different immersion conditions (of D and W) did not affect performance on tests of retention and transfer.

- The Narration (N) groups outperformed Text (T) groups on retention, transfer and/or program ratings scores.

- The Narration (N) and Narration and Text (NT) groups did not differ from each other on retention, transfer or program ratings scores. Adding redundant on-screen text to spoken explanations did not help or hurt students learning. Moreno and Mayer suspect that NT participants may have simply paid attention to narration alone because, when exploring a virtual environment, they are less likely to read a box containing text information that they can instead obtain by listening (Moreno & Mayer, 2002).
2.3.3 Social Agency Theory. The main thesis in social agency theory is that social cues in a multimedia message can prime the social conversation schema in learners. Once the social conversation schema is activated, learners are more likely to act as if they are in conversation with another person. Thus, at least to some extent, the social rules of human-to-human communication come into play (Mayer et al., 2003). Social agency theory was originally described in the context of human-to-computer interaction as opposed to human-to-human (or avatar-to-avatar) communication. Mayer et al. (2003) observed that once learners interpreted their interaction with a computer as social the rules of human-to-human communication came into play, making learners try harder to make sense of what the computer is saying by engaging in deep cognitive processing (Mayer et al., 2003).

When considering the relationship between learning and communication, perhaps even more interesting is Mayer’s theory that a learner can interpret a multimedia learning episode as either a case of information delivery or a case of social communication (Mayer et al., 2003). Further, the learner’s interpretation of the episode—as social communication versus information delivery—influences the type of schemas that are activated in the learner, the type of cognitive processing that occurs during learning, and ultimately the quality of the learning outcome (Mayer et al., 2003). In short, strong social cues support good performance on transfer tests and weak social cues lead to poor performance on transfer tests. Please note that this theory only applies to transfer tests and not necessarily to other styles of learning measures, like retention.

Despite its origin in studies of human-to-computer communication, social agency theory is relevant to the study conducted for this thesis, which sought to identify and understand both cognitive and communication differences in Group V, Group T, and Group VT participants. A major question about voice (Group V and
Group VT) and text only (Group T) participants is whether both voice and text (or one or neither) prime the social conversation schema: Does the presence of voice in a social virtual world, for instance, make learners more likely to act as if they are in conversation with another person? The availability of voice communication does not, of course, guarantee that learners will want to use it, or that it will prompt them to speak with others more than they might with only text communication available: Voice availability and one’s desire to speak with others are not the same or necessarily linked. Does voice use correlate with deeper cognitive processing and thus better results on post-activity cognitive tests? Or, does text act as speech (Baron, 1998), thus priming the social conversation schema to a similar degree the human voice does? Is social agency theory completely irrelevant where human-to-human communication and learning in a social virtual world are concerned? It is hoped that the study results presented here will provide some hints at answers to these questions.

Finally, it is worth noting that Mayer describes how the social agency model is consistent with wider research in discourse processing and classroom teaching (see Mayer et al., 2003, p. 419-20 for a brief review of this literature). This may imply that social agency theory is relevant in human-to-human (voice or text based) communication, instruction, and learning in a social virtual world, which more closely resembles classroom teaching than it does multimedia teaching via computer animation.

2.3.4 Information Equivalency Hypothesis. In contrast to the split-attention and dual-processing aspects of cognitive load theory, the information-equivalency hypothesis predicts no differences between groups that have concurrent auditory narration presented with corresponding visual diagrams, and groups that have printed text presented concurrently with corresponding visual diagrams (Mayer & Moreno, 1998). The hypotheses tested as part of this dissertation are stated in opposition to the information equivalency hypothesis, acknowledged here to ensure contrasting
theories and perspectives have been identified and considered.

2.4 Presence and Co-Presence

So far, this review of the literature has presented theories of the structure of working memory (Baddeley, 2002; Clark & Paivio, 1991), theory on the impact of modality on working memory and information processing streams (Penney, 1989), and cognitive load theory (Sweller, 1994). It has also reviewed studies that focused specifically on learning and modality, which this study aimed to expand. The final section of the literature review, presented next, more fully introduces the two remaining, key theoretical concepts to the dissertation study presented here: Presence, one’s sense of “being there” in a virtual world, and co-presence, one’s sense of “being there with others.” This section concludes the review of the literature, which will have covered research on cognitive load, learning, modality and presence—all of the dependent and independent variables in the study presented here—by its end.

Presence and co-presence are considered two major factors that differentiate virtual worlds from other online and multimedia applications, and are attributes on which a virtual world’s success is measured. According to Mikropoulos and Strouboulis (2004), for example, “Presence is the main attribute of virtual reality (VR) that differentiates it from other information technologies, giving learners an active role.” Slater et al. (2001) more succinctly explain why researchers bother to study presence at all: If humans are required to perform tasks within virtual environments, then surely it is beneficial for them to feel present in the environment in which the task is taking place.

Co-presence is considered an attribute of presence (i.e. a child to the parent concept of presence), but it is important to discuss co-presence as a distinct concept because it is very similar to social presence, and “satisfaction with entertainment
systems and with productive performance in collaborative virtual environments is based largely on the quality of the social presence they afford” (Biocca et al., 2003). Co-presence, of course, is not inherently a positive thing: With increased co-presence and highly social interactions can come negative attributes and experiences like biases, manipulation, and deception (Biocca et. al. 2003), just as in real life.

2.4.1 Definitions of Presence. Research on spatial presence (also known as “physical presence”) explores the sense of being in a virtual place (Biocca et. al. 2003). Biocca et al. (2003) uses the terms “telepresence,” “spatial presence,” and “physical presence” interchangeably, with each term describing the phenomenal sense of “being there” Biocca et al. (2003). A sense of “being in” a virtual space includes automatic responses to spatial cues (flinching in reaction to an oncoming bird in flight, for example) and mental models of mediated spaces (moving through a door and ascending a stairway) that create the illusion of place. For the purposes of this dissertation, only mediated spaces (social virtual worlds) were studied, not real world ones.

Unlike Biocca et al. (2003), Sheridan (1992) distinguishes between presence and telepresence: Presence is the sense of being within a computer-generated world (like Second Life) and telepresence is the sense of being at a real remote location (like a remote work site or battlefield) instead of a computer terminal. Like Biocca et al. (2003), Witmer and Singer (1998; 2005), equate presence and telepresence in their definition of presence, which they describe as the subjective experience of being in one place or environment even when one is physically situated in another (whether the environment is a virtual world or a real remote location does not matter). As applied to virtual environments specifically, however, presence refers to experiencing the computer-generated environment rather than the actual physical locale (Witmer & Singer, 1998).
The concept of a continuum can help us understand higher and lower degrees of presence, touched on briefly here. Schroeder (2006) points out that the difference between desktop and immersive virtual environments ("immersive" meaning "non-desktop" applications like VR caves and head-mounted displays) is "higher and lower degrees of presence... different technologies and their uses provide different degrees of absence or removedness from the physical world" (Schroeder, 2006). Earlier in this literature review, Moreno and Mayer’s (2003) findings (in Experiments 2 and 3) reinforced this statement: Varying levels of immersion correlated with desktop and head-mounted display environments.

The literature focused on presence includes many lengthy, ongoing debates about how to define and measure presence and co-presence. For the purposes of this study, Heeter’s (1992) definitions of presence are particularly helpful. He describes three kinds of presence:

1. Personal presence, a measure of the extent to which the person feels as if they are part of the virtual environment

2. Social presence, the extent to which a person feels that other beings (living or synthetic) also exist in the virtual environment

3. Environmental presence, the extent to which the environment itself acknowledges and reacts to the person or avatar in the virtual environment (Schuemie et al., 2001), which can be evaluated by measuring things like how quickly a scene is drawn (rendering speed), time between user action and environment response or feedback, lag (delay) and so on.

In summary, presence and social presence are psychological states that reflect the perceptions and feelings of people using the virtual environment, while environmental presence is an evaluation of the virtual environment itself and largely focused on its
technical performance, not on people using the environment. According to Biocca et al. (2003), “social presence refers to the sense of “being there together with another.””

A sense of being with others includes things like primitive responses to social cues (laughter, fear, panic) and automatically generated models of the intentionality of others (whether people, animals, agents, gods, and so on). In virtual environments, the others that we experience are primarily technologically mediated representations of other humans, including 3D avatars (Biocca et al., 2003) (as in Second Life).

In Heeter’s (1992) definition, social presence is one kind of presence. As Schuemie et al. (2001) noted, social presence is a special case of interaction because it involves interaction between two or more users. Despite this seemingly special or distinctive interaction, social presence is often considered part of presence, a component of the broader or umbrella term “presence.”

What is the difference between social presence and co-presence, anyway? It is not easy to pinpoint because the two terms are often referred to synonymously in the literature. Biocca et al. (2003) provided a hint, at least, by distinguishing between the two terms for primarily historical reasons (Biocca honors definitions from past work on presence in the real world on unmediated interactions). Biocca et al. (2003) note that, in the literature of unmediated interactions that take place in the real world, social presence is often a binary term “because social presence is treated as self-evident: The other simply is or is not present” (Biocca et al., 2003). In the real world, it’s easy to tell if another person is or is not in a room with you under normal circumstances. In Biocca et al.’s (2003) definition of social presence, then, social presence is not seen as a continuum but as binary: Another person is either there or not there, unambiguously.

Biocca et al. (2003) also remind us that the concept of co-presence is older than we might think, hearkening back to the work of Goffman (1959). Goffman
(1959) coined the term “co-presence” and emphasized the role of each of the five senses. According to Biocca et al. (2003), Goffman’s emphasis on the senses has become especially important in mediated interactions because the senses of the user are extended (or, it’s worth adding, potentially blocked) to some degree by technology. The representation of another person triggers a sensory impression of the other that exists on a continuum from the minimal to the intense (Biocca et al., 2003). Although social presence does, when considered in the context of mediated (instead of unmediated) environments, become synonymous with co-presence in regard to being there together, co-presence was included in this study for its continuum properties: Participants in this study subjectively measured co-presence on a Likert-style scale via Witmer and Singer’s (2005) Presence Questionnaire.

To summarize, then, the primary difference between social presence and co-presence is a binary state (for social presence in the real world, with others obviously present or not) or a continuum (for co-presence in virtual worlds, with others present at varying degrees). As Schroeder (2006) tells us, virtual environments provide people with the sensory experience of being in a place other than the one they are physically in, of being able to interact with that place, or of simply being there. Co-presence can then be defined as being there together. The study presented here employed, as did Mayer and Moreno (2003), the presence definitions and measures of Witmer and Singer (1998; 2005), described in the following section.

2.4.1.1 Witmer and Singer’s Measures of Presence. The definitions of presence and co-presence reviewed in the previous section still beg the question of “What exactly do we measure when we measure presence?” Since presence and co-presence have been defined, this section describes how the study presented here evaluated them using measures created by Witmer and Singer (1998; 2005) and Gerhard et al. (2001). Measures for presence are often based on the expected results of presence, so it can
be difficult to talk about one (i.e. measures or results) without also referring to the other.

According to Witmer and Singer (1998; 2005), presence is a subjective sensation or mental manifestation that is not easily amenable to objective physiological definition and measurement, like monitoring brain activity or heart rate, for example. Subjective report by participants is the essential basic measurement of presence (and was also used in the study presented here). As Schuemie et al. (2001) note, the most commonly used measures in presence research are based on subjective ratings through questionnaires. As noted in the definitions of presence provided in the previous section, presence usually refers to either attributes of people (feelings) or of the environment (usually technical responsiveness). Witmer and Singer believe that the strength of presence experienced in a virtual environment varies both as a function of individual differences and the characteristics of the environment itself:

Individual differences and abilities can enhance or detract from the presence experienced in a particular virtual environment (VE). Various characteristics of the VE may also support and enhance, or detract and interfere with, the presence experience. Hence, presence measures should assess these individual differences as well as characteristics of the VE that may affect presence (Witmer & Singer, 2005).

Witmer and Singer’s synopsis is certainly true of experiences in social virtual worlds, including Second Life. Because Witmer and Singer consider both individual differences and environmental characteristics in their Presence Questionnaire (Appendix F) and because this study builds upon the work of other researchers who have used Witmer and Singer’s Presence Questionnaire, the Presence Questionnaire was used in this study. Further, the results and measures described in this section pertain both to users and VE’s themselves, thereby incorporating Witmer and Singer’s (1998; 2005) definition of presence.
Witmer and Singer’s (1998; 2005) work on identifying and measuring factors that contribute to presence was among the earliest to cluster variables affecting presence into meaningful groups of similar items (Witmer & Singer, 2005). They (1998) initially produced and empirically tested four clusters or subscales of presence. All four factors reflect and measure both user perceptions (perceived ability and impressions) and the medium itself. They are:

- **Factor #1: Involvement**: The perceived ability to control events in a virtual environment; the responsiveness of the virtual environment to user-initiated actions; involvement in visual aspects and overall involvement in the virtual environment. A sample Factor #1 question is “How responsive was the environment to actions that you initiated or performed?”

- **Factor #2: Sensory Fidelity**: The degree to which the configuration of the virtual environment enables participants to sense different aspects of it, whether visual, auditory or haptic in nature. A sample Factor #2 question is “How much did the auditory aspects of the environment involve you?”

- **Factor #3: Adaptation/Immersion**: The perceived proficiency of interacting with and operating in the virtual environment, and how quickly the user adjusts to the experience. A sample Factor #3 question is “Were you able to anticipate what would happen next in response to the actions that you performed?”

- **Factor #4: Interface Quality**: Whether control or display devices interfere with or distract from task performance, and the extent to which participants feel they are able to concentrate on tasks. A sample Factor #4 question is “How much did the visual display quality interfere with or distract you from performing assigned tasks or required activities?”

In addition to the factors described by Witmer and Singer (1998), Schuemie
et al. (2001) discussed measures of presence that include task performance, responses and emotions (fear of heights and fear of public speaking to virtual audiences), and simulator sickness (positive and negative correlations between presence and nausea and dizziness). The study presented here, however, was not concerned with and thus not designed to evaluate measures of responses and emotions (for simple scope control) or simulator sickness, which is unlikely and not common in Second Life.

Thie and van Wijk (1998) discuss the additional presence factor of social cues but, in their study of 48 subjects with a desktop virtual environment, social cues had no significant effect on reports of social presence or presence. The authors suspect that this is because social cues were limited in the virtual environment: Users could not, for instance, pick their own avatar name or nickname, nor could they make gestures. Although Second Life provides gestures that can be deliberately invoked by its users, gestures do not necessarily respond naturally or immediately to social cues: The gesture of an avatar’s head nodding “yes,” for example, is cued by typing /yes and not by the human user nodding. Second Life is, like the environment in Thie and van Wijk’s (1998) study, limited in social cues. For this reason social cues were not evaluated or measured as part of this study.
CHAPTER 3

METHODS

This section describes the methods, instruments and analyses used to conduct the study presented here, which employed a single independent variable (communication modality via Group V, Group T and Group VT) and three dependent measures (presence, retention, and cognitive load). This section begins with an overview of the research design, statement of delimitations, and a description of online research methods employed, including study procedures, materials, a description of the target study population, and recruiting methods. Next, participant selection and assignment to condition (i.e. to Group V, Group T, or Group VT) and the characteristics of the final study population are described in detail. The Methods section concludes with a brief description of data coding and a review of the instruments used to collect data.

3.1 Hypotheses

The study presented here compared three communication modes (voice only, text only and voice and text concurrently) in Second Life, a social virtual world, to address the following hypotheses: $H^1$: Retention scores will be higher for voice participants (Group V) and voice-and-text participants (Group VT) than for text-only (Group T) participants.

$H^2$: Cognitive load will be lower for voice participants (Group V) and higher for text conditions, voice-and-text participants (Group VT) and text participants (Group T), as evidenced by subjective reports of mental strain and effort.

$H^3$: Voice will contribute to greater perceptions of presence.

$H^4$: Perceptions of presence will not correlate with deeper learning.
3.2 Research Design

Now that research questions have been defined and hypotheses stated, this section describes the empirical, quantitative approach taken to address them in detail.

3.2.1 Overview. One small pilot study and one main study were conducted from January through July 2009. The study was comprised of a consent form and screening questionnaire; instructional activity; a post-activity questionnaire; and data analysis, all of which are described in detail in this section and shown in Appendix A-I. The study design described in this section received approval from the IIT Institutional Review Board (IRB) in June 2008.

The research described here studied adult participants in an online, 3D environment (Second Life) to determine if the communication modality to which they were assigned affected cognitive load, perceptions of presence, and learning (specifically, retention of material covered in the instructional activity). All participants were randomly assigned within three groups of 20 participants each participant to one of three different communication modes: voice only (Group V), text only (Group T), and voice and text together (Group VT) and attended an Introduction to Building in Second Life class in an effort to discover which mode, if any, increased, decreased, or had no effect on cognitive load; which mode best supported learning as measured by retention tests; and whether communication modality had any impact on participants perceptions of presence.

3.2.1.1 Ethical Considerations.

In the interest of full disclosure, the author managed the implementation of integrated voice chat in Second Life and was a full-time employee at Linden Lab, creators of Second Life, from February 2007–August 2009. Though the author was obviously close to the voice project as a result of this, it is worth noting that no
voice research with Second Life residents had been designed or conducted to date, and no comprehensive voice user data had been collected or disseminated. Further, no research on communication (even text-based communication) in Second Life had been conducted.

That said, the author believes she had no prior knowledge that might have skewed or otherwise positioned her to expect or look for certain types of findings over others. In addition, research has indicated that researchers of games should play the games they are studying. If they do not, they cannot know what questions to ask, decipher the local language, understand the game mechanics, or have any sense of the social context of play (Williams, 2004). In this sense, at least, the author’s familiarity with Second Life may be considered critical to the research described here.

The addition of voice chat was a sensitive topic among some Second Life Residents at the time of its launch. Some Second Life Residents enjoy Second Life primarily as a fantasy environment happily separate from real life, and some also engage in adult activities while logged in. Accidentally speaking aloud, in one’s own voice, was a concern for Residents who did not want to betray their real identities or provide a means of identification that might extend to real life.

For these reasons, this study was conducted online and anonymously. Williams (2004) acknowledges that the inherently virtual nature of virtual environments represents a special challenge to academic researchers seeking original empirical data because it must often be conducted remotely and anonymously. Yet the same advantages and disadvantages of various research methods still apply (Williams, 2004). For the above reasons, online, anonymous research was deemed appropriate for the context, purpose, and focus of this dissertation, which was communication modality, learning, and perceptions of presence in virtual worlds. The solicitation of participants and data collection was conducted almost entirely online, both inside and outside of
Second Life.

3.3 Study Procedures

3.3.1 Overview. Study events took place in the following linear order between January and July 2009. Each step in the research process and its corresponding instrument or materials is described in detail in this section, immediately after this overview:

1. First, prospective study participants were referred to (and, one hopes, read) a study information web page (Appendix A).

2. Prospective study participants submitted consent forms (Appendix B), which followed the study information web page.

3. After submitting the consent form, potential subjects were directed to and completed a screening questionnaire (Appendix C).

4. Qualified subjects were notified by email (Appendix D) as to whether or not they had been selected for study participation. In this email, they were asked to attend a class in Second Life (the learning activity) on building 3D objects. The schedule of classes corresponded to the communication modality (test group) to which the participant was randomly assigned. In other words, a participant randomly assigned to the voice condition, Group V, received an email for a class to attend that was conducted with voice chat. The group assignment procedure is described in detail later in this section.

5. Using a typical computer set up (i.e. either a desktop and mouse, or laptop and mouse) participants attended one learning activity (Appendix E) in Second Life for approximately 45 minutes. During the Introduction to Building class, subjects learned and completed example exercises with an online instructor
(female avatar with female voice, kept constant throughout the study).

6. Following the Introduction to Building class, subjects completed an online questionnaire (Appendix F–H) with three sections on mental strain, presence, and retention (questions about material covered during the Introduction to Building class). First, participants completed two, seven-point, Likert-scale format questions that asked them to report perceived mental strain (Appendix H). Reports of mental strain were followed by a Presence Questionnaire (copyright Witmer & Singer, 1998), a seven-point Likert-scale format with 28 questions rating their experience in different areas of the 3D environment (Appendix F). After completing the online Presence Questionnaire, participants completed a learning (retention) assessment, a test of how well they remembered material from the learning activity (Appendix I).

A study information web page (Appendix A) was created at www.iit.edu/filiste (the author’s university homepage) to provide prospective participants with detailed study information and links to the consent form (Appendix B) and screening questionnaire (Appendix C). The web page also contained instructions, eligibility requirements, and IIT IRB approval and contact information. Pages were created in basic HTML and a database script automatically sent consent form signatures to an email account (studyinSL@gmail.com) used to hold them. The web page was moved online via IITs Hawk telnet/FTP. The website and email were free of charge (or at least included in the authors tuition payment).

Online surveys (the screening questionnaire and post-activity assessment) were powered by Qualtrics, a provider of academic survey software. A Qualtrics license for one calendar year was obtained for $90, paid for by the author. After the initial target sample size ($N = 63$) was reached, data was downloaded out of Qualtrics and into Microsoft Excel, where it was cleaned and cases (i.e. screening questionnaire results,
test condition, and post-activity survey) were linked together (all information on a particular participant combined into a single spreadsheet and located in the same row for that participant, or case). This spreadsheet information was then uploaded into SPSS for data coding (variable names, 0 for missing values, what right and wrong meant).

3.3.2 Data Collection.

3.3.2.1 Pilot Studies. Three pilot studies were conducted shortly after the study information web page was posted, with one pilot study conducted for each of the three conditions (V, T, or VT). Data for the pilot studies was not analyzed or retained, but was collected to determine if participants’ understood the questions being asked and to determine how long study completion would take.

The pilot studies resulted in small changes to the learning activity test script, and corrected a few trivial typographical errors in the post-activity survey instrument. As a result of the pilot studies, additional how-to statements were added to the learning script. The first such addition was instruction to participants about how to use the Camera Controls in Second Life, which change how Second Life users see and interact with the virtual world and 3D objects in it.

Camera Controls enable Second Life Residents to zoom in and out; pan left and right (the equivalent of walking around a scene or object in a circle); and to obtain an aerial, omniscient view of an area below or a bottom up/view, looking skyward. It was important to provide participants with a shared baseline level of knowledge about how to use Camera Controls because knowing how to use them could directly affect class performance: Camera Controls effect how difficult or easy it is to observe instruction, observe 3D objects, work with 3D objects, and generally build 3D content, which was the topic of the learning activity. Participants who already knew how to use Camera
Controls would have had a distinct advantage in successfully completing the learning activity, and potentially in retention: A limited view of 3D content could have affected participants’ responses to post-activity survey questions that involved choosing which 3D shapes (prims) to use to create an object, 3D rotation, and identification of 3D shapes in Second Life. A difference in participant knowledge of Camera Controls could have also affected reports of mental strain, quality of instruction, and reports of presence.

The second addition to the learning activity script, following the pilot studies, described how to use Second Life volume controls, how to mute oneself and other participants, and how to use Talk controls.

There can be a large degree of variability in participants’ computer system setups, even if all systems meet the minimum system requirements to run Second Life. Preferred volume levels and sound quality can vary with hearing ability, background and environmental noise levels in the participant’s physical location, and personal preference. Input and output tools can be comprised of varying combinations (and quality and manufacture) of headsets with or without microphones, internal computer microphones, speakers, and so on. Ensuring that participants knew how to hear others, knew how to hear them well by adjusting the volume at which participants heard others, and knew how to prevent themselves from sounding too loud to others ensured that V and VT group members in particular began their study sessions on more equal footing than they otherwise may have.

There was virtually no difference in the duration of the Introduction to Building classes: Regardless of modality, the class took approximately 40-50 minutes. The three pilot study participants were not included in the final sample (N = 60).

3.3.3 Target Study Population. Both new Second Life users (i.e. people who had
never used Second Life prior to the study) and existing Second Life users, known as Residents, were recruited for study participation. To this end, study participants were recruited from both the general online population and the Second Life population, in an effort to avoid a sample comprised primarily of university undergraduate students or Second Life Residents. The learning activity consisted of a 45-minute (on average) class on how to build 3D objects in Second Life.

A Second Life Resident is not necessarily an experienced creator of 3D content: The simple fact of having a Second Life account does not also mean that one knows how to build 3D objects in Second Life. One may very much enjoy flying around and exploring the virtual world, attending events, or shopping in Second Life without also knowing how to create content using the tools provided with the Second Life software. Of course, it is certainly possible that some existing Second Life users already know how to create 3D content and complete some or many of the activities taught during the Introduction to Building class. In order to acknowledge and account for the factor of Second Life experience, the screening questionnaire included three questions about experience, which asked if the participant had ever logged in to Second Life before; how long ago their Second Life account had been created; and how many total hours they had spent in Second Life. These questions were taken into account during the intercorrelations phase of data analysis, described later in this document.

3.3.4 Recruiting Materials and Methods. Study participants were recruited both offline and online. All recruiting materials, both online and offline, referred readers to the aforementioned study information web page (Appendix A) hosted on the IIT.edu domain. The web page described the study and outlined expectations for prospective participants, and also contained links to the consent form (see Appendix B) and screening questionnaire (Appendix C).

All recruiting materials and study communication informed participants that
they would receive $10 USD (or equivalent, if required in the form of another currency for international participants) as compensation following study completion. Study participants could choose to receive the $10 USD study compensation as either cash via PayPal; as an Amazon electronic gift certificate delivered by email; or as Linden Dollars, the virtual currency of Second Life, with which Second Life Residents purchase virtual goods for sale within Second Life.

The $10 USD compensation amount was reached after no study participants were successfully recruited when no compensation was offered and, later, when only three participants were recruited but failed to attend the study when $5 compensation was offered. The majority of study compensation was paid from the author’s personal income, and the IIT Lewis Department of the Humanities provided $140 worth of study compensation for participants who had taken part in the study through mid-May 2009.

To recruit from the general online population, ads were posted weekly on craigslist.org, a free online classifieds site, in the Volunteers section. Ads were posted in the Detroit, MI; Chicago, IL; Denver, CO; and San Francisco Bay Area geographic categories of craigslist. While any craigslist ad can conceivably be searched for across all geographic markets of the site, ads can only be associated with one geographic market at a time, per craigslist rules: “Please choose just ONE craigslist site for which your ad is most relevant - this should generally be the site closest to where you are located. If your ad is equally relevant to all locations, your ad does not belong on local craigslist sites at all, please find another service.” (craigslist, 2009)

Craigslist was used to recruit from the general online population because it is popular and accessed by a wide range of internet users. According to the Pew Internet and American Life Project, 49% of internet users have used online classified ads, or sites like Craigslist. Further, The number of online adults who have used online
classified ads has more than doubled in the past four years. Almost half (49%) of internet users say they have ever used online classified sites, compared with 22% of online adults who had done so in 2005. On any given day about a tenth of internet users (9%) visit online classified sites, up from 4% in 2005. These are just a few of the main findings of an April 2009 survey by the Pew Research Centers Internet and American Life Project survey.

Online ads were also posted to the Illinois Institute of Technology (IIT) online campus bulletin board, Bazaar, where anyone with an IIT e-mail account is eligible to post items on the site. Professors Susan Feinberg, Karl Stolley, and Kathryn Riley recruited undergraduate students during their Spring 2009 classes at the Illinois Institute of Technology. James Braman of Towson University also notified undergraduates about the study.

The social networking website, Facebook, was also used for study recruitment. The study was listed as an event by the author, with a possible immediate audience of 340 friends, academic and professional colleagues, and acquaintances, some of whom in turn posted the event to their own Facebook pages (which could in turn seen by their own networks) and/or to their Twitter feeds, the audience for which may be comprised only of friends or public (anyone who wants to read or subscribe to a particular person’s Twitter feed).

In addition, the study was listed as an in-world event in the Second Life classifieds. This ad told existing Second Life Residents when and where a study session (i.e. Introduction to Building class) was scheduled to take place, and referred them to the study information web page posted on the author’s IIT pages.

Offline, hardcopy fliers were posted on the bulletin boards of several libraries, specifically the Warren Public Library in suburban Detroit, MI and the Main branch
of the Detroit Public Library; the Main branch of the San Francisco Public Library; and the Rogers Park branch of the Chicago Public Library. Hardcopy fliers were also posted in a couple of restaurants and coffee shops, namely The Wake Up in St. Clair Shores, MI, and on community bulletin boards in Golden Gate Park, San Francisco, CA and on Castro St. in San Francisco.

One major omission in recruiting was the failure to track the origin of study participants: The screening questionnaire failed to ask how someone had heard about the study. For this reason, no quantitative data is available on which recruiting methods proved more and less successful. Anecdotally, the most successful recruiting method appeared to be the Second Life classifieds. After posting Second Life classifieds about the study, many more requests (primarily via IM in Second Life) for study validation (“Is this a real event?”) and details on how to participate were received than from any other recruiting venue.

3.3.5 Participant Anonymity. The information participants provided in order to participate in this study was completely anonymous: Real-world names were never provided because only avatar names were used, and any valid email address (which may give no hint of real world name or gender) could be used. It was hoped that the total anonymity of the participants’ real-world identities would decrease their potential fear or concern about associations between online and real life identities being made, and provide a safe environment that could support honest responses. Anonymity was specifically mentioned in the consent form (Appendix B).

3.3.6 Final Study Sample. This section describes how study participants were selected for participation and reports initial and final participation numbers.

All participants were required to submit a consent form (Appendix B) and screening questionnaire (Appendix C). The consent form required participants to
swear that they were 18 years of age or older. Though the screening questionnaire contained 25 questions, responses to four questions were considered the most important:

1. Prospective participants had to agree to accept the test condition (communication modality) to which they were assigned, and specifically if they would agree to use spoken voice communication: “If you are assigned to use voice in the Introduction to Building in Second Life class, will you accept assignment to your condition (either voice, text, or both)?”

2. They had to meet the minimum system requirements (asked across two questions about connection speeds and RAM): 1) How do you usually connect to the Internet? Choose one: a) Cable modem, b) DSL, c) telephone line/modem, d) T1, e) Other (please describe), and 2) How much memory does your computer have? Choose one: a) 16 MB b) 56 MB c) 512 MB or more d) 1 GB or more e) I dont know f) I dont understand the question.

Otherwise, Second Life may not have been able to run on their computers, and it was easier to disqualify these participants at the outset.

3. They had to have a high degree of availability (i.e. be available at many dates and times listed in the screening questionnaire) for study participation. The screening questionnaire asked respondents to check boxes beside dates and times they were available:

You will be asked to attend an Introduction to Building class in Second Life as part of this study. Please say when you will attend by selecting a class time and date below. If you are available on multiple days, please feel free to give more than one answer.

- Monday, May 11 at 6 PM Pacific / 8 PM Central / 9 PM Eastern
- Saturday, May 16 at 10 AM Pacific / 12 PM Central / 1 PM Eastern
- Saturday, May 16 at 4 PM Pacific / 6 PM Central / 7 PM Eastern
- Sunday, May 17 at 10 AM Pacific / 12 PM Central / 1 PM Eastern
- Sunday, May 17 at 4 PM Pacific / 6 PM Central / 7 PM Eastern
- None of the above/Other (Please suggest another date and time below)

Only high availability from participants, indicated with availability at multiple days and times, would allow for a truly stratified random assignment to condition. Otherwise, if based on limited availability, the sample would have been a convenience sample, based on when a participants could attend. As its name implies, a convenience sample would have certainly been much more convenient, and would have led to faster study completion but also weaker claims. In a random sample, each individual in the population has an equal probability of being selected. Nonprobability samples are those in which respondents are chosen based on their availability: With randomization, a representative sample from a population provides the ability to generalize to a population (Creswell, 2003), while a convenience sample does not.

In all, 149 prospective participants completed the consent form and screening questionnaire.

3.3.7 Participant Selection and Assignment to Condition. Table 3.1, and the paragraphs immediately following, describes the criteria for selection for study participation, identify screening cut points, and note the number of participants eliminated based on each cut point and the order in which they were eliminated.

The first cut point was agreement to participate in assigned condition. Of the 149 respondents, 24 did not agree to participate given the assigned test condition (”If you are assigned to use Voice in the Introduction to Building class, will you agree to use voice and speak aloud to other avatars?”) and were removed from the pool of
possible participants. All 24 people received an email stating that they had not been selected for study participation and thanking them for their interest. This left 125 participants.

The second cut point determined whether or not participants’ computer environments met minimum Second Life system requirements. Two people reported they usually connected to the Internet with a dial-up modem and were removed from the pool because a slow connection could have impeded their ability to participate (to stay connected to and/or not experience too much lag in Second Life) in the learning activity. In addition, one person reported having 16 MB or less of RAM, and three people reported having 56 MB or less of RAM. All four of these participants were removed from the pool because the minimum system requirement for running Second Life is 512 MB of RAM. In all, six people were omitted from the sample at this point.

This left 119 prospective participants who met all of the minimum requirements (willingness to participate in the assigned condition and met computer system requirements) for study participation. All 119 participants received an email notifying them of study acceptance (Appendix D), sent to the email address they’d provided

<table>
<thead>
<tr>
<th>Reason for Elimination</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total recruited</td>
<td>149</td>
</tr>
<tr>
<td>Did not agree to accept assigned condition</td>
<td>24</td>
</tr>
<tr>
<td>Did not meet minimum system requirements</td>
<td>6</td>
</tr>
<tr>
<td>Did not attend scheduled study session (no-show)</td>
<td>46</td>
</tr>
<tr>
<td>Attempted to but could not attend or complete study session</td>
<td>10</td>
</tr>
<tr>
<td>Total study participants</td>
<td>63</td>
</tr>
<tr>
<td>Removed due to missing data or as outliers</td>
<td>3</td>
</tr>
<tr>
<td><strong>Final study sample (N)</strong></td>
<td><strong>60</strong></td>
</tr>
</tbody>
</table>
on the screening questionnaire. None of the study acceptance emails bounced or were otherwise noted as undelivered to the author.

Of the 119 prospective participants, a total of 56 ultimately did not attend the learning activity at the scheduled time and date. The majority (41 people) simply did not show up at all and did not respond to the initial study acceptance email, or to one follow-up email asking if they could attend at another time and date. For this reason, no information is available on why they did not attend or any factors that may have prevented successful participation.

Several participants (10 people) very much wanted to participate, as stated in email correspondence and IMs, but had system issues: They could not stay connected to Second Life and could be seen logging on and off rapidly, making a connection and losing it immediately; were unable to successfully download and install the software; or had Second Life crash too frequently to stay connected. Despite extensive troubleshooting (and with a Linden Lab employee at the time, as was the author) these 10 individuals were unable to participate. It is possible that other prospective participants experienced similar issues, but this is pure speculation: There is no evidence of this as there was no communication about it.

In addition to the 10 people who were willing but unable to participate, three participants canceled by email when personal schedules intervened and did not reschedule or respond to follow-up email. This left a total sample size of 65, and 65 people total participated in the study. The next section describes how the final 65 participants were assigned to condition.

Later, during the data coding and analysis phase, three more cases were removed due to incomplete or missing data and two were removed as outliers. This left a final, total sample size of 60 participants, the bare minimum required to achieve
the desired power for statistical analysis.

### 3.3.7.1 Participant Assignment to Condition.

The study used a stratified random sample controlled for age and gender. While software does a wonderful job of creating random samples, assignment to condition took place through a very manual process that relied heavily on the use of Post-It notes and paper. Fortunately, the results of a Chi-square (shown in Table 3.2) subsequently indicated that assignment to condition was, indeed, random.

<table>
<thead>
<tr>
<th>Participant Attribute</th>
<th>Value</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.338</td>
<td>4</td>
<td>.855</td>
</tr>
<tr>
<td>Gender</td>
<td>1.481</td>
<td>2</td>
<td>.477</td>
</tr>
<tr>
<td>Race</td>
<td>.492</td>
<td>2</td>
<td>.782</td>
</tr>
</tbody>
</table>

*Note: Significant at the p <0.05 level.*

Each person had an assigned condition, as did each scheduled study session: The Introduction to Building classes were conducted using either a Voice (V), Text (T), or Voice and Text (VT) communication modality. Everyone attending a particular study session was automatically “in” their test condition simply by using the communication modality provided and stated at the beginning of class (“We won’t be using voice chat in today’s class so be sure to mute your mics if you usually use voice.”).

As consent forms and screening questionnaires arrived and were subsequently accepted, two fields were examined first: The participant’s avatar name and his or her answer to the question “What is your real-world gender (Male or Female)?” As a proud holder of an undergraduate Womens Studies Certificate, the author admits to having written the avatar name and email address of those who reported male on a
blue Post-It note and those who reported female on a pink Post-It note. The Post-It notes were then folded in half so the name and email address of the respondent were not visible and placed into two color-coded piles (one blue pile, one pink pile). The Post-It notes were then subdivided into roughly equivalent group sizes, usually of three to five people (Post-Its) per group. Each group was then randomly labeled V, T, or VT to indicate that all members of that group had been assigned to that condition. The underside of each Post-It note was also labeled with the assigned condition, just in case the clusters of Post-It notes stuck to the dining room table were rearranged or moved.

Next, scheduled study sessions were randomly labeled V, T, or VT: The Saturday study session at 10 AM became a voice class, for example, received a ”V” label beside it in a day planner, and so on for the VT and T conditions. This process designated what type of session each class would be and who would be in that session, based on the group (condition) to which they’d been randomly assigned and the condition randomly assigned to the study session. Then, the Post-It notes were opened and the avatar name of each participant was written in a notebook used to track group assignment, any communication with participants like changes in availability, and so on. After jotting this information down in one place, email messages (Appendix D) were composed and sent to participants, telling them the date and time at which their session would take place, what to do in advance, how to locate the instructor in Second Life, and so on.

A single email was sent to all the members of a particular group (Appendix D). A “dummy” email address created for the purpose of study-related communication (studyinSL@gmail.com) was used in the “To” field, while participant email addresses were kept private by being placed in the BCC field. This made study-related communication much easier to track and keep up with.
The review, sort, assign and communicate process described above was conducted on an ongoing basis from January 2009 through mid-July 2009. An exception to this was the month of May 2009, when a family emergency created too much day-to-day uncertainty for any study sessions to be reliably scheduled and attended by the instructor.

Nearly all study sessions were run on Saturdays and Sundays, usually with two sessions (at 10 AM and 4 PM Pacific) offered each day for a total of four possible weekend sessions. Online and offline ads were posted early in the week, usually no later than Wednesday, with group assignment and email communication to participants about acceptance and details (Appendix D) or rejection were sent by Thursday or Friday. Reminders about upcoming study sessions were sent 24 hours in advance of the scheduled session if the initial email had been sent 36 hours or more in advance of the session.

This shortened cycle (i.e. less than one full week from ad posting, to submission of consent forms and screening questionnaires, to acceptance, group assignment, and completion of the study session) improved attendance rates. Longer cycles, in which the screening questionnaire enabled participants to state availability for a study session 10-14 days in the future, resulted in high numbers of no shows. Longer cycles allowed too much time to pass and made it easier for participants to forget about a commitment, make better plans, or just have other real life events and tasks arise. The small number of people who responded to and qualified for the study on a weekly basis meant that no shows had an especially high impact: If even 1-2 people did not show up in a total group size of four people, the group size would be cut in half and no longer kept constant with other group sizes.

As in the pilot studies, there was virtually no difference in the duration of the Introduction to Building classes: Regardless of modality, the class took approximately
40-50 minutes.

3.3.7.2 Description of Independent Variable. This section briefly describes the assigned modality conditions further, by explaining what it meant for a class to be conducted in Voice, Text or both modalities concurrently. Two points are especially noteworthy. First, there was a certain lack of control over the participants condition: Communication modality could only be controlled (by the instructor) for the Introduction to Building class, not for all possible participant communication (a private conversation with another Second Life Resident, for example). If a Resident were assigned to Group T for the Introduction to Building Class, for instance, it is possible that they could opt to use voice communication in a private call with another Resident before, during or after the class.

This is because Second Life, by default, makes both voice and text communication available to all Residents at all times. Each Resident decides whether they will use voice communication, text communication or both depending on their personal preferences and communication needs. The Second Life viewer (software) is not officially available in voice-only or text-only configurations (though, since the viewer is open source, a programmer might choose to create such a version of the viewer). This means that the only reasonable and realistic way to control the modality condition for this study was through the instructor: The instructor used only voice, text or both modalities to conduct the class and asked participants to follow (which they appeared to do, in that no one was heard speaking during a text-only session, for instance).

Technically, however, it was possible for participants to communicate however they wanted, and participants used their assigned modalities to varying degrees. Some participants in voice sessions (Group V and Group VT) clearly had voice enabled (indicated by a white dot floating above avatars’ heads) and spoke aloud, while others listened but did not speak aloud much or at all. Some participants in Group V and
Group VT listened to the instructor’s voice instruction but combined modalities when asking questions, for example: If others were speaking aloud, a participant with a question might type it into the Local Chat window rather than interrupt the voice conversation, but might later ask another participant a question using voice chat. Obviously, this behavior cannot be neatly attributed to or explained by assigned modality condition alone, as typing rather than speaking may also indicate politeness in the form of a desire not to interrupt others and an awareness of the need for turntaking.

Second, the VT condition was simultaneous (concurrent) and not linear. Participants assigned to Group VT, for instance, did not experience some or all of the Introduction to Building class in the voice modality and then in the text modality, or vice versa. Instead, Group VT participants experienced both modalities concurrently. The instructor delivered the Introduction to Building class in two modalities simultaneously by pasting a full sentence into the IM window (from the learning activity script in Appendix E) as the sentence was spoken. Anecdotally, this created high cognitive load for the instructor who spoke from the script while copy-and-pasting sentences into the Local Chat window and, while providing instruction, received and had to respond to questions from students delivered in both voice and text chat. While difficult to manage, this concurrent VT condition was the best imitation to the work of Moreno and Mayer possible with the typical, default design of the Second Life viewer.

The instructor avatar was female, the instructor voice was female, and both avatar and instructor voice were kept constant throughout the study. The instructor script (Appendix E) was given by the same female instructor throughout the study in either voice chat, text chat or both concurrently.

When participants arrived in Second Life at the scheduled time the instructor
went through the test script. Participants sometimes asked questions before class started, usually about how to do something: Fly, control volume, adjust their microphone output volume, and so on. Even if individual questions were answered before class began, the test script was given in full so all study participants received the same information.

3.3.8 Final Study Sample Description.

3.3.8.1 Gender, Age and Race. Of the N = 60 total participants in the final sample 37 (62%) were female and 23 (38%) were male.

The ages of participants ranged from 18 to over 80 years of age, as detailed in Table 3.3 below:

<table>
<thead>
<tr>
<th>Age (Years)</th>
<th>Number of Participants</th>
<th>Percent of Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>18-24</td>
<td>6</td>
<td>9.7%</td>
</tr>
<tr>
<td>25-35</td>
<td>17</td>
<td>29%</td>
</tr>
<tr>
<td>36-45</td>
<td>12</td>
<td>21%</td>
</tr>
<tr>
<td>46-55</td>
<td>13</td>
<td>21%</td>
</tr>
<tr>
<td>56-65</td>
<td>8</td>
<td>12.9%</td>
</tr>
<tr>
<td>66-75</td>
<td>1</td>
<td>1.6%</td>
</tr>
<tr>
<td>75-86</td>
<td>3</td>
<td>4.8%</td>
</tr>
<tr>
<td>87+</td>
<td>0</td>
<td>0%</td>
</tr>
</tbody>
</table>

Note: These are the original age brackets including in the Screening Questionnaire (Appendix C). During the data coding and analysis phase, age brackets were compressed from the original eight (above) to three groups (see Table 3.4 below).

Because some of these groups had too few respondents to provide useful analysis, the groups were collapsed as follows later in the statistical analysis phase, as follows:
Table 3.4. Study Participants by Compressed Age Group (N = 60)

<table>
<thead>
<tr>
<th>Age (Years)</th>
<th>Number of Participants</th>
<th>Percent of Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;35</td>
<td>23</td>
<td>38%</td>
</tr>
<tr>
<td>&lt;55</td>
<td>25</td>
<td>42%</td>
</tr>
<tr>
<td>55+</td>
<td>12</td>
<td>20%</td>
</tr>
</tbody>
</table>

Note: Compressed age brackets to three groups from the original eight (see Table 3.3 above). Prior to this collapse, there were an insufficient number of participants in each age group: The older (66-75, 75-86, and 86+) groups and younger (18-24) group were each too small on their own for valid analysis. Reducing/combining the number of groups strengthened analysis.

3.3.8.2 Race. Participants were asked to select the race with which they most closely identified. Most of the participants were Caucasian: 51 (85%) were Caucasian and 9 (15%) were not (this after collapsing all minority respondents into “all minority” because the non-Caucasian group size was so small a part of the sample). Of the 9 minority participants, 2 (3.2%) were African American; 1 (1.6%) was Hispanic; 4 (8.1%) were Asian; 1 (1.6%) was Native American; and 1 (1.6%) was Pacific Islander (and added “Hawaiian” in the text field for “Other”).

3.3.8.3 Introversion and Extroversion. On a 4-point, Likert-style scale of introversion and extroversion, participants were asked “How introverted or extroverted are you?” The majority of participants described themselves as introverted: 41 (68%) of participants described themselves as extremely introverted; 19 (32%) as somewhat introverted; 24 (40%) as somewhat extroverted, and 3 (5%) as extremely extroverted.

3.3.8.4 Prior Second Life Use. Three aspects of prior Second Life use were considered in this study: Whether or not a participant had ever logged in to Second Life before; the age of their account; and the total time spent in Second Life.
(68%) participants had logged into Second Life at least once before the Introduction to Building class, while 19 (32%) had not logged in to Second Life before.

Table 3.5. Age of Second Life user account (N = 60) (Experience)

<table>
<thead>
<tr>
<th>Account Age</th>
<th>Number of Participants</th>
<th>Percent of Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;6 months</td>
<td>34</td>
<td>57%</td>
</tr>
<tr>
<td>6-12 months</td>
<td>5</td>
<td>8%</td>
</tr>
<tr>
<td>1 year</td>
<td>2</td>
<td>3%</td>
</tr>
<tr>
<td>2-3 years</td>
<td>13</td>
<td>22%</td>
</tr>
<tr>
<td>4+ years</td>
<td>6</td>
<td>10%</td>
</tr>
</tbody>
</table>

As described in Table 3.5, of the 60 total participants, 34 (57%) had created their Second Life account less than six months ago; 5 (8%) 6-12 months ago; 2 (3%) 1 year ago; 13 (22%) 2-3 years ago; and 6 (10%) four or more years ago.

Because some of these groups had too few respondents to provide useful analysis, and because they were continuous rather than categorical, the groups were collapsed (Table 3.6), first to four groups and then to two: <6 months and >6 months experience.

Table 3.6. Recoded Age of Second Life user account (N = 60) (Experience)

<table>
<thead>
<tr>
<th>Account Age</th>
<th>Number of Participants</th>
<th>Percent of Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;6 months</td>
<td>34</td>
<td>57%</td>
</tr>
<tr>
<td>&gt;6 months</td>
<td>26</td>
<td>43%</td>
</tr>
</tbody>
</table>

Table 3.7. Total Time Spent in Second Life Prior to Study Completion (N = 60) (Time)

<table>
<thead>
<tr>
<th>Time Spent</th>
<th>Number of Participants</th>
<th>Percent of Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;2 hours</td>
<td>28</td>
<td>47%</td>
</tr>
<tr>
<td>2-20 hours</td>
<td>9</td>
<td>15%</td>
</tr>
<tr>
<td>20+ hours</td>
<td>23</td>
<td>38%</td>
</tr>
</tbody>
</table>
In response to the question “How much time (total hours) have you spent in Second Life?” 60 participants, 28 (47%) reported that they had spent less than two hours total in Second Life; 9 (15%) had spent 2-20 hours total in Second Life; and 23 (38%) had spent 20 or more total hours in Second Life.

Table 3.8. Recoded Total Time Spent in Second Life Prior to Study Completion (N = 60) (Time)

<table>
<thead>
<tr>
<th>Time Spent</th>
<th>Number of Participants</th>
<th>Percent of Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;20 hours</td>
<td>28</td>
<td>62%</td>
</tr>
<tr>
<td>20+ hours</td>
<td>32</td>
<td>38%</td>
</tr>
</tbody>
</table>

3.3.8.5 Primary Interest in Second Life. Some readers may find the following of interest though this data was not used any further in this study. On a 4-point, Likert-style scale, 1 being “Not interested at all” and 4 being “Extremely interested,” participants were asked to state their primary interest in Second Life from among the following items: Socializing; building/scripting/content creation; running a business; providing education (teacher); receiving education (student); role-playing/fantasy activities; and Other (please describe). Results are displayed in Table 3.7:

Table 3.9. Frequencies of Primary Interest in Second Life (N = 60)

<table>
<thead>
<tr>
<th>Topic of Interest</th>
<th>Not At All</th>
<th>Somewhat</th>
<th>Fairly</th>
<th>Extremely</th>
</tr>
</thead>
<tbody>
<tr>
<td>Socializing</td>
<td>9 (15%)</td>
<td>20 (33%)</td>
<td>20 (33%)</td>
<td>7 (12%)</td>
</tr>
<tr>
<td>Building/Content Creation</td>
<td>13 (22%)</td>
<td>11 (18%)</td>
<td>15 (25%)</td>
<td>16 (27%)</td>
</tr>
<tr>
<td>Virtual Business</td>
<td>26 (43%)</td>
<td>13 (22%)</td>
<td>11 (18%)</td>
<td>6 (10%)</td>
</tr>
<tr>
<td>Providing Education</td>
<td>22 (37%)</td>
<td>13 (22%)</td>
<td>9 (15%)</td>
<td>10 (17%)</td>
</tr>
<tr>
<td>Receiving Education</td>
<td>11 (18%)</td>
<td>21 (35%)</td>
<td>11 (18%)</td>
<td>13 (22%)</td>
</tr>
<tr>
<td>Role-playing/Fantasy</td>
<td>27 (45%)</td>
<td>18 (30%)</td>
<td>7 (12%)</td>
<td>4 (7%)</td>
</tr>
</tbody>
</table>

*Note: Responses on a 4-point, Likert-style scale.*
In sum, study participants were most interested in building/scripting/content creation and obtaining education, and least interested in role-playing/fantasy activities and running a business. This may be due to selection bias: An Introduction to Building class was part of the study. It is not surprising, then, that people particularly interested in building and obtaining education might sign up to take, and receive $10 compensation for, a class about subject matter in which they might already be interested.

3.3.8.6 Video Game Use. Another item that may be of interest to other researchers is how many Second Life Residents also play video games. While this study by no means claims to answer that question conclusively, this data is available for the small, 60-person sample in this study. Of the 60 total participants, 34 (57%) reported that they also play video games: 0 (0%) reported that they play MMORPGs (World of Warcraft, Everquest); 10 (17%) play non-MMO online games (like Quake, Counterstrike and Team Fortress); 7 (12%) play PC games; 12 (20%) play console games (Wii, Xbox, PS2); and 5 (8%) play handheld (Nintendo DS, Sony PSP) games.

3.3.9 Data Coding. All identifying information (avatar names and email addresses) was removed from SPSS cases and each case was assigned a unique numerical ID from 1-65. Cases with incomplete data were removed, reducing the number of total cases from N = 65 to N = 62. Later, two outliers were removed (both male and both Group V members, which reduced the size of Group V to being the smallest), resulting in a final sample size of N = 60.

During scale reliability analysis (described in more detail in the following section), age brackets were reduced from seven to four because some age groups were too small and did not have enough people in each age group. Experience brackets were also reduced during analysis, from 1-5 and 1-4 to 1 and 2).
Items were scored (i.e. with points or as correct/incorrect) according to the instructions in Appendix I.

### 3.4 Instrumentation and Materials

All instruments (save the screening questionnaire) were part of a single, online post-activity questionnaire. Three dependent variables (presence, retention and cognitive load) were measured with the instrument shown in Appendix F-I. Two of the instruments used in this study are the creation of the author (the cognitive load and retention instruments in Appendix H and I), while one instrument, the Presence Questionnaire, was created and is copyrighted by Witmer and Singer with permission to make digital or hard copies of part or all of their work for personal, classroom, and commercial use, granted without fee provided that appropriate credit is given (Witmer & Singer, 1998).

Witmer and Singer’s Presence Questionnaire is frequently employed in virtual reality research and often cited in the literature on immersion and presence in virtual environments. The Presence Questionnaire (version 3.0) was employed in this study for many sound reasons: Because Moreno (2006) used it in her work on modality and presence, and this study aimed to expand on her work; because it is complete in that it measures multiple aspects of presence; because its validity is well established and no serious shortcomings of the instrument have been detected in more than a decade; because it measures presence data collected from participants immediately following exposure to a virtual environment, which also took place in this study; and because, upon reading it with high familiarity with Second Life, the author found that the items contained in the Presence Questionnaire applied neatly to Second Life without requiring alteration. Though Second Life is not the same (or same kind of) virtual environment in which the PQ was originally applied, the relevance of PQ questions to Second Life inspired the author to apply it to a social virtual world, just as Moreno
(2006) applied the PQ to an environment different than those studied by Witmer and Singer (1998; 2005).

3.4.1 Reliability of Witmer and Singer’s Presence Questionnaire. The first version of the PQ was created in 1998 to measure presence, but Witmer and Singer were also testing whether or not variables affecting presence could be classified into meaningful groups of similar items (Witmer & Singer, 2005), describing aspects of presence, a term that has been seeking and evolving in its definition for many years. If they could do this, Witmer and Singer would be able to produce subscales that could measure and thus evaluate specific aspects of presence, such as immersion and involvement.

Witmer and Singer have described their tests of reliability on the PQ at length in the literature (Witmer & Singer, 2005), briefly reviewed here to state scale reliability for the PQ. Reliability helps to establish a reasonably consistent scale and Witmer and Singer used principal-components analysis with direct oblimin rotation to establish PQ reliability.

Witmer and Singer (1998) conducted a cluster analysis of Version 2.0 of the PQ, which contained 32 items, using PQ data from four experiments (N = 152). Only PQ items that contributed to the reliability of the scale (19 items) were retained for the cluster analysis. The 13 items that were removed from the scale reduced scale reliability as measured by Cronbach’s alpha. This cluster analysis (of 19 items) identified the following PQ subscales: Involved/Control (the ability to control virtual environment events and responsiveness of the virtual environment to the user); Natural (the extent to which interactions felt natural); Interface Quality (whether devices or UI interfere with or distract users from task performance); and Resolution (examining objects closely and from multiple viewpoints) (Witmer & Singer, 2005).
These 19 questions served as the foundation for Version 3.0 of the PQ, which also included new or revised items for a total of 32 items evaluated for their utility as potential additions to the PQ. Three of these 32 items were dropped because they tended to reduce PQ scale reliability when included, leaving a 29-item, reliable PQ scale with Cronbach’s alpha 0.91 (N = 325) (Witmer & Singer, 2005). Subsequent cluster analysis identified six PQ factors (subscales), four of which were found in both this study and the 1998 study (described in the paragraph immediately above). The two new factors are Consistent with Expectations (the extent to which the environment reacts to user-initiated actions as expected) and Haptic/Visual Fidelity (the degree to which the configuration of the virtual environment permits users to examine, search for and manipulate multiple objects using their sense of touch) (Witmer & Singer, 2005).

The 28-item Presence Questionnaire (Appendix F) was measured on a 7-point, Likert-type scale ranging from “No/None” to “High/Extremely” on presence variables. Example items include “How well could you move or manipulate objects in the virtual environment?” and “Were you able to anticipate what would happen next in response to the actions that you performed?” See Appendix F for the complete instrument. Please note that one item from the PQ was omitted (“How well could you actively survey or search the virtual environment using touch?”) because Second Life does not provide, by default, haptic (touch-based) interaction.

### 3.4.2 Reliability of all Presence Items.

The PQ was supplemented with 11 post-experiment questions from Gerhard (2001), all of which are presented in Appendix G. These questions are more attitudinal in nature and measure immersion, communication, involvement and awareness. They were added because they were considered highly salient based on the design of the study presented here, but also because they more directly ask participants about social presence (co-presence) than
does the PQ. The 7-point, Likert-style scale was applied to these questions in order to keep response format consistent. Example items include “How responsive were the avatars of other participants to verbal communication that you initiated?” and “How compelling was your sense of other participants being present?”

Because the addition of these questions could have compromised the established reliability of the PQ, reliability was tested for all presence items as part of this study and was strong with Cronbach’s alpha 0.95.

The total Presence score was a sum of all responses to all presence questions, first the PQ items followed by Gerhard’s (2001) items. Presence subscales were not included or evaluated as part of this study, as Gerhard’s presence subscales may not have mapped directly to Witmer and Singer’s presence factors.

### 3.4.3 Reliability of Cognitive Load Measures

The author created the cognitive load and retention measures (Appendix H–I). The cognitive load measures were brief and typical of cognitive load evaluations, in which study participants are asked to report their perceived levels of mental effort and strain. The Cognitive Load assessment was comprised of two 7-point, Likert-type scale options that ranged from “Not at all” to “Extremely” and took subjective measures of mental strain. These two items were: “How difficult was the Introduction to Building class?” and “How much mental strain did you experience during the Introduction to Building class?”

### 3.4.4 Reliability of Retention Measures

The Retention instrument was originally comprised of nine, and later eight, primarily multiple-choice style questions designed to measure different types of learning outcomes: Declarative knowledge, concept learning, and procedural knowledge (Smith & Ragan, 2005). Example items included “Which of the following shapes are “prims” in Second Life? Please check all that apply.” and “Below is the list of steps required to construct a table in Second
Life. The steps are listed in the incorrect order. Please re-order the steps shown below to successfully build the table shown here (a table with an oval top and four cylindrical legs). Select the #1 button for the first step you should take, with #8 marking the last step in the procedure.” See Appendix I for the complete retention instrument and for scoring instructions. The total Retention score was the sum of all correct answers.

To establish internal scale reliability for each of these three scales, reliability was calculated using Cronbach’s alpha for the dependent variables (retention, presence, and cognitive load). All of the items that constituted a particular measure were entered and a Cronbach’s alpha was run for that measure, resulting in the internal consistency of that particular scale. The procedure and results of scale reliability for all items on the test instrument are presented in Table 3.8:

Table 3.10. Cronbach’s alphas for Presence, Cognitive Load and Retention Scale Reliability

<table>
<thead>
<tr>
<th>Scale</th>
<th>Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retention</td>
<td>0.51</td>
</tr>
<tr>
<td>Cognitive Load</td>
<td>0.94</td>
</tr>
<tr>
<td>Presence</td>
<td>0.95</td>
</tr>
</tbody>
</table>

The original retention alpha was 0.495, too low to be considered reliable. Factor loadings indicated that one question (“Look at the image below. Three lines are shown: a red line, a green line, and a blue line. Each line represents one angle, or axis. Choose the color of the line that represents the Z axis.”) reduced scale reliability when included, so it was not retained, bringing Cronbachs alpha to 0.51, which still did not meet the standard of .07 desired reliability. Upon further examination, it was noted that reliability was underestimated: The retention scale was not a continuous measure, but dichotomous, for which another measure of reliability would be more appropriate. This limitation is also noted in the Discussion section.
CHAPTER 4
RESULTS

This chapter initially reports on all variables (independent, dependent, and demographic), described in terms of their relationship to each other. Next, the independent variable of communication modality (Group V, T, or VT) is described as it relates to the three dependent variables of cognitive load, retention, and presence. Variables with significant intercorrelations are then analyzed and described further.

First, to determine if demographic variables might affect observed outcomes, descriptive statistics and Pearson’s two-tail correlations were run for all variables (independent, dependent and demographic) (Table 4.1). As described in the Methods section, this study used a stratified random sample controlled for age and gender, and a Chi-square test confirmed that the assignment to condition was indeed random and not significant for age (4, N = 60) = 1.34, p = .855), gender (2, N = 60) = 1.48, p = .477) or race (2, N = 60) = .49, p = .782). Other demographic variables (time and experience) were not included in the stratified random control or Chi-square test. To review the relationship between all variables (the independent variable, dependent variables, and demographic variables (time, experience, age, race, or gender)), a correlation matrix is presented in Table 4.1.

The results of the Pearson’s correlations (Table 4.1) were positive and significant for retention and presence (r(60) = .36, p = .01), for experience and modality (r(60) = .28, p = .05), for time and presence (r(60) = .41, p = .01), and for time and retention (r(60) = .48, p = .01).

The results of the Pearson’s correlations were negative and significant for cognitive load and retention (r(60) = -.26, p = .05), for experience and presence (r(60) = -.31, p = .05), for experience and retention (r(60) = -.39, p = .01), for time and
cognitive load ($r(60) = -.36, p = .01$), and for time and experience ($r(60) = .50, p = .01$).
Table 4.1. Correlation matrix for Independent, Dependent and Demographic Variables (N=60)

| Variable       | M   | SD  | 1    | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   |
|----------------|-----|-----|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| V/T/VT         | 2.05| .81 | 1.00 |     |     |     |     |     |     |     |     |     |
| Presence       | 185.47| 27.14| -.12 | 1.00|     |     |     |     |     |     |     |     |
| Retention      | .72 | .17 | -.25 | .36**| 1.00|     |     |     |     |     |     |     |
| Cognitive Load |    |     |      |     |     |     |     |     |     |     |     |     |
| Experience     | 3.80| 1.54| .28* | -.31*| -.39**| .10 | 1.00|     |     |     |     |     |
| Time           | 1.92| .93 | -.15 | .41**| .48**| -.36**| -.50**| 1.00|     |     |     |     |
| Age            | 3.25| 1.54| .09  | -.11 | -.03 | .02  | .01  | .00  | 1.00|     |     |     |
| Gender         | 1.62| .49 | .13  | .12  | -.11 | .04  | -.17 | .15  | -.03 | 1.00|     |     |
| Race           | 1.43| 1.18| .08  | -.02 | .11  | -.08 | -.18 | -.04 | -.23 | .12 | 1.00|     |

*Correlation is significant at the .05 level (2-tailed).

**Correlation is significant at the .01 level (2-tailed).
4.1 Proposed Analyses

The following sections describe the independent and dependent variables that were addressed for each hypothesis and the analysis conducted for each hypothesis. Means and standard deviations were calculated for the independent variable of communication modality (Group V, Group T and Group VT) on each of the three dependent measures (presence, retention, and cognitive load) (Table 4.2).

4.1.1 Test of Hypothesis 1. \( H^1 \): Retention scores will be higher for voice participants (Group V) and voice-and-text participants (Group VT) than for text-only (Group T) participants. The results of a one-way ANOVA to test this hypothesis were significant \( (F(2, 54) = 3.53, p = .04) \). Retention scores were highest for Group T \( (M = 5.38, SD = 1.14) \), lower for Group V \( (M = 5.19, SD = 0.89) \), and lowest for Group VT \( (M = 4.49, SD = 1.32) \).

4.1.2 Test of Hypothesis 2. \( H^2 \): Cognitive load will be lower for voice participants (Group V) and higher for text conditions, voice-and-text participants (Group VT) and text participants (Group T), as evidenced by subjective reports of mental strain and effort. The results of a one-way ANOVA to test this hypothesis were significant \( (F(2, 54) = 4.58, p = .01) \). Cognitive load was lowest for Group T \( (M = 3.33, SD = 2.08) \), higher for Group VT \( (M = 4.14, SD = 2.26) \), and significantly higher for Group V \( (M = 5.56, SD = 2.57) \).

4.1.3 Test of Hypothesis 3. \( H^3 \): Voice will contribute to greater perceptions of presence. The results of a one-way ANOVA to test this hypothesis were not significant \( (F(2, 54) = .65, p = .53) \). Voice use did not contribute to greater perceptions of presence. Group T had the highest presence scores \( (M = 188.70, SD = 24.14) \) while Group V had just slightly lower presence scores \( (M = 188.07, SD = 27.33) \) and Group VT the lowest \( (M = 180.00, SD = 30.11) \).
4.1.4 Test of Hypothesis 4. \( H^4: \) Perceptions of presence will not correlate with deeper learning. The results of a Pearson’s two-tail correlation to test this hypothesis were positive and significant: \( r(60) = .36, p = .01 \). Perceptions of presence did correlate with retention.

Table 4.2. Means and standard deviations for Cognitive load, Retention and Presence by Communication modality

<table>
<thead>
<tr>
<th>Condition</th>
<th>V</th>
<th>T</th>
<th>VT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Cognitive Load</td>
<td>18</td>
<td>5.56</td>
<td>2.57</td>
</tr>
<tr>
<td>Retention</td>
<td>18</td>
<td>5.19</td>
<td>.90</td>
</tr>
<tr>
<td>Presence</td>
<td>18</td>
<td>188.07</td>
<td>27.33</td>
</tr>
</tbody>
</table>
To summarize, a one-way analysis of variance was completed for each dependent variable (cognitive load, retention, and presence) to answer hypotheses 1-3. The results of the ANOVA between groups were significant for cognitive load ($F(2, 54) = 4.58, p = .01$) and retention ($F(2, 54) = 3.53, p = .04$) but were not significant for presence ($F(2, 54) = .65, p = .53$).

The results of each one-way ANOVA are presented graphically in Figures 4.1 - 4.3.
Figure 4.2. Means plot from a 1-way ANOVA of Retention and Communication modality
Figure 4.3. Means plot from a 1-way ANOVA of Presence and Communication modality
Next, to determine if the means were significantly different, post-hoc analysis (Tukey) was run for all dependent variables (cognitive load, retention and presence) by condition (Group V, Group T, and Group VT) (Table 4.3).

Table 4.3. Post-hoc Differences in Means of Cognitive load, Retention and Presence by Condition

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean Difference</th>
<th>Std. Error</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive Load</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V and T</td>
<td>1.11*</td>
<td>.37</td>
<td>.01</td>
</tr>
<tr>
<td>V and VT</td>
<td>.11</td>
<td>.37</td>
<td>.15</td>
</tr>
<tr>
<td>T and V</td>
<td>-1.11*</td>
<td>.37</td>
<td>.01</td>
</tr>
<tr>
<td>T and VT</td>
<td>-.40</td>
<td>.35</td>
<td>.50</td>
</tr>
<tr>
<td>VT and V</td>
<td>-.71</td>
<td>.37</td>
<td>.15</td>
</tr>
<tr>
<td>VT and T</td>
<td>.40</td>
<td>.35</td>
<td>.50</td>
</tr>
<tr>
<td>Retention</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V and T</td>
<td>-.03</td>
<td>.05</td>
<td>.86</td>
</tr>
<tr>
<td>V and VT</td>
<td>.10</td>
<td>.05</td>
<td>.15</td>
</tr>
<tr>
<td>T and V</td>
<td>.03</td>
<td>.05</td>
<td>.86</td>
</tr>
<tr>
<td>T and VT</td>
<td>.13*</td>
<td>.05</td>
<td>.04</td>
</tr>
<tr>
<td>VT and V</td>
<td>-.10</td>
<td>.05</td>
<td>.15</td>
</tr>
<tr>
<td>VT and T</td>
<td>-.13*</td>
<td>.05</td>
<td>.04</td>
</tr>
<tr>
<td>Presence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V and T</td>
<td>.63</td>
<td>8.77</td>
<td>1.00</td>
</tr>
<tr>
<td>V and VT</td>
<td>8.06</td>
<td>8.77</td>
<td>.63</td>
</tr>
<tr>
<td>T and V</td>
<td>.63</td>
<td>8.77</td>
<td>1.00</td>
</tr>
<tr>
<td>T and VT</td>
<td>8.70</td>
<td>8.43</td>
<td>.56</td>
</tr>
<tr>
<td>VT and V</td>
<td>-8.06</td>
<td>8.77</td>
<td>.63</td>
</tr>
<tr>
<td>VT and T</td>
<td>-8.70</td>
<td>8.43</td>
<td>.56</td>
</tr>
</tbody>
</table>

*The mean difference is significant at the .05 level.*
Post-hoc analysis showed mean differences were significant ($p < .05$) between groups VT and V on retention ($p = .03$), between groups T and VT on retention ($p = .04$), and between groups V and T on cognitive load ($p = .01$).

Then, exploratory analysis was conducted for time and experience, as the correlation matrix (Table 4.1) showed a relationship between these two extraneous variables, the independent variable (communication modality), and the three dependent variables (cognitive load, retention, and presence). While not hypothesized, strong correlations (both positive and negative) indicated that time and experience should be treated as second independent variables in future analyses. Six 2-way ANOVAs were run with the independent variables of communication modality and either time or experience, and one of the dependent variables (cognitive load, retention, or presence).

The results for each of the six 2-way ANOVAs are presented in Tables 4.4-4.12 and Figures 4.4-4.9. For each of the six 2-way ANOVAs, a table of descriptive statistics, a line graph, and a between-subjects effect table is presented.
Table 4.4. Descriptive Statistics from a 2-way ANOVA of Cognitive load, Communication modality, and Experience

<table>
<thead>
<tr>
<th>Communication Modality</th>
<th>Experience</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>&lt;6 months</td>
<td>3.29</td>
<td>1.32</td>
<td>7</td>
</tr>
<tr>
<td>V</td>
<td>&gt;6 months</td>
<td>2.45</td>
<td>1.21</td>
<td>11</td>
</tr>
<tr>
<td>V</td>
<td>Total</td>
<td>2.77</td>
<td>1.29</td>
<td>18</td>
</tr>
<tr>
<td>T</td>
<td>&lt;6 months</td>
<td>1.82</td>
<td>1.31</td>
<td>11</td>
</tr>
<tr>
<td>T</td>
<td>&gt;6 months</td>
<td>1.50</td>
<td>.67</td>
<td>10</td>
</tr>
<tr>
<td>T</td>
<td>Total</td>
<td>1.67</td>
<td>1.04</td>
<td>21</td>
</tr>
<tr>
<td>VT</td>
<td>&lt;6 months</td>
<td>2.16</td>
<td>.91</td>
<td>16</td>
</tr>
<tr>
<td>VT</td>
<td>&gt;6 months</td>
<td>1.80</td>
<td>1.79</td>
<td>5</td>
</tr>
<tr>
<td>VT</td>
<td>Total</td>
<td>2.07</td>
<td>1.13</td>
<td>21</td>
</tr>
<tr>
<td>Total</td>
<td>&lt;6 months</td>
<td>2.28</td>
<td>1.23</td>
<td>34</td>
</tr>
<tr>
<td>Total</td>
<td>&gt;6 months</td>
<td>1.96</td>
<td>1.21</td>
<td>26</td>
</tr>
<tr>
<td>Total</td>
<td>Total</td>
<td>2.14</td>
<td>1.22</td>
<td>60</td>
</tr>
</tbody>
</table>
Figure 4.4. Means plot from a 2-way ANOVA of Cognitive load, Communication modality, and Experience
Table 4.5. Test of Between-Subjects Effects from a 2-way ANOVA of Cognitive load, Communication modality, and Experience

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Sq.</th>
<th>df</th>
<th>Mean Sq.</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squ.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>16.09a</td>
<td>5</td>
<td>3.22</td>
<td>2.43</td>
<td>.046</td>
<td>.184</td>
</tr>
<tr>
<td>Intercept</td>
<td>246.49</td>
<td>1</td>
<td>246.49</td>
<td>186.28</td>
<td>.00</td>
<td>.78</td>
</tr>
<tr>
<td>V1T2VT3</td>
<td>14.37</td>
<td>2</td>
<td>7.18</td>
<td>5.43</td>
<td>.01</td>
<td>.17</td>
</tr>
<tr>
<td>Experience</td>
<td>3.30</td>
<td>1</td>
<td>3.30</td>
<td>2.50</td>
<td>.12</td>
<td>.04</td>
</tr>
<tr>
<td>V1T2VT3 * Experience</td>
<td>.72</td>
<td>2</td>
<td>.36</td>
<td>.27</td>
<td>.76</td>
<td>.01</td>
</tr>
<tr>
<td>Error</td>
<td>71.45</td>
<td>54</td>
<td>1.32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>362.75</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>87.55</td>
<td>59</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*R Squared = .184 (Adjusted R Squared = .108)*
To summarize, cognitive load was lowest for Group T ($M = 1.67$) regardless of experience. There was no significant difference between Group T participants who had created their Second Life accounts <6 months ago ($M = 1.82$) and >6 months ago ($M = 1.50$).

Table 4.6. Descriptive Statistics from a 2-way ANOVA of Cognitive load, Communication modality, and Time

<table>
<thead>
<tr>
<th>Communication Modality</th>
<th>Time</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>&lt;20 hours</td>
<td>2.77</td>
<td>1.33</td>
<td>13</td>
</tr>
<tr>
<td>V</td>
<td>&gt;20 hours</td>
<td>2.80</td>
<td>1.30</td>
<td>5</td>
</tr>
<tr>
<td>V</td>
<td>Total</td>
<td>2.77</td>
<td>1.29</td>
<td>18</td>
</tr>
<tr>
<td>T</td>
<td>&lt;20 hours</td>
<td>2.38</td>
<td>1.38</td>
<td>8</td>
</tr>
<tr>
<td>T</td>
<td>&gt;20 hours</td>
<td>1.23</td>
<td>0.38</td>
<td>13</td>
</tr>
<tr>
<td>T</td>
<td>Total</td>
<td>1.67</td>
<td>1.04</td>
<td>21</td>
</tr>
<tr>
<td>VT</td>
<td>&lt;20 hours</td>
<td>2.34</td>
<td>1.15</td>
<td>16</td>
</tr>
<tr>
<td>VT</td>
<td>&gt;20 hours</td>
<td>1.20</td>
<td>0.45</td>
<td>5</td>
</tr>
<tr>
<td>VT</td>
<td>Total</td>
<td>2.07</td>
<td>1.13</td>
<td>21</td>
</tr>
<tr>
<td>Total</td>
<td>&lt;20 hours</td>
<td>2.50</td>
<td>1.25</td>
<td>37</td>
</tr>
<tr>
<td>Total</td>
<td>&gt;20 hours</td>
<td>1.57</td>
<td>0.93</td>
<td>23</td>
</tr>
<tr>
<td>Total</td>
<td>Total</td>
<td>2.14</td>
<td>1.22</td>
<td>60</td>
</tr>
</tbody>
</table>
Figure 4.5. Means plot from a 2-way ANOVA of Cognitive load, Communication modality, and Time
Table 4.7. Test of Between-Subjects Effects from a 2-way ANOVA of Cognitive load, Communication modality, and Time

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Sq.</th>
<th>df</th>
<th>Mean Sq.</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squ.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>23.60(a)</td>
<td>5</td>
<td>4.72</td>
<td>3.99</td>
<td>.004</td>
<td>.27</td>
</tr>
<tr>
<td>Intercept</td>
<td>218.21</td>
<td>1</td>
<td>218.21</td>
<td>184.26</td>
<td>.000</td>
<td>.77</td>
</tr>
<tr>
<td>V1T2VT3</td>
<td>10.14</td>
<td>2</td>
<td>5.07</td>
<td>4.28</td>
<td>.02</td>
<td>.14</td>
</tr>
<tr>
<td>Time</td>
<td>6.87</td>
<td>1</td>
<td>6.87</td>
<td>5.80</td>
<td>.02</td>
<td>.10</td>
</tr>
<tr>
<td>V1T2VT3 * Time</td>
<td>3.53</td>
<td>2</td>
<td>1.77</td>
<td>1.50</td>
<td>.23</td>
<td>.05</td>
</tr>
<tr>
<td>Error</td>
<td>63.95</td>
<td>54</td>
<td>1.18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>362.750</td>
<td>59</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>87.546</td>
<td>59</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(a\) R Squared = .270 (Adjusted R Squared = .202)
To summarize, cognitive load was highest for Group V ($M = 2.77$) and lower for Group T ($M = 1.67$) and VT ($M = 2.07$), without a significant difference between Group T and VT. There was a significant between-groups effect for communication modality and time ($p = .23$). The text condition (Group T) and time spent in Second Life (>20 hours) correlated with a greater decrease in cognitive load. This implies that the text condition may have been more effective for reducing cognitive load among participants who had spent more time in Second Life.

While this is strictly speculative, the correlation between time spent in Second Life, text communication, and reductions in cognitive load may relate to how long ago the participants who had spent more time in Second Life created their accounts: Text communication has been around the longest in Second Life, while voice communication is a newer addition dating back to August 2007.

Table 4.8. Descriptive Statistics from a 2-way ANOVA of Retention, Communication modality, and Experience

<table>
<thead>
<tr>
<th>Communication Modality</th>
<th>Experience</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>&lt;6 months</td>
<td>.68</td>
<td>.14</td>
<td>7</td>
</tr>
<tr>
<td>V</td>
<td>&gt;6 months</td>
<td>.78</td>
<td>.11</td>
<td>11</td>
</tr>
<tr>
<td>V</td>
<td>Total</td>
<td>.74</td>
<td>.13</td>
<td>18</td>
</tr>
<tr>
<td>T</td>
<td>&lt;6 months</td>
<td>.74</td>
<td>.21</td>
<td>11</td>
</tr>
<tr>
<td>T</td>
<td>&gt;6 months</td>
<td>.80</td>
<td>.10</td>
<td>10</td>
</tr>
<tr>
<td>T</td>
<td>Total</td>
<td>.77</td>
<td>.16</td>
<td>21</td>
</tr>
<tr>
<td>VT</td>
<td>&lt;6 months</td>
<td>.57</td>
<td>.15</td>
<td>16</td>
</tr>
<tr>
<td>VT</td>
<td>&gt;6 months</td>
<td>.88</td>
<td>.06</td>
<td>5</td>
</tr>
<tr>
<td>VT</td>
<td>Total</td>
<td>.64</td>
<td>.19</td>
<td>21</td>
</tr>
<tr>
<td>Total</td>
<td>&lt;6 months</td>
<td>.65</td>
<td>.18</td>
<td>34</td>
</tr>
<tr>
<td>Total</td>
<td>&gt;6 months</td>
<td>.81</td>
<td>.10</td>
<td>26</td>
</tr>
<tr>
<td>Total</td>
<td>Total</td>
<td>.72</td>
<td>.17</td>
<td>60</td>
</tr>
</tbody>
</table>
Figure 4.6. Means plot from a 2-way ANOVA of Retention, Communication modality, and Experience
Table 4.9. Test of Between-Subjects Effects from a 2-way ANOVA of Retention, Communication modality, and Experience

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Sq.</th>
<th>df</th>
<th>Mean Sq.</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Sq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>.616^a</td>
<td>5</td>
<td>.12</td>
<td>6.09</td>
<td>.000</td>
<td>.36</td>
</tr>
<tr>
<td>Intercept</td>
<td>28.83</td>
<td>1</td>
<td>28.83</td>
<td>1425.25</td>
<td>.000</td>
<td>.96</td>
</tr>
<tr>
<td>V1T2VT3</td>
<td>.02</td>
<td>2</td>
<td>.01</td>
<td>.59</td>
<td>.56</td>
<td>.02</td>
</tr>
<tr>
<td>Experience</td>
<td>.31</td>
<td>1</td>
<td>.31</td>
<td>15.47</td>
<td>.000</td>
<td>.22</td>
</tr>
<tr>
<td>V1T2VT3 * Experience</td>
<td>.16</td>
<td>2</td>
<td>.08</td>
<td>3.90</td>
<td>.03</td>
<td>.13</td>
</tr>
<tr>
<td>Error</td>
<td>1.09</td>
<td>54</td>
<td>.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>32.46</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>1.71</td>
<td>59</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^aR Squared = .361 (Adjusted R Squared = .301)
While the text condition (Group T) increased retention slightly but not significantly for both experience levels, retention scores were highest for Group VT participants who were also more experienced (>6 months) ($M = .88$). Group VT had the opposite effect on less experienced participants, for whom retention dropped significantly ($M = .57$). For retention, the between-subjects effects for communication modality and experience were significant ($p = .03$).

Table 4.10. Descriptive Statistics from a 2-way ANOVA of Retention, Communication modality, and Time

<table>
<thead>
<tr>
<th>Communication Modality</th>
<th>Time</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>&lt;20 hours</td>
<td>.74</td>
<td>.15</td>
<td>13</td>
</tr>
<tr>
<td>V</td>
<td>&gt;20 hours</td>
<td>.74</td>
<td>.06</td>
<td>5</td>
</tr>
<tr>
<td>V</td>
<td>Total</td>
<td>.74</td>
<td>.13</td>
<td>18</td>
</tr>
<tr>
<td>T</td>
<td>&lt;20 hours</td>
<td>.70</td>
<td>.23</td>
<td>8</td>
</tr>
<tr>
<td>T</td>
<td>&gt;20 hours</td>
<td>.81</td>
<td>.09</td>
<td>13</td>
</tr>
<tr>
<td>T</td>
<td>Total</td>
<td>.77</td>
<td>.16</td>
<td>21</td>
</tr>
<tr>
<td>VT</td>
<td>&lt;20 hours</td>
<td>.59</td>
<td>.18</td>
<td>16</td>
</tr>
<tr>
<td>VT</td>
<td>&gt;20 hours</td>
<td>.79</td>
<td>.13</td>
<td>5</td>
</tr>
<tr>
<td>VT</td>
<td>Total</td>
<td>.64</td>
<td>.19</td>
<td>21</td>
</tr>
<tr>
<td>Total</td>
<td>&lt;20 hours</td>
<td>.67</td>
<td>.19</td>
<td>37</td>
</tr>
<tr>
<td>Total</td>
<td>&gt;20 hours</td>
<td>.79</td>
<td>.10</td>
<td>23</td>
</tr>
<tr>
<td>Total</td>
<td>Total</td>
<td>.72</td>
<td>.17</td>
<td>60</td>
</tr>
</tbody>
</table>
Figure 4.7. Means plot from a 2-way ANOVA of Retention, Communication modality, and Time.
Table 4.11. Test of Between-Subjects Effects from a 2-way ANOVA of Cognitive load, Communication modality, and Time

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Sq.</th>
<th>df</th>
<th>Mean Sq.</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Sq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>.406&lt;sup&gt;a&lt;/sup&gt;</td>
<td>5</td>
<td>.08</td>
<td>3.37</td>
<td>.01</td>
<td>.238</td>
</tr>
<tr>
<td>Intercept</td>
<td>25.85</td>
<td>1</td>
<td>25.85</td>
<td>1071.74</td>
<td>.000</td>
<td>.952</td>
</tr>
<tr>
<td>V1T2VT3</td>
<td>.03</td>
<td>2</td>
<td>.02</td>
<td>.72</td>
<td>.50</td>
<td>.026</td>
</tr>
<tr>
<td>Time</td>
<td>.13</td>
<td>1</td>
<td>.13</td>
<td>5.39</td>
<td>.02</td>
<td>.091</td>
</tr>
<tr>
<td>V1T2VT3 * Time</td>
<td>.08</td>
<td>2</td>
<td>.04</td>
<td>1.61</td>
<td>.21</td>
<td>.056</td>
</tr>
<tr>
<td>Error</td>
<td>1.30</td>
<td>54</td>
<td>.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>32.46</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>1.71</td>
<td>59</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>*R Squared = .238 (Adjusted R Squared = .167)*
To summarize, retention scores were approximately the same for participants in the voice condition (Group V), regardless of time spent in Second Life. The text condition (Group T) correlated with increased retention, but only for those participants who had spent more time in Second Life (\( M = .81 \)). As a result, a causal relationship cannot be determined between communication modality and retention, or experience and retention: Both correlated with improvements in learning. For retention, between-groups effects were significant for communication modality and time spent in Second Life (\( p = .21 \)).

Table 4.12. Descriptive Statistics from a 2-way ANOVA of Presence, Communication modality, and Experience

<table>
<thead>
<tr>
<th>Communication Modality</th>
<th>Experience</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>&lt;6 months</td>
<td>174.56</td>
<td>26.50</td>
<td>7</td>
</tr>
<tr>
<td>V</td>
<td>&gt;6 months</td>
<td>196.66</td>
<td>25.28</td>
<td>11</td>
</tr>
<tr>
<td>V</td>
<td>Total</td>
<td>188.07</td>
<td>27.33</td>
<td>18</td>
</tr>
<tr>
<td>T</td>
<td>&lt;6 months</td>
<td>180.03</td>
<td>26.37</td>
<td>11</td>
</tr>
<tr>
<td>T</td>
<td>&gt;6 months</td>
<td>198.23</td>
<td>18.15</td>
<td>10</td>
</tr>
<tr>
<td>T</td>
<td>Total</td>
<td>188.69</td>
<td>24.14</td>
<td>21</td>
</tr>
<tr>
<td>VT</td>
<td>&lt;6 months</td>
<td>176.32</td>
<td>32.44</td>
<td>16</td>
</tr>
<tr>
<td>VT</td>
<td>&gt;6 months</td>
<td>191.80</td>
<td>18.94</td>
<td>5</td>
</tr>
<tr>
<td>VT</td>
<td>Total</td>
<td>180.00</td>
<td>30.11</td>
<td>21</td>
</tr>
<tr>
<td>Total</td>
<td>&lt;6 months</td>
<td>177.16</td>
<td>28.66</td>
<td>34</td>
</tr>
<tr>
<td>Total</td>
<td>&gt;6 months</td>
<td>196.33</td>
<td>20.91</td>
<td>26</td>
</tr>
<tr>
<td>Total</td>
<td>Total</td>
<td>185.47</td>
<td>27.14</td>
<td>60</td>
</tr>
</tbody>
</table>
Figure 4.8. Means plot from a 2-way ANOVA of Presence, Communication modality, and Experience.
Table 4.13. Test of Between-Subjects Effects from a 2-way ANOVA of Presence, Communication modality, and Experience

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Sq.</th>
<th>df</th>
<th>Mean Sq.</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Sq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>5704.20a</td>
<td>5</td>
<td>1140.84</td>
<td>1.63</td>
<td>.167</td>
<td>.13</td>
</tr>
<tr>
<td>Intercept</td>
<td>1817630.34</td>
<td>1</td>
<td>1817630.34</td>
<td>2600.25</td>
<td>.000</td>
<td>.98</td>
</tr>
<tr>
<td>V1T2VT3</td>
<td>248.82</td>
<td>2</td>
<td>124.41</td>
<td>.18</td>
<td>.837</td>
<td>.01</td>
</tr>
<tr>
<td>Experience</td>
<td>4527.71</td>
<td>1</td>
<td>4527.71</td>
<td>6.48</td>
<td>.014</td>
<td>.11</td>
</tr>
<tr>
<td>V1T2VT3 * Experience</td>
<td>90.19</td>
<td>2</td>
<td>45.10</td>
<td>.07</td>
<td>.938</td>
<td>.00</td>
</tr>
<tr>
<td>Error</td>
<td>37747.21</td>
<td>54</td>
<td>699.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2107293.56</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>43451.40</td>
<td>59</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) R Squared = .131 (Adjusted R Squared = .051)
To summarize, the text condition (Group T) and correlated with increases in perceptions of presence for both experience levels in Second Life. For presence, between-groups effects were significant for experience ($p = .014$) but were not significant for communication modality and experience in Second Life ($p = .938$).

Table 4.14. Descriptive Statistics from a 2-way ANOVA of Presence, Communication modality, and Time

<table>
<thead>
<tr>
<th>Communication Modality</th>
<th>Time</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>&lt;20 hours</td>
<td>182.79</td>
<td>29.73</td>
<td>13</td>
</tr>
<tr>
<td>V</td>
<td>&gt;20 hours</td>
<td>201.80</td>
<td>13.99</td>
<td>5</td>
</tr>
<tr>
<td>V</td>
<td>Total</td>
<td>188.07</td>
<td>27.33</td>
<td>18</td>
</tr>
<tr>
<td>T</td>
<td>&lt;20 hours</td>
<td>174.82</td>
<td>26.93</td>
<td>8</td>
</tr>
<tr>
<td>T</td>
<td>&gt;20 hours</td>
<td>197.23</td>
<td>18.47</td>
<td>13</td>
</tr>
<tr>
<td>T</td>
<td>Total</td>
<td>188.69</td>
<td>24.14</td>
<td>21</td>
</tr>
<tr>
<td>VT</td>
<td>&lt;20 hours</td>
<td>175.51</td>
<td>31.69</td>
<td>16</td>
</tr>
<tr>
<td>VT</td>
<td>&gt;20 hours</td>
<td>194.40</td>
<td>20.67</td>
<td>5</td>
</tr>
<tr>
<td>VT</td>
<td>Total</td>
<td>180.00</td>
<td>30.11</td>
<td>21</td>
</tr>
<tr>
<td>Total</td>
<td>&lt;20 hours</td>
<td>177.92</td>
<td>29.45</td>
<td>37</td>
</tr>
<tr>
<td>Total</td>
<td>&gt;20 hours</td>
<td>197.61</td>
<td>17.49</td>
<td>23</td>
</tr>
<tr>
<td>Total</td>
<td>Total</td>
<td>185.47</td>
<td>27.14</td>
<td>60</td>
</tr>
</tbody>
</table>
Figure 4.9. Means plot from a 2-way ANOVA of Presence, Communication modality, and Time
Table 4.15. Test of Between-Subjects Effects from a 2-way ANOVA of Presence, Communication modality, and Time

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Sq.</th>
<th>df</th>
<th>Mean Sq.</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Sq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>6121.08&lt;sup&gt;a&lt;/sup&gt;</td>
<td>5</td>
<td>1224.22</td>
<td>1.77</td>
<td>.134</td>
<td>.14</td>
</tr>
<tr>
<td>Intercept</td>
<td>1711891.39</td>
<td>1</td>
<td>1711891.39</td>
<td>2476.33</td>
<td>.000</td>
<td>.98</td>
</tr>
<tr>
<td>V1T2VT3</td>
<td>473.74</td>
<td>2</td>
<td>236.87</td>
<td>.34</td>
<td>.711</td>
<td>.01</td>
</tr>
<tr>
<td>Time</td>
<td>4908.54</td>
<td>1</td>
<td>4908.54</td>
<td>7.10</td>
<td>.010</td>
<td>.12</td>
</tr>
<tr>
<td>V1T2VT3 * Time</td>
<td>35.68</td>
<td>2</td>
<td>17.84</td>
<td>.03</td>
<td>.975</td>
<td>.00</td>
</tr>
<tr>
<td>Error</td>
<td>37330.325</td>
<td>54</td>
<td>691.30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2107293.56</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>43451.40</td>
<td>59</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>\(R^2 = .141\) (Adjusted \(R^2 = .061\))
The voice condition (Group V) correlated with increased perceptions of presence for participants who had spent both less and more time in Second Life. For presence, between-subjects effects were significant for time ($p = .010$) but not for communication modality and time ($M = .975$).

To summarize the findings of the six 2-way ANOVAs, significant between-subjects effects were found for cognitive load and time ($p = .23$), for retention and time ($p = .21$), and for retention and experience ($p = .03$).

Table 4.16. Summary Table of Between Subjects Effects from six 2-way ANOVAs

<table>
<thead>
<tr>
<th>Variables</th>
<th>Sum of Sq.</th>
<th>df</th>
<th>Mean Sq.</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Sq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive Load &amp; Time</td>
<td>3.53</td>
<td>2</td>
<td>1.77</td>
<td>1.50</td>
<td>.23</td>
<td>.05</td>
</tr>
<tr>
<td>Cognitive Load &amp; Experience</td>
<td>.72</td>
<td>2</td>
<td>.36</td>
<td>.27</td>
<td>.76</td>
<td>.01</td>
</tr>
<tr>
<td>Retention &amp; Time</td>
<td>.08</td>
<td>2</td>
<td>.04</td>
<td>1.60</td>
<td>.21</td>
<td>.06</td>
</tr>
<tr>
<td>Retention &amp; Experience</td>
<td>.16</td>
<td>2</td>
<td>.08</td>
<td>3.90</td>
<td>.03</td>
<td>.13</td>
</tr>
<tr>
<td>Presence &amp; Time</td>
<td>35.68</td>
<td>2</td>
<td>17.84</td>
<td>.03</td>
<td>.98</td>
<td>.00</td>
</tr>
<tr>
<td>Presence &amp; Experience</td>
<td>90.19</td>
<td>2</td>
<td>45.10</td>
<td>.07</td>
<td>.94</td>
<td>.00</td>
</tr>
</tbody>
</table>
CHAPTER 5
DISCUSSION

This chapter begins with a review of the goals and contributions of the study presented here and moves on to a review and discussion of the results by hypothesis. Strengths, weaknesses, and limitations of the study are also presented, as are directions for future research are also presented.

5.1 Contributions The study presented here is an empirical, quantitative, and statistically significant one of learning and communication modality in a publicly available, popular social virtual world. As of this writing, research on this topic is still somewhat rare. This study also extended the work of Mayer and Moreno (1998) on the application of the modality principle to multimedia learning, specifically by:

- Finding that experience with and time spent in the virtual environment of Second Life have a significant impact on cognitive load, retention and presence, and should be controlled for in group assignment in future studies that aim to focus on modality.

- Using a longer learning episode (an average of 45 minutes), per Mayer and Moreno’s (1998) note that their learning episodes were brief.

- Allowing hands-on interaction by study participants (or, at least, the hands of their avatars).

- Expanding narration in the form of a live human voice, not just a recorded narration.

- Studying communication modality in a social situation, in which there is the opportunity for interaction between other students rather than one student viewing one animation alone.
• Studying adults, not only college students, from ages 18-80+

• Not only studying how-it-works content but also how-to content, with hands on learning.

• Studying simultaneous communication modes. The proposed study was comprised of three groups: Group V (voice only), Group T (text only), and Group VT (voice and text simultaneously).

• Supplementing existing research on cognitive load theory, the split-attention effect, and dual-processing theory by: (1) examining the split-attention effect and dual-processing theory in a social virtual world rather than a website-based or paper-based instructional environment, and (2) by using a single independent variable (communication modality) and multiple dependent measures (cognitive load, retention and presence initially, and later experience and time).

• Providing additional refutation of the information-equivalency hypothesis, which predicts no differences between the narration and text groups on any of the tests when identical information is presented to both groups.

• Calling attention to the need to broaden the goals of technical communicators and instructional designers to include social virtual worlds in their work, particularly since platforms like Second Life have a substantial educational community.

5.2 Review of Results

This study sought to determine whether or not communication modality (assignment to Group V, T, or VT) had an impact on cognitive load, retention and presence. The most significant finding is that experience and time had significant effects on measures of cognitive load, retention, and presence.
Retention scores were approximately the same for participants in the voice condition (Group V), regardless of the time (either <20 hours or >20 hours) participants had spent in Second Life. The text condition (Group T) correlated with increased retention, but only for those participants who had spent more time in Second Life (i.e. >20 hours, $M = .81$). Based on these results, it is impossible to determine whether time spent in Second Life, the communication modality, or both led to improvements in learning. For retention, between-groups effects were significant for communication modality and time spent in Second Life ($p = .21$).

While the text condition (Group T) increased retention slightly but not significantly for both experience levels, retention scores were highest for Group VT participants who were also more experienced (>6 months) ($M = .88$). Group VT had the opposite effect on less experienced participants, for whom retention dropped significantly ($M = .57$). Based on this, Group VT is not a recommended, initial communication modality for learners who are less experienced with Second Life. For retention, the between-subjects effects for communication modality and experience were significant ($p = .03$).

The voice condition (Group V) correlated with increased perceptions of presence for participants who had spent both less and more time in Second Life. For presence, between-subjects effects were significant for time ($p = .010$) but not for communication modality and time ($M = .975$). The text condition (Group T) and correlated with increases in perceptions of presence for both experience levels in Second Life. For presence, between-groups effects were significant for experience ($p = .014$) but were not significant for communication modality and experience in Second Life ($p = .938$).

Significant between-subjects effects were found for cognitive load and time ($p = .23$), for retention and time ($p = .21$), and for retention and experience ($p = .03$).
These findings are not surprising: According to Eyring et al. (1993), task familiarity (an individual’s possession of declarative knowledge and procedures relevant to performance of a given task) can be gained through prior experience with the task, or through experience with tasks similar in terms of declarative knowledge and procedures. Eyring et al. (1993) cite Ackerman (1989), who suggested that previous experience may allow an individual to pass more rapidly through early stages of skill acquisition and reduce cognitive demands.

It is possible and, based on the findings about the impact of time and experience presented here, that more experienced Second Life users had learned how to better use the Second Life interface (which was necessarily a strong component of the learning activity) and may have also known other content included in the learning activity script.

5.2.1 Discussion of Hypothesis 1. \( H^1: \) Retention scores will be higher for voice participants (Group V) and voice-and-text participants (Group VT) than for text-only (Group T) participants. The results of a one-way ANOVA to test this hypothesis were significant (\( F(2, 54) = 3.53, p = .04 \)). Retention scores were highest for Group T (\( M = 5.38, SD = 1.14 \)), lower for Group V (\( M = 5.19, SD = 0.89 \)), and lowest for Group VT (\( M = 4.49, SD = 1.32 \)).

According to Mayer and Morenos (1998) cognitive affective theory of learning with media (CATLM), cognitive load theory, and dual-processing theory, using a visual modality to present both pictorial (in this case, the 3D objects in the Introduction to Building class) and verbal information (text-based instruction) can create an overload situation for the learner. More memory capacity is available when dual modalities are used (Penney, 1989). According to the dual-processing theory of working memory, learners must work harder to learn in a no-voice (text-based) instruction than a voice-based instruction.
Further, Moreno and Mayer (2002), in their study of virtual environments with a concurrent narration and text condition (Group NT), found that instruction with a concurrent auditory condition resulted in deeper learning. Based on the similarity of the Group VT condition in this study to Moreno and Mayers (2002) Group NT condition, a similar result was expected.

The results presented here, however, indicate that retention scores were significantly higher for Group T compared to Group V, and that there was no significant difference in performance between Group T and Group VT. These findings are the opposite of what was hypothesized based on the theoretical and experimental research covered in the review of the literature, and present a significant departure from a substantial body of sound, empirical research that has demonstrated the beneficial effects of presenting instructional material in a mixed auditory and visual mode. In this case, the mixed auditory and visual mode (Group VT) did not have significantly higher retention scores than Group T.

The results of the tests for hypothesis 1 do, however, support those of Mousavi et al.’s (1995) sixth experiment in which two groups, a visual-visual mode and an auditory-auditory mode, were studied to see if visual processes were inherently more demanding. Instead, Mousavi et al. (1995) found the opposite: A significant difference favored the visual-visual group (which was most like Group T in the study presented here). Mousavi et al.’s (1995) results suggest that the auditory-auditory mode may have been more intrinsically demanding of cognitive load.

The study data presented here contain no information on the intrinsic demand of different communication modalities, but it is possible that variation in intrinsic demand may have contributed to the results found in this study. Future research should delve into this possibility more deeply.
To summarize the findings for hypothesis 1, in a departure from Moreno and Mayer’s (2002) findings in Experiment 3A, receiving information as on-screen text rather than narration did not significantly hinder learning among the participants in this study: Group T, which received information as on-screen text only, had the highest retention scores of the three groups.

5.2.2 Discussion of Hypothesis 2. \( H^2 \): Cognitive load will be lower for voice participants (Group V) and higher for text conditions, voice-and-text participants (Group VT) and text participants (Group T), as evidenced by subjective reports of mental strain and effort. The results of a one-way ANOVA to test this hypothesis were significant \( (F(2, 54) = 4.58, p = .01) \). Cognitive load was lowest for Group T \((M = 3.33, SD = 2.08)\), higher for Group VT \((M = 4.14, SD = 2.26)\), and significantly higher for Group V \((M = 5.56, SD = 2.57)\).

Why might cognitive load have been higher for voice conditions, a departure from consistent findings in the literature that usually show the opposite? It would be careless not to explore the most simple and obvious explanation first: It is possible that cognitive load was not actually higher but appeared that way in measurements. The text modality automatically offers, via the Second Life user interface, something critical to retention that voice does not: Transcripts.

Due to the nature of remote, online research, it was impossible for the instructor to control whether or not participants in Groups T and VT retained full or partial text transcripts of the material in the Introduction to Building Class. If participants deliberately did so (and even if they did not do so deliberately but simply happened to leave the Second Life viewer window open while they brought up the post-activity questionnaire in a browser window, and subsequently discovered they could go back to the Second Life viewer window and refresh their memories), the Group T and Group VT participants would have had an advantage. As a result, Groups T and VT could
have scored more highly on the retention items, thus providing the result of higher cognitive load for Group V. While the instructor could have asked all participants to leave Second Life before completing the post-activity questionnaire, text transcripts still could have been saved after this request was made.

If Group V’s higher cognitive load is not due to a simple lack of text transcripts, another possible explanation is that auditory load may be intrinsically more demanding (as discussed in the review of the literature), or that voice (auditory communication) in combination with social virtual worlds creates higher auditory load than might be expected for other multimedia learning contexts. When looking at the user interface (UI) of a social virtual world, it is easy to see how the visual channel might be overloaded: The world itself is comprised of 3D content and, in Second Life as of this writing, there are dozens of UI elements available (and possibly in use) at any given moment: There is a bottom menu, a communication menu, a top level menu bar (similar in appearance to that of a web browser), various buttons and preferences, and the ability to right-click on one’s own and other avatars and surface still more pie menus. Given this quantity of visual elements, its only more difficult to understand why the auditory channel, rather than the visual, would be the channel to increase cognitive load.

The nature of the activity may have played a role in contributing to higher auditory load. The instructor’s script for the Introduction to Building class had complimentary information to the actions taking place in Second Life. It is possible that the hands-on nature of the activity—doing something hands-on while trying to attend to the verbal information in the auditory channel—created a higher cognitive load situation than might a more passive learning activity, such as observing an animation or listening to a speaker. The hands-on activity may have had the effect of creating an environment more like that of multitasking than one of focused learning, partic-
ularly since the user activity (the way in which a particular participant completed a learning task) could have deviated from voice instructions, either intentionally out of curiosity about something else or accidentally by clicking on a UI element unrelated to the activity at hand.

In this way, a participant’s visual stream could have become more dissonant from the auditory stream. According to Penney’s separate streams hypothesis, however, different content coming in through two streams really shouldn’t matter. In fact, this situation should have had the effect of freeing up working memory. In Penney’s (1989) work, lists of different items presented simultaneously to different sensory modalities improved short-term memory retention.

Further research should investigate differences in cognitive load after instruction of the same content (script) is delivered differently: More passively, in a typical speaking narrator, listening student configuration, and more actively, with a more hands-on format.

5.2.3 Discussion of Hypothesis 3. \( H^3 \): Voice will contribute to greater perceptions of presence. The results of a one-way ANOVA to test this hypothesis were not significant \( (F(2, 54) = .65, p = .53) \). Voice use did not contribute to greater perceptions of presence. Group T had the highest presence scores \( (M = 188.70, SD = 24.14) \) while Group V had just slightly lower presence scores \( (M = 188.07, SD = 27.33) \) and Group VT the lowest \( (M = 180.00, SD = 30.11) \).

This hypothesis was based on prior findings (Salinas, 2004) that voice contributed to co-presence (the sense of being with others in a virtual world), and that co-presence was a factor of presence. One possible explanation (though not supported by any of the data collected as part of this study) is that the combination of real-world voices with in-world avatars may actually lessen the "fantasy impression of avatars,
reminding participants more of the real world than the virtual one.

Conflicts in prior research still exist and are worthy of further study: Some research (Salinas, 2004) indicated that modality (specifically voice) could impact perceptions of presence. Other research (Moreno & Mayer, 2002) found that modality did not impact perceptions of presence, while the level of immersion made possible with the media environment (a desktop vs. head-mounted display) did affect presence. The results of this study support the findings of Moreno and Mayer (2002) and do not extend those of Salinas (2004).

Future research should evaluate perceptions of presence with real voices and voice fonts, which are filters or morphs that make human voices sound different (male-to-female, female-to-male, human-to-robot, and so on). As of this writing, voice fonts are a planned, forthcoming feature for Second Life. Research that examines perceptions of presence for groups that hear unaltered human voices and groups that hear voices to which voice fonts have been applied may provide more sound data on whether or not hearing a voice that sounds less real increases, decreases, or does not alter users perception of presence in a social virtual world. Comparisons between two voice groups (Real Voice (RV) and Voice Font (VF)), studied along with with Text Only, Real Voice and Text (concurrently) and Voice Font and Text (concurrently) groups might also provide more substantive data on communication modality and perceptions of presence in virtual worlds.

It is also possible that the overall level of communication was low. Participants were not required to engage in any conversation, whether text or voice based, while attending the Introduction to Building class as part of this study. There is no data on this factor because it is difficult to track (the instructor could not see each participants computer screen, for example, and had no way of knowing if someone was or was not engaged in one-to-one or other group communication as part of or outside of class).
The only chat visible to the instructor is in the Local Chat window, visible by all avatars in a given range. It is possible that very little communication took place, and thus that modality less noticed and thus less salient to the study experience than it may have been in other contexts (like a virtual support group for recently diagnosed cancer patients, for example, in which speaking with others may be a more a primary interest).

Related to this, little or no conversation may have contributed to a decrease in participants’ sense of social presence or co-presence: A sense of being with others includes things like primitive responses to social cues, like laughter. While this sort of data was not recorded as part of this study, anecdotal recollection brings to mind only two livelier classes that evidenced this sort of primitive response. Instructor impressions, however, are not nearly as important as whether participants perceived themselves, in their own experience, as having had conversations that included more spontaneous, primitive responses. Unfortunately, such data was not collected as part of this study but should be considered in future research as a possible explanation for presence measures.

5.2.4 Discussion of Hypothesis 4.  *H4: Perceptions of presence will not correlate with deeper learning.* The results of a Pearson’s two-tail correlation to test this hypothesis were positive and significant: \((r(60) = .36, p = .01)\). Perceptions of presence did correlate with retention.

5.3 Limitations

It is important to note that none of the studies discussed throughout this dissertation are directly comparable (apples to apples). Each study is different from the other in one or more major ways, whether in the type of multimedia environment used; the number and type of participants; the presence, cognitive load, and learning
measures used; the content and nature of the instructional activity; the natural or synthetic nature of the instructor (and by extension the voices of animated agents or human instructors); and more. While the study presented here sought to build and expand on prior research on modality, cognitive load, presence and learning, this study did not directly replicate another. This is not to say that the results discussed here are not valid so much as a reminder of the very limited context in which they were found.

It is also important to remember that the study results presented here are based on a small sample (N = 60). According to Cohen, this size was the minimum needed to obtain power for the planned statistical analyses. Replicating this study with a larger sample may strengthen and/or change the results. Doing so may, however, prove time consuming if the recruitment and no-show rates found in this study are found by others: An elapsed period of six months was required to obtain N = 60. Still, a replicated or similar study that uses a larger sample size with striated random assignment based on experience and time—instead of race, gender and age—is a worthwhile direction for future work.

There is a rich history of work, much of it in linguistics, on voice and gender. Gender, of the avatar and voice of the instructor, were kept constant in this study and the study of voice and gender was beyond the scope of this study. Future research should see if the voice of the instructor impacts cognitive load, retention, and/or presence: Do participants in virtual worlds retain information better when the instructor is male or female? Do participants’ perceptions of presence increase or decrease when they hear voices of the same gender that they are?

The retention measures used in this study could have been, and should become, stronger: The retention instrument was brief in terms of the number of questions it contained (though not in the time required to complete it) and focused on Second
Life-specific content, which it is difficult to judge the robustness of. This study should be replicated with non-Second Life specific content (ideally multiple types of subject matter) to see if the results for retention, presence and cognitive load hold across different types of subject matter. Varying subject matter, of course, introduces the possibility for high experience or existing knowledge on the part of participants, so those factors should be examined and, if necessary, controlled for.

While not a result of the study, its worth pointing out how difficult it was for people to use Second Life to participate in a study. Despite detailed web pages and email messages (which, like so many, most likely were not read) many people thought they could access Second Life through a web page. We think of online studies as easier to conduct and this might be true if web surveys are involved. When an online study requires more effort than completing an online survey, researchers should be prepared for potentially higher rates of no-shows and to offer higher compensation than they may have planned on giving in order to reach their N.

In addition, this study confined itself to surveying, interviewing, and observing Residents within Second Life. A study of additional social virtual worlds like There, Kaneva or Active Worlds, and the comparison of communication habits between users of each, is an interesting direction for further research and one worth pursuing at a later time. Studying different virtual environments, however, would have increased the number of variables involved and was beyond the scope of this dissertation. This limitation to Second Life, however, may decrease the generalizability of findings to MMORPGs and to other social virtual worlds, such as There.

This study dealt only with short-term, not long-term, memory, in that the post-activity assessment always took place immediately after the Introduction to Building class. No participants, for example, attended the class and then returned to the post-activity assessment hours, days, or weeks later. Finally, the prior spatial
abilities of participants were not evaluated or taken into account.
CHAPTER 6
GLOSSARY OF KEY TERMS

1. Co-presence: One's sense of "being in a social virtual world with others."

2. Integrated voice communication: A voice application that is part of and fully integrated with a virtual world, as opposed to voice communication via a third party, external voice application such as Skype or Ventrilo (e.g., an external VoIP application that is completely separate from the virtual world).

3. Massive Multiplayer Online Rollplaying Game (MMORPG): Throughout this dissertation, online games will be distinguished from other types of virtual worlds by being referred to as massive multiplayer online role-playing games (MMORPGs). Similar terms include Massive Multiplayer Online Games (MMOGs).

4. Multimodal interaction: Multimodal interaction provides users with multiple modes of interfacing with a given computer system. The most common multimodal interface combines a visual modality (e.g., a display and input device, like a mouse) with a voice modality (e.g., speech recognition for voice input). For the purposes of this dissertation, multimodal interaction will refer to an interface that combines a visual modality with a voice modality, but this voice modality is not limited or related to speech recognition. The voice modality implied here is person-to-person (or rather, avatar-to-avatar) real-time voice communication via integrated (in world) voiceover IP (VoIP).

5. Presence: One's sense of being in a virtual world (versus in the real world or another location).

6. Second Life: An online, persistent, 3D, social virtual world that is not a game. Human users take the form of avatars (3D representations of themselves) and
communicate with one another in live voice chat, typed text chat, or both. Second Life is available (for free) at www.secondlife.com.

7. Social virtual world  This term can mean many things, but for the purposes of this dissertation it will be used as shorthand for a 3-D, online, collaborative, multiplayer virtual world that is not a game.

8. Virtual Environment (VE) or Virtual Reality (VR)  According to Wikipedia and common usage in the research literature, virtual reality (VR) is a technology that allows a user to interact with a computer-simulated environment. Most VR environments are primarily visual experiences, displayed either on a computer screen or through special or stereoscopic displays. Some simulations include additional sensory information, such as sound through speakers or headphones, and haptic systems may include tactile information. Users can interact with a virtual environment or a virtual artifact (VA) either through the use of standard input devices such as a keyboard and mouse, or through multimodal devices such as a wired glove, the Polhemus boom arm, and the omnidirectional treadmill.

Second Life could, according to this broad definition, be classified as VR in that it allows users to interact with a computer-simulated environment, and users can interact with it through the use of standard input devices, most commonly keyboard and mouse and not usually multimodal devices like wired gloves. More common usages of VR, however, tend to include most of all of the attributes described above, and most of these are not attributes of Second Life. For this reason, then, Second Life will be described as a social virtual world rather than a virtual reality application.
APPENDIX A

STUDY INFORMATION WEB PAGE
Research Project on Communication and Learning in Second Life

Principle Investigator: Stephany Filimon Contact me via email at: filiste@iit.edu or Stephany@lindenlab.com

**Goal** This research project is an effort to better understand communication in Second Life and how it impacts learning.

For a few years now, the media has written about the potential of social virtual worlds like Second Life for things like remote meetings and conferences, collaboration, and education. Many of these claims are unfounded and have little or no data to support them. This project is one effort to investigate the learning potential of, and better understand communication in, Second Life in a sound empirical way.

Your participation in this research project will provide an improved and reliable understanding of learning and communication in Second Life, and will help universities, educators, and Linden Lab to improve Second Life and better support these activities.

**Methods** This project is comprised of four main parts and takes part entirely online. You dont need to leave home or go anywhere to participate. You will be asked to:

1. Read and complete a consent form that expresses your willingness to participate in this research project.

2. Complete a brief Screening questionnaire on your Second Life experience, communication preferences, and computer equipment and settings.

3. Attend a single, one-hour class called Introduction to Building in Second Life.

4. After the class ends, complete one questionnaire and a few additional
questions about what you learned in class.

The data gathered and analyzed will consist only of statistical measures of questionnaire responses, measures of learning from the building class, and communication transcripts (voice, text, or both) from the building class.

**Your Anonymity** No identifying information about your avatar (besides the avatar name, which is shown in Second Life in your avatar name tag), your Second Life usage, or real-life data is needed or desired for participation in this study. A valid Second Life name is needed only for verification, to ensure that the name is a valid Second Life account and that you can log in to Second Life using that avatar name. During the data analysis phase of this project, all data will be tagged with unique numerical identifiers. No data will be tied to either your Second Life (avatar) name or real-world identity. We don’t want avatar names to skew the way researchers think about the data they need to analyze!

**Expectations**

**Communication Mode:** After submitting the Consent Form and Screening Questionnaire, you will be randomly assigned to one of three possible groups. Your group may be one that prefers to use text communication, voice communication, or both. Since group assignment is random, you MUST agree to use the communication mode used by your group. Please do NOT complete the Consent Form and Screening Questionnaire if you are not willing to or cannot use text communication, voice communication, or both for any reason.

**Class Attendance:** You will be asked to attend one of several available Introduction to Building classes based on your group assignment. You may choose the class time and date that is most convenient for you, but you must attend one of the classes listed. Your class list is based on the group to which you were randomly
assigned.

Please do NOT complete the screening questionnaire if you cannot commit to attending the Introduction to Building class for approximately one hour sometime during the next few weeks. Your attendance will be noted during the class. You only need to attend once.

**Time:** The total time expected from you is 2 hours (from screening questionnaire to post-class questions)

**Valid Email Address:** You will receive an email at the email address you provided in the Screening Questionnaire. This email will contain additional information on participation, a schedule (dates and times) of available Introduction to Building classes, and other relevant information.

**Frequently Asked Questions (FAQ)**

Q: Can my friends participate too?
A: Definitely! We need more people to join the study, so please give them the URL of this web page if you think you know someone who might be interested. Your friend can then complete the Consent Form and Screening Questionnaire and see if they meet the requirements.

Q: Who are you?
A: I am a Ph.D. candidate at the Illinois Institute of Technology, doing this project as part of my dissertation. I am also a fulltime employee at Linden Lab in San Francisco, and have been since February 2007.

Q: What will you do with this research?
A: As I write more about this work, I will publicly post any papers and welcome your comments. I also welcome your feedback at anytime at filiste@iit.edu or stephany@lindenlab.com. I am also often available in Second Life during standard
business hours on Pacific Time as Stephany Linden.

**Q: How do I know this is the real deal?**

A: If you have any questions about the validity of this study, please contact the Illinois Institute of Technology Institutional Review Board (IRB). You are welcome to contact Glenn Krell MPA, CRA, Executive Officer of the IIT Institutional Review Board (IRB) at Email: irb@iit.edu or Phone: (312) 567-7141

Thank you again for your interest in participating!
APPENDIX B

CONSENT FORM
Agreement to Participate In Second Life Communication Study

Stephany Filimon, Ph.D. Candidate
Lewis Department of the Humanities
Illinois Institute of Technology
3300 Federal St.
Chicago, IL 60616-3793

Description of Study

I understand that I have been asked to participate in a study of online communication and learning in Second Life as part of the above persons (henceforth, the researcher’s) dissertation research. My participation in this study will help the researcher to better understand some of Second Life’s strengths and weaknesses for instruction. I have been asked to spend about one hour in Second Life (online) participating in this test, and approximately 30-45 minutes completing online surveys and a post-activity assessment. This will involve my doing some or all of the following things:

- Using a computer
- Reading
- Logging in to Second Life, a 3D virtual environment
- Interacting with researchers in Second Life
- Potentially interacting with other students in Second Life
- Filling out questionnaires

The researcher will record learning activity transcripts, text and/or voice transcripts. The researcher will not have access to my real (legal) name. Neither my
real name, nor any other aspect of my real-world identity, will be associated with
my avatar name. Because I will be in Second Life, only my avatar will attend and
appear in the learning activity. When the researcher describes her work to other
people, she will only use my avatar name and will not use my real name.

**Risks and Benefits Expected**
The study will not do me any harm. I do not expect to experience any stress,
strain, or psychical or psychological injury beyond that I would normally encounter
using a computer. I may benefit from learning how to create and modify 3D
objects. The results may help the researcher to improve the design of Second Life
for instructional purposes.

**Negligence**
I absolve the Illinois Institute of Technology (IIT) of responsibility for any injuries
or medical conditions that I may suffer by during the course of my study
participation, unless those injuries or medical conditions are due to IIT’s negligence.

**Anonymity**
I understand that any information about me that is obtained from this study,
including what I say, will be anonymous. The researcher will not have access to my
real name, only to my avatar name. I do not need to appear anywhere in person in
order to participate in this study. Only my avatar will appear in Second Life and
the learning activity. Only my avatar name could be accessed and referenced in
parts of this study.

**Right to Refuse or End Participation**
I understand that my participation in this study is strictly voluntary. I understand
that I may refuse to participate in this study and/or stop participating at any time.
Certification

I certify that I have read and that I understand the foregoing, that I have been given satisfactory answers to my inquiries concerning this study, and that I have been advised that I am free to withdraw my consent and to discontinue participation in the project or activity at any time.

I herewith give my consent to participate in this study with the understanding that such consent does not waive any of my legal rights, nor does it release the researcher or any agent thereof from liability for negligence. I understand that I shall remain anonymous in all written and verbal reports of this study. I acknowledge that I have received and read this consent form.

Name

Date (If you cannot obtain satisfactory answers to your questions or have comments or complaints about your treatment in this study, please contact Stephany Filimon at 945 Battery St., San Francisco, CA 94111 Email: filiste@iit.edu or Phone: (415) 243-9000. You may also contact Glenn Krell MPA, CRA, Executive Officer of the IIT Institutional Review Board (IRB) at Email: irb@iit.edu or Phone: (312) 567-7141
APPENDIX C

SCREENING QUESTIONNAIRE
Please note that you must have a valid Second Life resident (user) name and an active Second Life account before completing this questionnaire. If you need to create a Second Life account, you can get a free one at www.secondlife.com.

1. What is your Second Life Resident Name? This will not be used for any identifying purposes, and only for automatic validation.

2. How old are you?
   a. 18-24
   b. 25-35
   c. 36-45
   d. 46-55
   e. 56-65
   f. 66-75
   g. 76-85
   h. 85+

3. Have you logged in to Second Life Before? If you answered no, please skip to question number 6.
   Yes/No

4. If you answered yes to question #2, how long have you been in Second Life?
   Please choose one.
   a. <2 hours = Little to no experience (newbie)
   b. 2-20 hours = Medium level of experience
   c. 20+ hours = High level of experience

5. Do you belong to any Groups in Second Life?
   Yes/No

6. If you answered yes to question #4, how many SL Groups do you currently
belong to? You don’t need to name them but you can if doing so makes it easier to answer the question. Knowing the names of groups is not important to this research and this information will be discarded.

7. What is your primary interest in Second Life? Please rate the list of items below in order of interest to you on a scale from 1-4 (1 = Not interested at all; 2 = Somewhat interested; 3 = Fairly interested; 4 = Extremely interested):
   a. Socializing
   b. Building/Scripting/Content Creation
   c. Running a business
   d. Providing education
   e. Receiving education
   f. Role-playing/fantasy activities
   g. Just doing this to participate in this study
   h. Other (please describe)

8. Do you use voice chat in SL? Please choose one:
   a. Always
   b. Sometimes
   c. Never

9. If you are assigned to use voice in the Instruction to Building class, will you accept assignment to your condition (either voice, text, or both)?
   Yes/No

10. Do you own a headset with a mic for speaking?
    Yes/No
11. What kind of computer will you use to participate in this study?
   a. Desktop PC/Windows
   b. Desktop Mac
   c. Desktop PC/Linux
   d. Laptop PC/Windows
   e. Laptop Mac
   f. Laptop PC/Linux
   g. Other (please describe)

12. How do you usually connect to the Internet? Choose one.
   a. Cable modem
   b. DSL
   c. Telephone line/modem
   d. T1
   e. Other (please describe)

13. How much memory does your computer have? Choose one.
   a. 16 MB
   b. 56 MB
   c. 512 MB or more
   d. 1 GB or more
   e. I dont know.
   f. I dont understand the question.

14. What do you usually use for your computer input device?
   a. Mouse
   b. Space Navigator Mouse
   c. Keyboard
   d. Neither/Other (please describe)
15. What time zone do you live in? For example, U.S. Pacific.

16. Please select the race with which you most closely identify.

17. What is your real-world gender? (M or F)
APPENDIX D

STUDY ACCEPTANCE AND REMINDER EMAIL MESSAGES
Hello,

Thank you for submitting a Consent Form and Screening Questionnaire for the study on communication and learning in Second Life, at www.iit.edu/ filiste. You have been accepted for participation in this study.

Please attend the Introduction to Building Class in Second Life on **Sunday, May 10 at 10 AM Pacific / 12 PM Central / 1 PM Eastern.** Please send me an email if you cannot attend, as this impacts study data and group size. Please read on for other important details on your participation.

What happens next?

1) First, make sure to create a Second Life account and download the Second Life Viewer, if you haven’t already. **You need an avatar name and the Viewer in order to enter Second Life.** You can download the Second Life Viewer at:

   http://secondlife.com/support/downloads.php

   You may wish to spend some time familiarizing yourself with Second Life and how to use it, but this is not required.

2) Please email me your avatar name when you have one. **I have no other way to locate you in Second Life and help you** move around in the virtual world.

3) Please log in to Second Life bf a few minutes before class is scheduled to begin on Sunday, May 10 at 10 AM Pacific / 12 PM Central / 1 PM Eastern. You may wish to add ”Stephany Linden” to your Second Life Friends List, to make it easier to find and communicate with me.

4) Using your avatar name, I will find you in Second Life shortly before class begins and teleport you to Pixie Island, which is a private area where the Introduction to Building Class will be held. Text (typing) and/or voice (speaking) communication will be used in class.
5) After class ends, you will be asked to complete an online survey (similar in nature to the Screening Questionnaire you’ve already submitted). The total expected time for these activities is approximately one hour (60 minutes). Before submitting the final survey, you’ll be asked to specify if you’d like $10 USD cash reimbursement via PayPal or a $10 USD Amazon email gift certificate.

Please feel free to email me if you have any questions in the meantime. Thank you for taking time to participate in this study!

Best,

Stephany Filimon

SL: Stephany Linden

email: filiste@iit.edu or stephany@lindenlab.com

study information: www.iit.edu/~filiste

mobile: 415 123-4567

AIM: stephanyfilimon

The reminder email was comprised of the above information plus this message:

Hello,

This is just a reminder about the Second Life Study that will take place this Sunday, May 10 at 4 PM Pacific / 6 PM Central / 7 PM Eastern.

All of the details about Sunday’s study are below. If you have not yet sent me your avatar name, please make sure to so. This is the only way for me to find you in Second Life and bring you to the class.

Please feel free to contact me with any questions you may have in the meantime, using any of the contact information at the bottom of this email.
All best,

Stephany
APPENDIX E

INSTRUCTOR SCRIPT FOR INTRODUCTION TO BUILDING CLASS
Welcome to the Introduction to Building in Second Life class, which covers the fundamentals of building 3D objects in Second Life. By the end of this class, you will know:

1. How to use tools in the Second Life interface to create basic 3-D objects, like cubes and spheres
2. How to combine basic 3D objects into more complex ones, such as using multiple rectangles to create a sofa.

The Object Editor (also called the "Build menu," which is what we’ll use in class) gives you access to the tools you need to create objects in Second Life. We’ll open it now, via View >Build. Now we’ll select a basic shape ("prim").

Right-click the ground you’re standing on, and choose Create to open the Object Editor. You can also press Ctrl+4. This works for us because we are standing on land that permits the creation of objects. Not all land does, depending on what the owner wants. So remember: You can only create objects on land that permits the creation of objects. Land that prohibits object creation is marked No Build. You can identify no-build land by the icon (a yellow cube in a red circle) that appears at the top of your screen when standing on no-build land.

The Build menu (Object Editor) should be open. Now, you will choose the type of basic shape you wish to create. Basic shapes are called "primitives" or "prims" for short. They are the building blocks of all content you see in Second Life. If you ever took an art class in which you had to look at a sketch or a painting, and reduce it to more primitive shapes (like spheres for a person’s head and joints, for example), you’ll understand the idea of prims.

In the Build menu, choose the basic shape you wish to create (cube, sphere, and so on). Any shape will do for now. Next, click the location inworld where you
wish to build your object. The shape (prim) you chose should appear in the location you selected (most likely with a resounding "whoosh" sound).

Now we’ll cover moving and modifying prims. First, you’ll learn to move your prim. When the "hand" icon is selected in the Build menu, you can click and drag your prim to move it around. Dragging moves the object along the horizontal (X/Y) plane. Hold down Ctrl to drag the object vertically (Z) as well. Clicking on the red (X), green (Y), and blue (Z) arrows lets you drag the object only along those axes.

You can also change the size of your prim: Hold down Ctrl+Shift to bring up the sizing box. If you’ve done it correctly, you should see a glowing white effect, with small cubes in four colors: red, green, blue and white. Click and drag one of the white corner boxes to scale the entire object proportionally. Click and drag a red, green or blue box to resize a prim’s length/width/height without changing the other dimensions.

Prims can also be rotated. This is helpful when, in building an object or building, or a chair as we’ll do later, you need to angle one piece to fit better with another. It’s also helpful for flipping prims. To rotate your prim: Hold down Ctrl to bring up the rotation sphere. It is a glowing white sphere with three colored rings (rotation axes) - red, green and blue. Click and drag anywhere within the sphere to rotate the object freely along all three axes. Click and drag a specific ring (red, green or blue) to rotate the object only around that axis.

In the three examples above, you learned how to move, change, and rotate your prim by clicking and dragging. But there is another way to do these same things with a higher level of accuracy: By entering specific values. Instead of clicking and dragging, you can enter numerical values to move, change, and rotate your prim. We’ll do that now.
Click on the More button in the Build menu, then click on the Object tab. Make sure your Object tab is highlighted. Along the left side, you will see boxes with X, Y, and Z beside them. Here, you can enter the specific XYZ coordinates to move, resize, or rotate the object. Changes to these values are always based on the center of the object (which is the point where the red/green/blue axes meet).

Now, we’ll make many objects into one object by linking them together. As we learned earlier, objects in Second Life are built out of basic shapes called primitives. You can link several primitives (prims) together to create one cohesive object, like a building. A linked object is, for all intents and purposes, considered one object: It has one name, acts as one object (for example, if you attempt to move the object, it will move as one entity and not multiple pieces), and it cannot be broken apart unless you Unlink it yourself. Linking is what makes multiple prims act as one coherent piece in Second Life, and we’ll do that now.

If you do not already have two prims inworld, please create a second one by selecting it in the build menu and clicking inworld. If you are not in the object Editor already, right-click any object and choose Edit, or open the Editor with Ctrl-3. Click on any empty area to make sure you do not have any objects selected. Hold down the Shift key, and click on each prim you want to link together one at a time. You can also click Edit and drag a yellow box (as you would if you were to crop a photo in another piece of software).

You must select the entire prim: Dragging the yellow box around the part of the prim will not select the whole prim. You will know you’ve selected an entire prim, and ALL the prims you intended to select, if they glow pale blue. If a prim does not pale blue it means you did not fully select it. Please let me know if you need help.

Once you’re sure you’ve selected all the prims you’d like to link together, go to
the Tools menu (along the top of the Second Life viewer) and select Link, or just press Ctrl+L. Don’t worry if you make a mistake: You can undo a mistake by selecting Tools-Unlink or press Ctrl+Shift+L to break the object apart into separate prims again.

We’ll use linking when we create a chair from separate prims, which we’ll do now. Take a moment to think about what you might want your chair to look like. A big, orange sample chair is shown beside me. It was created with two HemiSpheres, for the seat and back, and four cylinders, for legs. If you’d like, you can build a chair just like the sample one or design and build your own. As long as it includes a seat, back and legs that you later link together, you’ll have used the fundamental concepts of creating 3D objects in Second Life.

First, we’ll make the chair seat. Select a HemiSphere if you’re making a chair like the big orange one and rez it inworld. Now, stretch or resize the HemiSphere to be the size you want, just like we did earlier in class. The easiest way was to hold down Ctrl+Shift to bring up the sizing box, see a glowing white effect, with small cubes in four colors, and click and drag one.

Once your HemiSphere is the right size, make a copy of it by right-clicking on it. You’ll see a pie menu. Click ”More” on the pie menu and then click ”Take Copy.” Your object has been copied into your Inventory, and you’ll need to retrieve your copy from your Inventory to make the back of the chair.

To retrieve the object you copied, click on the Inventory button in the lower right corner of the screen. The Inventory menu will appear. Go to the My Inventory Objects folder. You’ll probably see a list of multiple objects that all have the same name ”Object.” It’s not very helpful. The first object item on the top of the list will have the most recently copied item in it. Click on the first object in the list. Drag
right to the ground near your avatar.

Click the copied object (the one you just dragged on to the ground) and check "Rotate" on the Object Editor menu. Click on the red rotation strip. A sphere will appear but remember, don’t worry: Your cube/rectangle/HemiSphere/whatever prim you selected has not turned into a sphere. You’re still editing the prim you chose.

Drag the rotation strip downward to rotate the support piece in the same direction. Your prim will now be angled to look like the back of a chair. Don’t forget to use the arrow keys in the Camera Controls as you need to to change your view of the object and make it easier for you to see what you’re doing. Then make sure Position is on, click your prim, and drag the chair back to the chair seat, bringing the two objects close together so it looks like a chair seat with an angled chair back. Once the prims are close together, rotate them as needed to adjust the angle of your chair back.

Now we’ll add some chair legs. Right-click on the ground and select Create. Choose the cylinder from the available prims in the Object Editor (or any prim you’d like to make chair legs from). Uncheck the Position Radio button. Check the Stretch option. Select/click on the Cylinder. Stretch the cylinder up and down (to make it longer/taller and skinnier). Make the chair leg the appropriate size given how high your chair seat is off of the ground (to fit the space between the chair seat and the ground).

Highlight the Edit icon in the Object Editor. Draw a rectangle around your chair a lot like select all by drawing a rectangle around it. With the chair selected, click on Position and move the chair over the leg (most likely using the blue and red lines to go up and over). Move the chair over to rest on top of the chair leg you just finished. Use Position on the chair and chair leg to bring them closer together as
needed, as you did for the chair seat and back.

Make sure your chair and all parts of it are de-selected (make sure you’re not clicking on any of them right now). To make sure you’re not, you can just click an empty patch of ground.

Make sure your chair leg is the exact size you want it to be, because you’ll be making exact copies of it. Now right-click on the chair leg. The pie menu will appear. Select ”More” from the pie menu. More options will appear in the pie menu. Select ”Take Copy” at the top of the pie. Click on the ”Inventory” button (while you still see little white dots coming out of your avatar’s arm). The copy of the chair leg will now be in your My Inventory <Objects folder and will be the first ”Object” named ”Object” in the list. Select the first Object in the list and drag it to the ground.

Do this two more times, so you have a total of four chair legs: The original one that is already ”attached” to the chair, and the three copies you just made and dragged to the ground. Close the Inventory menu. Using ”Position”, select one of the chair legs you just dragged out of the Objects folder. Slide it over to rest on the ground beneath the chair. Move the rest of the legs under desired corners of the chair seat. When your chair legs are properly positioned and the chair looks the way you want it to, make sure Position is on.

Now, we’re going to link the prims together to make one chair that behaves as one would expect a chair to (i.e. it allows you to sit on it). Linking is the equivalent of gluing chair parts together in real life. We want to make sure the chair stays together and moves around as a single object, without parts coming off. Drag the mouse over the chair to highlight all of its pieces. Then click Tools  Link to link all the components into a single object. Click on the chair. You’ll see it highlighted in light blue.
Now your avatar can sit down: Right-click on your new chair and select "sit here."
APPENDIX F
POST-ACTIVITY PRESENCE QUESTIONNAIRE
In the online version of this questionnaire, each of the questions below was followed by seven radio buttons in a Likert-style scale, from "least" to "most". Each radio button corresponded to a numerical value on a seven-point scale: The first button of the seven (corresponding to "not at all"), for example, had a value of "1" while the fifth box ("moderately") had a value of "5", and so on. Participants could only select one button on the seven-point scale. The total survey score is the sum of all responses.

Directions
For each question below, please select the button that most accurately describes your experience in the Introduction to Building class you just completed.

1. How much were you able to control events?

2. How responsive was the environment to actions that you initiated (or performed)?

3. How natural did your interactions with the environment seem?

4. How much did the visual aspects of the environment involve you?

5. How much did the auditory aspects of the environment involve you?

6. How natural was the mechanism which controlled movement through the environment?

7. How compelling was your sense of objects moving through space?

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

9. Were you able to anticipate what would happen next in response to the actions that you performed?
10. How completely were you able to actively survey or search the environment using vision?

11. How well could you identify sounds?

12. How well could you localize sounds?

13. How well could you actively survey or search the virtual environment using touch? (omitted)

14. How compelling was your sense of moving around inside the virtual environment?

15. How closely were you able to examine objects?

16. How well could you examine objects from multiple viewpoints?

17. How well could you move or manipulate objects in the virtual environment?

18. How involved were you in the virtual environment experience?

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

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

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

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

23. How much did the control devices interfere with the performance of assigned tasks or with other activities?
24. How well could you concentrate on the assigned tasks or required activities rather than on the mechanisms used to perform those tasks or activities?

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

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

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

28. Was the information provided through different senses in the virtual environment (e.g. vision, hearing) consistent?
APPENDIX G
ADDITIONAL PRESENCE ITEMS FROM GERHARD
In the online version of this questionnaire, each of the questions below was followed by seven radio buttons in a Likert-style scale, from "least" to "most". Each radio button corresponded to a numerical value on a seven-point scale: The first button of the seven (corresponding to "not at all"), for example, had a value of "1" while the fifth box ("moderately") had a value of "5", and so on. Participants could only select one button on the seven-point scale. The total survey score is the sum of all responses.

1. How stimulating was the design of the virtual world (Second Life)?

2. How natural was the mechanism that controlled the actions of your avatar?

3. How responsive were the avatars of other participants to verbal communication that you initiated? (Option: I did not initiate verbal communication with other avatars.)

4. How natural did your communication with other participants seem? (Option: I did not communicate with other avatars.)

5. How compelling was your sense of being present in a virtual world?

6. How compelling was your sense of other participants being present?

7. How credible were the avatars of other participants with respect to representing human beings?

8. How aware were you of the existence of your own avatar?

9. How easy was it to distinguish between the avatars of different participants?

10. How easy was it to control your avatar?

11. How well could you concentrate on communication and the assigned task rather than on the mechanisms used to perform these?
APPENDIX H

COGNITIVE LOAD MEASURES
In the online version of this questionnaire, each of the questions below was followed by seven radio buttons in a Likert-style scale, from "least" to "most". Each radio button corresponded to a numerical value on a seven-point scale: The first button of the seven (corresponding to "not at all"), for example, had a value of "1" while the fifth box ("moderately") had a value of "5", and so on. Participants could only select one button on the seven-point scale. The total survey score is the sum of all responses.

1. How difficult was the Introduction to Building class?

2. How much mental strain did you experience during the Introduction to Building Class?

3. Please describe the quality of the instruction you received.
APPENDIX I

RETENTION MEASURES
Questions designed to assessed learning are presented below. The following eight attributes are listed below each question:

- **Learning Objective:** A brief statement that describes what participants should be able to do after completing the learning activity (instruction).

- **Learning Outcome:** The type of knowledge that participants will demonstrate by answering the question correctly. Possible learning outcomes of the learning activity presented here are: attitudes (see questions 1-3 above), declarative knowledge, concepts, and procedural learning.

- **Assessment Item:** The question participants are asked to answer.

- **Directions:** Instructions to participants on how to complete the question.

- **Question Form:** For the assessment instrument presented here, multiple choice with foils (distractors) is the most common question form; ordering steps in a procedure is also employed.

- **Number of Items:** The number of answers and foils (distractors).

- **Response Characteristics:** Statement of correct and incorrect answers (foils) for third-party/independent scoring.

- **Scoring Instructions:** The number of points to assign to correct and incorrect answers (foils) for third-party/independent scoring.

**Note:** In its final format, Question #4 consisted of images rather than a text list containing the names of shapes, as shown here.

**Question #4:** Which of the following shapes are "prims" in Second Life?
Please check all that apply:
Cube
Prism
Cylinder
Cone
Sphere
Trapezoid
Square
Circle
Triangle
Rectangle

**Learning Objective:** Given a collection of shapes (images), the participant can correctly identify which shapes are prims, the building blocks of all 3D content in Second Life.

**Learning Outcome:** Declarative Knowledge: Participants can summarize or recall information from the learning activity. Declarative knowledge objectives reflect whether learning will be recognition (choosing from options) or recall, verbatim or paraphrased, and listed or summarized (Smith & Ragan, 2005). For this question, declarative knowledge learning will be recognition and listed (in that participates will choose images from a list, or collection). Concept Learning: Concepts reflect learners’ ability to classify and label ideas, objects and events as examples or non-examples of a concept (Smith & Ragan, 2005). For this question, prims are a concept that is the most fundamental concept to building in Second Life.

**Assessment Item:** Which of the following shapes are "prims" in Second Life?
Directions: Look at the collection of shapes shown below. Check the box beside each shape that is a prim in Second Life.

Question Form: Multiple choice with five foils (distractors). All of the distractors are shapes that are similar to prims in Second Life, but are 2D instead of 3D and are not shown in the Build menu (Object Editor window) participants see and use during the learning activity. Question #11 contains information that would help participants answer this question (#4). For this reason, participants will not be able to return to prior questions after they have submitted answers and moved on to subsequent sections.

Number of Items: 10 (5 correct answers with 5 foils)

Response Characteristics: The correct answers are: Cube, cylinder, prism, sphere, cone The incorrect answers (distractors) are: Square, circle, triangle, rectangle, trapezoid

Scoring Instructions: Each correct answer a participant selects is worth 1 point. Since there are 5 correct answers, the maximum possible number of points for correct answers is 5 (5 correct answers x 1 point each = 1 point). Each correct answer a participant misses (fails to select) is worth 0 points. Each foil selected is also worth 0 points. Since there are five foils, the maximum possible number of points for selected foils is 0 (5 foils x 0 points each = 0 points).

Question #5 Read the incomplete sentence below. Choose one phrase that most accurately completes the sentence.

The symbol of a _______indicates land on which 3D objects cannot be built.
A. White hand in a red circle
B. Yellow box in a red circle
C. Yellow cylinder in a red circle
D. Blue house in a red circle

**Learning Objective:** Given a list of descriptions of symbols, the participant can correctly identify which symbol designates a land area in Second Life, an area in which they could repeat the skills they've just learned. Learners will demonstrate the ability to identify a necessary condition for repeating the skill they've just learned.

**Learning Outcome:** Declarative Knowledge: Participants can summarize or recall information from the learning activity. Declarative knowledge objectives reflect whether learning will be recognition (choosing from options) or recall, verbatim or paraphrased, and listed or summarized (Smith & Ragan, 2005). For this question, declarative knowledge learning will be recognition and listed (in that participate will choose the correct answer from a list).

**Assessment Item:** The symbol of a _____ indicates land on which 3D objects cannot be built.

**Directions:** Read the incomplete sentence below. Choose one phrase that most accurately completes the sentence.

**Question Form:** Multiple choice with three foils (distractors). All of the distractors are non-existent symbols that sound similar in appearance to the real symbol pointed out during the learning activity. The foils are not symbols that any of the participants would have seen.

**Number of Items:** 4 (1 correct answer with 3 foils)

**Response Characteristics:** The correct answer is: B. Yellow box in a red circle The incorrect answers (foils) are: A, C and D

**Scoring Instructions:** The correct answer (B) is worth 1 point. The incorrect answers (A, C and D) are worth 0 points.
**Question #6** After making a copy of an object, you can retrieve it by following which path? Please choose one item that best reflects your answer.

A. Inventory > My Inventory > Objects  
B. Edit > Preferences > Graphics  
C. World > Region/Estate > Ground Textures  
D. View > Beacons > Physical Objects

**Learning Objective:** Given a list of possible paths to retrieve a copy of a 3D object, the participant can correctly recall the path to the objects location.

**Learning Outcome:** Declarative Knowledge: Participants can summarize or recall information from the learning activity. Declarative knowledge objectives reflect whether learning will be recognition (choosing from options) or recall, verbatim or paraphrased, and listed or summarized (Smith & Ragan, 2005). For this question, declarative knowledge learning will be recognition, paraphrased, and listed.

**Assessment Item:** After making a copy of an object, you can retrieve it by following which path?

**Directions:** Read the following statement and the list of items below it. Choose only one item.

**Question Form:** Multiple choice with three foils (distractors)

**Number of Items:** 4 (1 answer with 3 foils)

**Response Characteristics:** The correct answer is:

A: Inventory > My Inventory > Objects  
The incorrect answers (foils) are: B, C, and D

**Scoring Instructions:** The correct answer (A) is worth 1 point. The incorrect answers (B, C and D) are worth 0 points.
Question #7 Read the statement below. Choose the answer that most accurately completes this statement. Choose only one answer.

You can change the size of a prim by:
A. Clicking and dragging to resize the prim’s length, width or height
B. Clicking and dragging a white corner box to scale the length, width and height of the object proportionally
C. Entering specific X, Y, and Z coordinates to resize the object
D. All of the above

Learning Objective: Given a list of possible methods for resizing 3D objects, the participant can identify the correct resizing techniques in Second Life.

Learning Outcome: Declarative Knowledge: Participants can summarize or recall information from the learning activity. Declarative knowledge objectives reflect whether learning will be recognition (choosing from options) or recall, verbatim or paraphrased, and listed or summarized (Smith & Ragan, 2005). For this question, declarative knowledge learning will be recognition and paraphrased. Concept Learning: Concepts reflect learners ability to classify and label ideas, objects and events as examples or non-examples of a concept (Smith & Ragan, 2005). For this question, resizing is a fundamental concept to building with prims in Second Life (moving from a more basic shape to a more specific one).

Assessment Item: You can change the size of a prim by:

Directions: Read the statement below. Choose the answer that most accurately completes this statement. Choose only one answer.

Question Form: Multiple choice with no foils (distractors)

Number of Items: 4 (4 technically correct answers with no foils): Though
each answer is technically correct, one answer is the most correct answer. The others are correct but to a lesser degree; true statements but not the ideal, most accurate answer.

**Response Characteristics:** The correct answer is: D. All of the above. The incorrect (incomplete) answers are: A, B, and C. A, B, and C are technically correct, but D is the most complete, accurate response.

**Scoring Instructions:** The best answer, D, is worth 1 points. Other answers (A, B and C) are worth .5 points.

**Question #8** Look at the image below. Three lines are shown: a red line, a green line, and a blue line. Each line represents one angle, or axis. Choose the color of the line that represents the Z axis. A. Red B. Green C. Blue

**Learning Objective:** Given an image containing horizontal (X/Y) and vertical (Z) axes, the participant will be able to correctly identify the Z (vertical) axis.

**Learning Outcome:** Declarative Knowledge: Participants can summarize or recall information from the learning activity. Declarative knowledge objectives reflect whether learning will be recognition (choosing from options) or recall, verbatim or paraphrased, and listed or summarized (Smith & Ragan, 2005). For this question, declarative knowledge learning will be recognition.

**Concept Learning:** Concepts reflect learners ability to classify and label ideas, objects and events as examples or non-examples of a concept (Smith & Ragan, 2005). For this question, three dimensionality and axes are a fundamental concept to building and changing prims, and to understanding all 3D content.

**Assessment Item:** Choose the color of the line that represents the Z axis.

**Directions:** Look at the image below. Three lines are shown: a red line, a
green line, and a blue line. Each line represents one angle, or axis. Choose the color of the line that represents the Z axis.

**Question Form:** Multiple choice

**Number of Items:** 3 (1 correct answer and 2 incorrect, with no foils). It is difficult to show a foil since this question is visual in nature, and the 2 incorrect answers are not foils for the concept of vertical axis; they are horizontal, and thus not a good foil for vertical.

**Response Characteristics:** The correct answer is: C: Blue The incorrect answers are A and B.

**Scoring Instructions:** The correct answer (C) is worth 1 point. The incorrect answers (A and B) are worth 0 points.

**Question #9** Look at the same image you examined for question #5, above. The image contains three lines: a red line, a green line, and a blue line. Each line represents one angle, or axis. Along which axis would you stretch the cylinder if you wanted to make it taller? Choose only one answer. A. Red B. Green C. Blue

**Learning Objective:** Given an image containing horizontal (X/Y) and vertical (Z) axes, the participant will be able to correctly identify the Z (vertical) axis (the axis that one would use to make a given object taller).

**Learning Outcome:** Declarative Knowledge: Participants can summarize or recall information from the learning activity. Declarative knowledge objectives reflect whether learning will be recognition (choosing from options) or recall, verbatim or paraphrased, and listed or summarized (Smith &Ragan, 2005). For this question, declarative knowledge learning will be recognition. Concept Learning: Concepts reflect learners ability to classify and label ideas, objects and events as examples or
non-examples of a concept (Smith & Ragan, 2005). For this question, three dimensionality and axes are a fundamental concept to building and changing prims, and to understanding all 3D content.

It is important to provide multiple assessment items to assess key concepts in learning. This questions gives participants another opportunity to show that they can identify the vertical axis and a way in which the axis is useful, even if participants dont know that the vertical axis is also the Z axis. Confusion with terminology could be a reason why participants answer the previous question incorrectly.

**Assessment Item:** Along which axis would you stretch the cylinder if you wanted to make it taller?

**Directions:** Look at the same image you examined for question #5, above. The image contains three lines: a red line, a green line, and a blue line. Each line represents one angle, or axis. Along which axis would you stretch the cylinder if you wanted to make it taller? Choose only one answer.

**Question Form:** Multiple choice

**Number of Items:** 3 (1 correct answer and 2 incorrect, with no foils). It is difficult to show a foil since this question is visual in nature, and the 2 incorrect answers are not foils for the concept of vertical axis; they are horizontal, and thus not a good foil for vertical.

**Response Characteristics:**
The correct answer is: C: Blue
The incorrect answers are A and B.

**Scoring Instructions:**
The correct answer (C) is worth 1 point.
The incorrect answers (A and B) are worth 0 points.

**Question #10** Below is the list of steps required to construct a table in Second Life. The steps shown here are listed in the incorrect order. Please re-order the list of steps to reflect the order in which you would need to perform them to successfully build a table with an oval top and three cylindrical legs. Begin with #1 for the first step and finish with #11 for the last step in the procedure. Select the Cylinder prim in the menu.

1. Make two copies of the table leg.
2. Click the location in world where you want to build your object.
3. Highlight all pieces of the table and link them together to make a single object.
4. Click to create the tabletop.
5. Stretch the sphere to make the tabletop.
6. Create one table leg.
7. Retrieve the object copies and place them beside the original table leg.
8. Click the location in world where you want to build your object.
9. Open the Object Editor.
10. Select the Sphere prim.

**Learning Objective:** Given a list of steps in random order, participants will be able to re-order the steps to create the correct procedure to build a given 3D object.

**Learning Outcome:** Procedural Learning: Describes what learners can do to demonstrate that they can successfully complete a procedure defined by a procedural rule. Participants should be able to demonstrate that they can recall the steps of a procedure, apply the steps in order, and confirm that the end result is reasonable (Smith & Ragan, 2005).
Assessment Item: Please re-order the list of steps to reflect the order in which you would need to perform them to successfully build a table with an oval top and three cylindrical legs.

Directions: Below is the list of steps required to construct a table in Second Life. The steps shown here are listed in the incorrect order. Please re-order the list of steps to reflect the order in which you would need to perform them to successfully build a table with an oval top and three cylindrical legs. Begin with #1 for the first step you would take, with #7 marking the last step in the procedure.

Question Form: Ordered steps/ordered procedure

Number of Items: 11 steps

Response Characteristics: Based on the content given in the class, there are two "strongly" correct sequences in which to order the steps of the procedure. There are two correct sequences because participants can build either the tabletop first (before the table legs) or the table legs first, before the tabletop. Either choice affects the outcome in the same way, and neither sequence is more correct than the other. The two sequences that will be considered correct answers for the purposes of scoring this question are:

Correct Sequence #1
Open the Object Editor.
Select the Cylinder prim in the menu.
Click the location in world where you want to build your object.
Create one table leg.
Make two copies of the table leg.
Retrieve the object copies and place them beside the original table leg.
Select the Sphere prim.
Click the location in world where you want to build your object.
Click to create the tabletop.
Stretch the sphere to make an oval tabletop.
Highlight all pieces of the table and link them together to make a single object.

Correct Sequence #2
Open the Object Editor.
Select the Sphere prim.
Click the location in world where you want to build your object.
Stretch the sphere to make an oval tabletop.
Click to create the tabletop.
Select the Cylinder prim in the menu.
Click the location in world where you want to build your object.
Create one table leg.
Make two copies of the table leg.
Retrieve the object copies and place them beside the original table leg.
Highlight all pieces of the table and link them together to make a single object.

Scoring Instructions: 11 answers with no foils (distractors). Answers are correct only in the context of order.

Using the sequences shown above (in the Response Characteristics section)
give: 1 point for each item that is in the correct position 0 points for each item that is not in the correct position (i.e. a position in the list that does not match either of the two possible sequences given above).

Exception Note: The Stretch the sphere to make an oval tabletop step can correctly appear at almost any point in the process: The placement of this step is correct as long as it occurs AFTER Click the location in world where you want to build your object and BEFORE Highlight all pieces of the table and link them together to make a single object. Because of this, this step in the process essentially amounts to one free/given point for this question.

Note: As with Question #4 (above), in its final format Question #11 consisted of images rather than a list of the names of shapes, as shown here. Images are not yet shown because I cannot get the images to appear in LaTeX yet.

**Question #11** You are preparing to build a covered bridge for a friend in Second Life. Your friend would like you to build a covered bridge that looks the one shown in the image below. In order to build the covered bridge shown here, you must first identify the prims (the more basic shapes) you will use to build it.

Look carefully at the covered bridge shown above. Select each prim you will use to create the bridge from the collection of shapes shown below. Select as many prims as you need in order to create the bridge. If one prim will be used more than once, only select the prim once.

**Learning Objective:** When shown a completed 3D object, participants will be able to identify the component parts (prims) of which the object is comprised.

**Learning Outcome:** Domain-Specific Problem Solving: Indicates that participants can assess the problem situation, determine which rules are applicable, and synthesize the rules to solve a particular problem in a specific domain (Smith & Ra-
In this case, the domain is highly specific: breaking more complex objects down into primitive shapes in order to build them in Second Life.

**Assessment Item:** Select each prim you will use to create the bridge from the collection of shapes shown below.

**Directions:** Look carefully at the covered bridge shown above. Select each prim you will use to create the bridge from the collection of shapes shown below. Select as many prims as you need in order to create the bridge. If one prim will be used more than once, only select the prim once.

**Question Form:** Multiple choice with no foils

**Number of Items:** 13 (6 correct and 7 incorrect)

**Response Characteristics:** Correct answers are all of the following (all because the bridge cannot be constructed without each of these required prims): Cube Prism Pyramid Tetrahedron HemiCylinder Award 1 point for each correct answer.

Incorrect answers (shapes that would not create parts of the bridge shown here) are: Cylinder Cone HemiCone Sphere Torus Ring Tube Award 0 points for each incorrect answer. Scoring Instructions: Award 5 points for each correct answer. Award 0 points for each incorrect answer.

**Instrument Scoring Instructions (to obtain total retention score)**
Follow the Scoring Instructions provided with each item (question) above. Add (sum) all of the points for a total final score.

**Mastery Criteria** Given the nature of content being assessed (i.e. it is not life or death instruction like air traffic control or intensive care nursing), a mastery level of 70% or higher is considered acceptable.
BIBLIOGRAPHY


