Running Head: Learning Efficiency and Efficacy in a MUVE

Learning Efficiency and Efficacy in a Multi-User Virtual Environment

Doug Hearrington

Kennesaw State University

Abstract

This study used the commercial multi-user virtual environment (MUVE) known as Second Life as an experimental platform for online course delivery with 17 graduate students. Participants interacted, constructed virtual buildings, and attended class in the virtual environment. General computer self-efficacy, MUVE self-efficacy, and learning efficiency were measured. Results indicate increased performance, decreased mental effort, improved learning efficiency, increased self-efficacy, and high self-reported engagement after a six-week usage period. Findings imply that Second Life has potential as a distance-learning platform, but that preparing educators to create simple simulations in the current version of this MUVE is likely to require a significant amount of scaffolding over a period of several weeks.

Learning Efficiency and Efficacy in a Multi-User Virtual Environment

In recent years, an increasing number of colleges, universities, and educational institutions have started to use multi-user virtual environments (MUVE) to host their online content and provide virtual educational interactions. Second Life, a popular free and commercially available MUVE, has attracted the attention of educational institutions world-wide and is currently being used by at least 142 colleges and universities, 41 for profit and non-profit educational organizations, 8 libraries, and 4 museums (SimTeach, 2008). Second Life is also finding increasing favor among secondary schools and is being utilized for educational purposes by schools across the globe. This study utilized the MUVE provided by Second Life.

A MUVE is a form of Virtual Reality (VR), a computer-based technology that provides visual, aural and tactile stimuli of a virtual world generated in real time (Sanchez, Lumbreras, & Silva, 2001). Users of a MUVE are provided with a 360-degree, three-dimensional environment in which users experience three “presence layers” seemingly causing physical and virtual realism to combine in the virtual space to produce an immersive experience that “conveys a feeling of being there and a strong sense of co-presence when other avatars are present (Warburton, 2009, p. 6).” The three layers of presence are physical, communication, and status (Warburton, 2009). A physical presence is created by the visual and physical proximity of avatars to one another. The communication presence is created by spatially enhanced voice – one can sense the direction from which another’s voice is originating – communication, synchronous text chat, and asynchronous text communications such as email and group notices. Finally, a status presence is created through various tools that allow one to know when a friendly avatar is “in-world” or offline (Warburton, 2009). Presence helps to establish a first-person experience when one is in a

MUVE. This experience has been said to enable direct, subjective and personal knowledge to develop (Sanchez et. al, 2001).

Barriers to learning in the Second Life MUVE have been documented in the literature. They include a “high learning curve,” interface difficulty, difficulty building, perceptions that working in the Second Life MUVE was too time-consuming, and technical difficulties (Sanchez, 2009). Additionally, some users have expressed disappointment with the environment because of their expectations that it was more like a playing a game. When the Second Life turned out to not be a game, these users became bored and frustrated (Sanchez, 2009).

Positive aspects of using the Second Life MUVE have been noted as well. Users enjoy creating and designing their avatar, the feelings of creativity and accomplishment they experience when building in the environment (Sanchez, 2009). Additionally, users have reported having a strong attachment to their avatar and enjoyed communicating with others via their avatar. Some have noted that the sense of enjoyment and creativity they experience in this MUVE outweighs the sense of frustration they feel from the complexity of the user interface and the technical issues associated with using Second Life (Sanchez, 2009).

The immersive nature of a MUVE combines these physical, social, and cultural dimensions to provide a space in which compelling simulations and role-playing activities may take place (Warburton, 2009). There is also evidence that collaborative interactions in MUVEs can result in conceptual change (Jackson and Fagan, 2000). However, research on the use of modern MUVEs as distance education platforms is in its infancy and little is known about the “learning curve” educators may experience when trying to create simple simulations in such an environment or whether the computer self-efficacy of an educator has any relationship with her or his ability to learn to build a simulation in a MUVE. This study advances the field in this area by providing a first step in this direction. In addition, this study represents an attempt to pilot test an instrument to measure self-efficacy for using a MUVE.

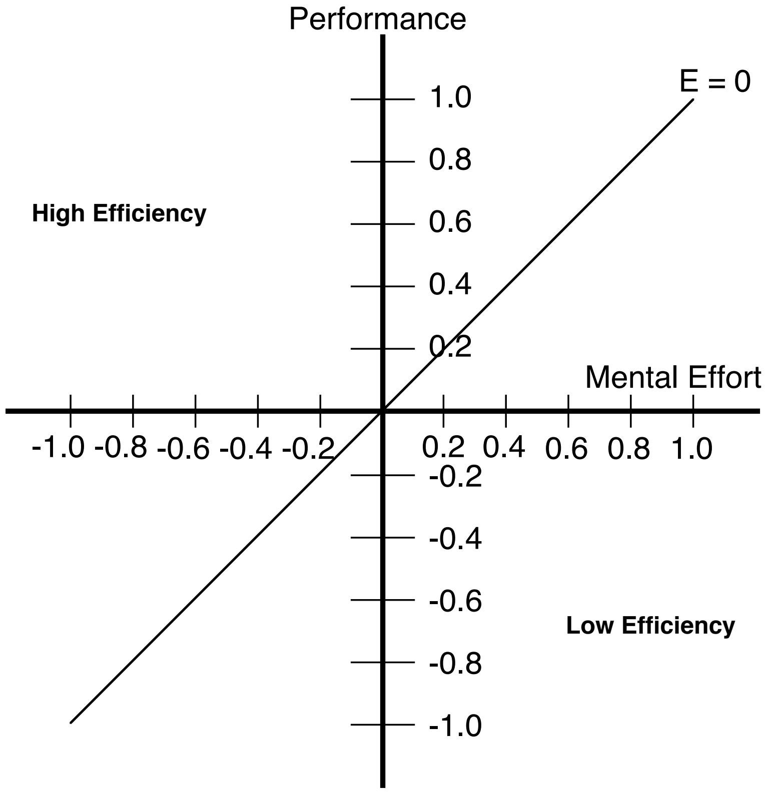
Perceived self-efficacy is one's belief in one's ability to complete actions required to produce a result or to accomplish a given task. Self-efficacy influences the careers people pursue, the level of effort they invest in a given endeavor, their resilience to adversity, and the level of accomplishments they achieve (Bandura, 1997). Therefore, it is possible that self-efficacy will be related to learning efficiency or the perceived difficulty of using the environment. For this purpose, a measure of general computer self-efficacy was administered and a specific self-efficacy instrument related to using a MUVE was created and administered by the researcher.

Since users interact with a MUVE's three-dimensional world and other users by means of avatars, these environments may present a higher level of cognitive load than the more typical two-dimensional, non-avatar, computer user interfaces with which most are familiar. Cognitive load refers to the level of difficulty a learner perceives when performing or learning a given task. Cognitive load theory is based on information processing theory in which an individual must process information using short-term memory in order to place it into long-term memory for later use. If the cognitive load is too great, the information will not be processed into long-term memory efficiently and learning will be inhibited. Cognitive load theory is a set of principles used by researchers to study the three types of load: (a) intrinsic load, which is the mental work imposed by the complexity of a task and is largely determined by one's goals; (b) germane (relevant) load, is the mental work imposed by an instructional activity that benefits the learning of the task; and (c) extraneous (irrelevant) load is mental work that is irrelevant to the learning goal and wastes limited mental resources (Clark, Nguyen, & Sweller, 2006). Cognitive load is operationalized in this study as perceived mental effort (Clark, Nguyen, & Sweller, 2006). The level of perceived mental effort can enhance or inhibit one's learning efficiency.

Any task that can be accomplished, such as learning to perform a mathematical calculation or learning to build a chair in a virtual world, can be measured in terms of learning efficiency. Learning efficiency is the relationship between a measurement of achievement, such as a test score or the amount of time it takes a learner to correctly perform a task, and the perceived mental effort (PME) of the learner. To measure learning efficiency, achievement and PME scores are converted to Z scores. Both are plotted on a Cartesian grid with PME on the X axis and achievement on the Y-axis. Z scores for PME that are above the mean are plotted to the right of the origin of the grid. Z scores for achievement (completion time) that are above the mean are reversed, because negative scores represent higher achievement. The hypothetical line of zero efficiency runs diagonally from a point at the lower left of the grid (quadrant III), through the origin, to the upper right of the grid (quadrant I) along a line that would extend through points (-1, -1) and (1, 1). See Figure 1 to see an example of the learning efficiency grid.

Figure 1

The Learning Efficiency Grid



Low learning efficiency occurs when something is learned slowly with great mental effort. High learning efficiency occurs when something is learned quickly with low mental effort (Clark, Nguyen, & Sweller, 2006). Thus, learning efficiency can potentially be used to calculate the learning "curve" of a given task over time and to compare different programs to one another.

While learning efficiency is a function of achievement and mental effort, it cannot tell us, by itself, how much more quickly one can perform a task after initial familiarity and over time. This may be accomplished by calculating the learning curve. Learning curve is the measurement of time to complete a specific task correctly over time, after practice. An illustration of the learning curve, taken from *The Learning Curve Deskbook* (Teplitz, 1991), assumes a piano student was learning to play the "Minute Waltz" by Chopin. The first time she played the piece, it took three minutes to play. Her second attempt took 2.37 minutes. Attempt number three took

2.6 minutes. The fourth attempt took under two minutes. The rate of improvement of the student, calculated each time the attempt doubled (attempt 1, 2, 4, 8, etc.), is 21 percent. This means that each doubling results in an improvement of 79 percent. Although the improvement itself gets smaller and smaller, rates of improvement tend to remain the same. This has been shown to remain constant in a variety of learning, manufacturing, and business situations (Teplitz, 1991). The learning curve has been used in manufacturing to help calculate production time and cost, to forecast labor requirements, and as a metric by which managers monitor production (Yelle, 1979). The same concept may also be applied to calculate the "forgetting curve" of a task (Bailey, 1989). The learning curve has been used to study the improvement in learning of computer-aided design (CAD) students over time (Hamade, Artail, & Jabar, 2005), and the improvement in disease pattern recognition and diagnosis by medical students over time (Williams, Klamen, & Hoffman, 2008). The learning curve is useful for describing and studying tasks requiring both procedural and declarative knowledge (Hamade, Artail, & Jabar, 2005). Therefore, it makes sense to apply the learning curve to the learning of the mostly procedural knowledge required to perform tasks in a virtual world. Learning curve may be calculated using various slope formulae (Yelle, 1979; Teplitz, 1991) or as a percentage of improvement.

Based on the increased interest in MUVEs in higher education as a distance learning platform and the increasing use of MUVEs at all levels of education for building and using simulations, this study was designed to answer these questions:

1. How efficient are new participants at creating and working in the MUVE and how does their efficiency change over time?
2. What is the relationship between participant general computer self-efficacy and MUVE self-efficacy before and after using the environment for a period of time? Does using the

MUVE result in increased GCSE and MUVE-SE? Does GCSE or MUVE-SE predict learning efficiency?

3. What were the participants' impressions of their user experience within the Second Life MUVE?

Method

Participants

Data were collected from 17 graduate students enrolled in a masters degree program in educational leadership in a southeastern state university. The students were enrolled in the same section of a class and therefore represent a convenience sample. Two of the participants were male. Eight identified themselves as African American and the remaining seven identified themselves as Caucasian. Other demographic characteristics of the participants are summarized in Table 1.

Table 1

Participant demographics

	Mean	Standard Deviation	Range
Age in years	32.6	6.3	25 - 47
Years of teaching experience	9.0	5.7	1 - 21

Instruments

Cognitive load was measured for two defined tasks within the MUVE using an established nine-point scale ranging from one (1), very little mental effort, to nine (9), representing a great deal of mental effort to achieve a task (Clark, Nguyen, & Sweller, 2006). These two tasks, the Maze Task and the Chair-Building Task, are described below.

General computer self-efficacy (GCSE) was measured by the Computer User Self-Efficacy Scale (Cassidy & Eachus, 2002). This scale consists of two parts: (a) individual characteristics and (b) computer self-efficacy items. The individual characteristics section contains seven items that ask the participant about whether they have attended a computer course, own a computer, basic demographics (age and gender), experience with computers, and types of software packages they have used. The computer self-efficacy section contains 30 items, each with a five-point Likert scale ranging from strongly disagree to strongly agree. About half of the items were positively worded and half were negatively worded. The negatively worded items were reverse coded for analysis purposes. Example items include, “computers are far too complicated for me,” and, “most difficulties I encounter when using computers, I can usually deal with.” The internal consistency of the 30-item scale has been reportedly very high (*alpha* = 0.97, *N* = 184). Test-retest reliability has been reportedly high and statistically significant ($r = .086$, $N = 74$, $p < 0.0005$) (Cassidy & Eachus, 2002).

Multi-user virtual environment self-efficacy was measured using the researcher created MUVE Self-Efficacy (MUVE-SE) instrument. This study represents an initial pilot test in the development of the MUVE-SE. The MUVE-SE contains 18 items designed to measure a participant’s efficacy to perform tasks typically required in the Second Life MUVE when used as a learning environment. Example items include, “I believe I can teleport to other locations,” and “I believe I can move objects I create in Second Life.” Initial responses to the 18 items were on a “Yes” or “No” binary scale. Participants responding “Yes” to any item were then asked to rate their level of confidence in their ability on a ten-point scale with a one representing low confidence and a ten representing high confidence (Bandura, 1997).

The “Maze Task” was devised by the researcher as a means of measuring the basic skills required to utilize a MUVE. A maze was designed that required participants to navigate their avatars through a door, turn and walk in various directions, and fly over and land on the other side of a wall. Additionally, a sign in the maze required them to take a picture of their avatar’s face and email it to the researcher to measure mastery of changing the camera point of view. Next, participants were to answer a question by writing their answer on a note card of their own creation. They then had to deposit the note card into a drop box, and obtain a different note card from a dispenser. Finally, participants had to click on a teleporter to transport to another location within the virtual world. Each of these tasks are commonly performed when using the Second Life MUVE for educational purposes.

Finally, the “Chair-Building Task” was devised by the researcher as a means of measuring participants’ basic building skills. Participants were given a model of a chair and were asked to replicate the design. Instructions indicated that the chair did not have to have identical measurements or dimensions, just an identical design. The chair was to include four legs, a square seat, and a back. The back was to consist of two vertical braces connected by two horizontal slats.

Participants also attended two classes, synchronously, entirely in the virtual environment and, outside of regularly scheduled class time, met with their teammates to discuss their assignments and work on their building project. During class sessions the instructor presented information verbally and using a PowerPoint slide show. The instructor also showed students a video during class in the MUVE. Participants also conducted small-group discussions in the

MUVE. Finally, participants uploaded a Power-Point presentation into the MUVE and loaded it into a perpetual slide-show viewer outside of their building. This was done to explain their design and their rationale for their decisions to their classmates and their instructor.

Design and Procedure

This pilot study used a one-group pretest-posttest design. Participants were introduced to the MUVE in class and were given a guided three-hour practice session in which they were introduced to all of the basic skills measured in the maze and chair tasks. During this time, participants built a chair identical to the one in the Chair-Building Task for practice.

After this introductory session, measurements for GCSE, MUVE-SE, the Maze Task and the Chair-Building Task were taken. The start and end times of the tasks were recorded as were the participants' perceived mental effort (PME) associated with both of these tasks.

Learning efficiency was measured using the following formula (Clark, Nguyen, & Sweller, 2006):

$$\frac{\text{Average Performance in Z Scores} - \text{Average Difficulty Rating in Z Scores}}{\sqrt{2}}$$

To conduct the learning efficiency calculation, performance was measured on the maze and chair tasks based on the amount of time it took participants to complete each task. The difficulty of each task was measured using the perceived mental effort scale (Clark, Nguyen, & Sweller, 2006). Learning efficiency was calculated for each of two performance tasks: (a) navigation through a maze in the MUVE, and (b) building a simple chair within the MUVE. Learning efficiency was calculated after participants had spent three hours in the MUVE (pre-test) and then after spending six weeks using the MUVE (posttest).

During the six weeks in which participants used the MUVE, they worked in teams of three or four to build a simulated classroom. Each member of the team designed and helped construct the simulated learning environment, contributing approximately 100 objects to its construction. These objects are known in Second Life as "prims" which is short for "primitive" objects. A "prim" starts out as a basic three-dimensional shape, such as a cube, and is then transformed by participants into a variety of shapes that, when combined together, form complete objects. For example, a bookshelf with three shelves, two sides, and a base may consist of six "prims."

At the end of the study, participants again completed the GCSE, MUVE-SE, Chair-Building Task, and Maze Task. Additionally a questionnaire was administered that was designed by the researcher to elicit their impressions of the MUVE as a learning environment and as a place in which to build virtual simulations.

Finally, a learning curve was calculated for the percentage of improvement in achievement (completion time), perceived mental effort, and learning efficiency over the six-week period of this investigation.

Analysis

The quantitative data were analyzed using SPSS and Microsoft Excel. Open-ended questionnaire responses were transcribed, coded, entered into a qualitative research program, HyperResearch, and analyzed using the constant comparative method to shed light on the quantitative findings.

A combination of procedures was used to answer research question one, "How efficient are new participants at creating and working in the MUVE and how does efficiency change over time?" First, pre-test and posttest learning efficiency was calculated for both the Chair-Building hair and Maze Tasks. Paired sample t-tests were conducted to determine whether pre and posttest scores on both tasks (completion time, perceived mental effort, and learning efficiency) were statistically different from one another. Effect sizes were calculated using Cohen’s *d*. Additionally, the learning curve was calculated based on the percentage of improvement participants achieved over the six-week period of the study.

Research question two contained three parts. Pearson correlation analyses and paired sample t-tests were performed to answer part one of the question, “what is the relationship between participant general computer self-efficacy and MUVE self-efficacy before and after using the environment?” Paired sample t-tests were used to answer part two of the question, “does using the MUVE result in increased GCSE and MUVE-SE?” Finally, linear regression analyses were used to answer the third part of question two, “Does GCSE or MUVE-SE predict learning efficiency?” Effect sizes were calculated using Cohen’s *d*.

An analysis of open-ended survey questions for trends and commonalities was conducted to answer research question three, “what were the participants’ impressions of their user experience within the Second Life MUVE?”

Results

Learning Efficiency

Learning efficiency was measured for two tasks – the Maze Task and the Chair-Building Task – using completion time as a measure of performance and perceived mental effort as a measure of cognitive load. These measurements were taken for the participants after three hours of experience (pre-test) in the MUVE and again after six weeks of experience (posttest). Means and standard deviations are provided in Table 2.

Table 2

Completion times (performance) and perceived mental effort (PME) scores

	Maze Task		Chair-Building Task	
Measurement	*M*	*SD*	*M*	*SD*
Pre-test				
Performance	12.88	6.92	16.55	5.66
PME	5.41	2.06	7.41	1.94
Posttest				
Performance	6.65	2.77	11.79	4.79
PME	4.35	1.91	4.94	2.01

Note. n = 17 for the pre-test and posttest.

The results of two-tailed, paired-sample t-tests comparing the Maze Task pre-test performance and PME scores with the Maze Task posttest performance and PME scores showed that participants' performance times were significantly better on the posttest ($t_{(16)} = 4.33, p < 0.01$) and had a large effect size ($d = 1.14$). However, participants' perceived mental effort was not significantly lower on the posttest ($t_{(16)} = 1.62, p = 0.12$).

Two-tailed, paired-sample t-tests comparing the Chair Task pre-test performance and PME scores with Chair Task posttest performance and PME scores indicated that participant's performance times were significantly better on the posttest ($t_{(16)} = 3.01, p < 0.01$) with a large effect size ($d = .88$). Results also indicated that participants' perceived mental effort was significantly lower on the posttest ($t_{(16)} = 4.02, p < 0.01$) with a large effect size ($d = 1.21$).

Means and standard deviations for learning efficiency are presented in Table 3. A two-tailed, paired-sample t-test indicated that participants were significantly more efficient in the

Maze Task ($t_{(16)} = 3.14$, $p < 0.01$) after six weeks in the MUVE, with a large effect size ($d = -.90$). Similarly, participants were significantly more efficient in the Chair-Building Task ($t_{(16)} = 4.08$, $p < 0.001$) after six weeks, with a large effect size ($d = -1.20$). Contrary to typical effect size interpretation, due to the use of standardized scores and the way learning efficiency is calculated, the negative signs on these two effect sizes indicates that a large positive improvement was realized.

Table 3

Pre and posttest learning efficiency scores

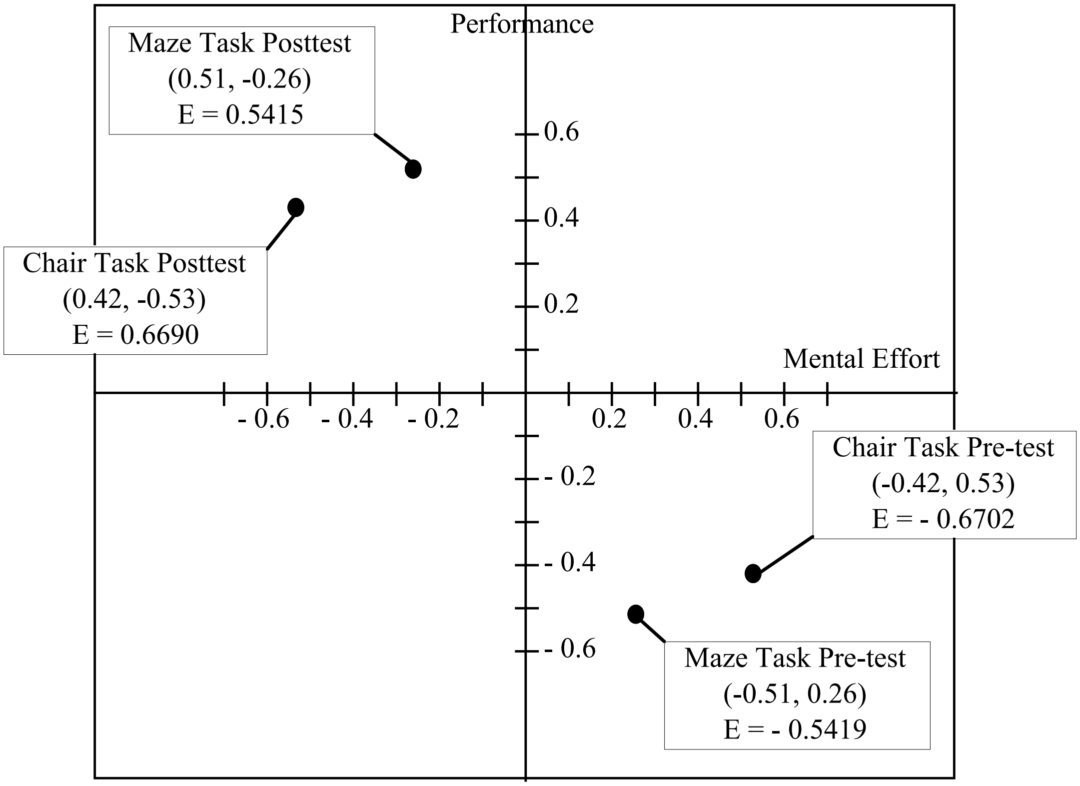
	Maze Task		Chair-Building Task	
	M	*SD*	*M*	*SD*
Pre-test	-0.5419	1.4222	-0.6702	1.0468
Posttest	0.5415	0.9379	0.6690	1.1728

Note. n = 17 for the pre-test and posttest.

Mean learning efficiency pre-test and posttest scores are plotted in Figure 2 for the Maze Task and Chair-Building Task. Both learning efficiency scores indicated low efficiency on the pre-test and higher achievement based on the posttest scores.

Figure 2

Learning Efficiency Plot



Learning curve values, in percentages, are presented in Table 4. These values show the percentage of improvement for the Maze and Chair-Building tasks as measured by achievement (completion time), perceived mental effort, and learning efficiency over the six-week duration of the investigation and as weekly averages.

Table 4

Mean learning curve improvement and weekly improvement percentages

	Maze Task		Chair-Building Task	
	Total Improvement	Weekly Improvement	Total Improvement	Weekly Improvement
Achievement	48 %	8 %	28 %	4.6 %
PME	19 %	3 %	33 %	5.5 %

Learning efficiency	1.6 %	.2 %	1.9 %	.3 %

Self-Efficacy

The 30-item instrument used to measure General Computer Self-Efficacy (GCSE) and the 18-item Multi-User Virtual Environment Self-Efficacy (MUVE-SE) instrument were tested for reliability using Chronbach's alpha. The GCSE instrument showed a high degree of reliability when it was administered at the beginning of the investigation (alpha = .840, *n* = 17) and again when it was administered at the end of the study (alpha = .860, *n* = 17). The MUVE-SE instrument showed a high level of reliability when it was administered at the start of the investigation (alpha = .936, *n* = 17) and then again when it was administered at the end of the investigation (alpha = .927, *n* = 17).

Means and standard deviations for the GCSE and MUVE-SE measures are provided in Table 5.

Table 5

GCSE and MUVE-SE means and standard deviations

	GCSE		MUVE-SE	
	M	*SD*	*M*	*SD*
Pre-test	156.18	12.52	101.29	38.95
Posttest	162.47	11.52	157.76	19.03

n = 17

Pearson correlations were performed to determine the relationship between GCSE and MUVE-SE pre and posttest scores. Two-tailed tests for significance were calculated within SPSS for each correlation. Results indicate that pre-test GCSE scores are not significantly correlated with pre-test MUVE-SE scores ($r = .097$, $p = .712$, $n = 17$). Additionally, posttest GCSE scores are not significantly correlated with posttest MUVE-SE scores ($r = .301$, $p = .240$, $n = 17$).

Linear regression analyses indicate that pre-test GCSE scores do not predict pre-test learning efficiency on either the Maze Task ($F_{(1, 16)} = 0.159$, p = .696), or the Chair-Building Task ($F_{(1, 16)} = 1.749$, p = .206). Similarly, pre-test MUVE-SE scores do not predict pre-test learning efficiency on the Maze Task ($F_{(1, 16)} = 0.811$, p = .382) or the Chair-Building Task ($F_{(1, 16)} = 0.020$, p = .890). Neither the GCSE or the MUVE-SE demonstrated the ability to predict learning efficiency when administered as a pre-test.

Another set of linear regression analyses were performed to determine whether the GCSE instrument or the MUVE-SE instrument could predict learning efficiency when administered after participants had spent six-weeks (posttest) using the MUVE. Results indicated that posttest GCSE scores did not predict posttest learning efficiency on the Maze Task ($F_{(1, 16)} = 3.235$, p = .092). However, posttest GCSE scores did predict learning efficiency for the posttest Chair-Building Task ($F_{(1, 16)} = 4.726$, p = .046) with an R^2 of .24. This indicates that a posttest administration of the GCSE instrument predicted 24 percent of the variance in posttest learning efficiency, a moderate amount. The same analyses for the effects of posttest MUVE-SE on posttest Maze Task learning efficiency were statistically significant ($F_{(1, 16)} = 19.802$, p < .001) with a strong R^2 of .569. The posttest MUVE-SE predicted over 56 percent of the variance in posttest Maze Task learning efficiency. The linear regression analysis for posttest MUVE-SE and Chair-Building Task learning efficiency was also statistically significant ($F_{(1, 16)} = 6.663$, p = .021) and had an R^2 of .308. This indicates that the posttest MUVE explained 30 percent of the variance in posttest Chair-Building Task learning efficiency. The MUVE-SE, when administered after participants spent six weeks in the multi-user virtual environment, appeared to explain a moderate amount of the variance in participants' learning efficiency on both the Maze and Chair -BuidlingTasks.

Paired-sample t-tests show that participants' improvement in general computer self-efficacy was significant ($t_{(16)} = 2.25$, $p < 0.05$) and had a moderate effect size ($d = -0.52$). Participants' improvement in self-efficacy for using a multi-user virtual environment was also statistically significant ($t_{(16)} = 7.16$, $p < 0.001$) and had a large effect size ($d = -1.84$). It does appear that using the Second Life MUVE for a period of six weeks does result in increased general computer self-efficacy as well as increased self-efficacy related to using the MUVE.

Participant Impressions

The 17 participants were asked to fill out a survey containing 14 open-ended items designed to elicit the details of their experiences in the MUVE. Any names mentioned here are pseudonyms. The first item asked about the presence of any technical issues that may have hindered or prevented use of the environment. No one reported difficulty downloading and installing the software on their home computers. The vast majority of participants reported no difficulty creating their account to use the MUVE or logging-in the first time. However, designing their personal avatar did take some time as participants spent up to an hour customizing their avatar's appearance. Two commented that they enjoyed the experience of customizing the appearance and dress of their avatars. Six reported instances when the Second Life client software froze and eight reported the monitor's display orientation spontaneously flipping sideways because of video cards that were incompatible with the client software. Finally, while not a technical problem with the MUVE's software, several participants, who are also in-service teachers, reported not being able to access the environment from their school computers due to Internet filters put in place by their school system.

As far as enjoyment of the virtual world experience, a mixture of impressions were reported. The majority were positive (*n* = 6) or neutral (*n* = 7), while others were negative (*n* = 4). Examples of positive comments about enjoying the MUVE included some remarking on their high level of engagement when using the MUVE, enjoyment of holding class meetings in the MUVE, and the creativity enabled by the tools in the MUVE. Mya said, "Very engaging, I would love to do it more." Kelly's remark was one of the most positive about the experience of interacting and creating in the MUVE:

> I <u>loved</u> this [participant underlined the word loved]. I loved the fact that I could chat with my professor and classmates very easily. I enjoyed being creative in building, purchasing my materials, and then uploading pictures of my choice.

Finally, Michelle said, "The interacting and building were fun. I really enjoy class online in the virtual environments."

Some of the comments about liking the user experience were more neutral. One participant reported that she, "Enjoyed it immensely, however it was an acquired enjoyment." Angelina explained, "It was very difficult at first. However, after creating more and more it became easy. I liked being able to chat with my classmates." Emma reported a bit of frustration that detracted from her user experience when participating in a synchronous class in the MUVE when she replied, "It was interesting. It was very frustrating when some people were able to talk and when we were unable to hear those who were talking." Apparently side conversations and comments from multiple participants frustrated her.

Finally, the negative comments related to the user experience were concerning such things as feeling ill when using the environment, unspecified frustration, and technical difficulties. Two participants reported that using the MUVE caused them to experience actual motion sickness. One commented that she experienced so much downtime due to technical issues on her home computer that she was very frustrated. However, she did not explain what issues caused the downtime. When asked a follow-up question about any ill effects they may have experienced, an additional four participants indicated they had spent so much time looking into the monitor when using the MUVE that they experienced eye strain, blurred vision, or headaches.

In regards to the impact of designing and building a learning environment in the MUVE, participants overwhelmingly reported that the three-dimensional environment required an enhanced level of planning and afforded them the opportunity to really show the application of research to their design. Perhaps the most representative participant statement on this topic was:

> Having to actually create the space brought the thought process to a deeper level than just merely stating what would be in the environment.

The last item on which the participants responded concerned their impressions of their own engagement when in the MUVE. Of the 17 participants, 13 subjectively reported having a high level of engagement while in the MUVE. Beyonce reported:

> Highly engaged. My team members were building at the same time so we communicated on how to build and collaborated on what to build where. Also, I had to keep in mind what our goals were in terms of key aspects for our environment so my building had a purpose.

Kelly said she was, “So tuned in that I barely noticed anything going on around me. I had to make a special effort to not get so in the zone that I lose [sic] track. I loved this [participant underlined the word loved] & very much enjoyed building." When in the environment itself, especially when working with their group, these 13 participants clearly reported a high level of engagement when in the MUVE.

The remaining four participants reported low engagement for various reasons. Two felt intimidated by the intricacy of the environment and the complexity of the user-interface. One said she did not like it because she felt like she was playing a video game and she dislikes video games. Finally, another stated that she did not like Second Life [no specific reason given] and felt that it had no place in her life. Therefore, there appeared to be some resistance to the idea of using this MUVE to complete a class assignment.

Summary and Conclusions

Summary

Participants improved over the six-week investigation on the Maze Task in terms of reduced perceived mental effort (mean improvement = 1.06) and performance (mean improvement = 6.23 minutes) and the results were statistically significant for performance ($t_{(16)} = 4.33$, $p < 0.01$) with a large effect size ($d = 1.14$). However, perceived mental effort was not significantly lower on the posttest for the Maze Task indicating that after six weeks participants still perceived the environment to be a challenging one with which to interface. This is consistent with participant responses to the open-ended questionnaire indicating a level of frustration with performing tasks such as changing one's camera angle (angle of view from an avatar's perspective) and landing one's avatar where one intends it to land. It is also consistent with the literature (Sanchez, 2009).

In addition to improvements in performance and perceived mental effort, participants improved their learning efficiency during the investigation. T-tests indicated that participants exhibited statistically significantly improved learning efficiency on the Maze Task ($t_{(16)} = 3.14$, $p < 0.01$) and the Chair-Building Task ($t_{(16)} = 4.08$, $p < 0.001$), with large effect sizes of .90 and -1.20 respectively. This finding indicates that in-service teachers can become efficient at using and creating in the environment over a six-week time period. Participant self-reported engagement in the environment likely played a role in increasing learning efficiency, but since engagement was not empirically measured, this study cannot shed any light on this relationship.

The learning curve percentages are useful to know. They seem to provide useful data as to how rapidly a user will learn to build and interact in the MUVE. Participants experienced the greatest improvement in terms of the time it took them to complete tasks (48 percent for the Maze Task, 28 percent for the Chair-Building Task), but also experienced double-digit reductions in perceived mental effort (Maze Task = 19 percent, Chair-Building Task = 33 percent). Learning efficiency improved at the lowest rate of all (Maze Task = 1.6 percent, Chair-Building Task = 1.9 percent). Again, this finding is consistent with participant reports of frustration related to using and building within the MUVE. These reports and findings are also consistent with the literature (Sanchez, 2009). Perhaps additional guided practice and tutorial videos would have helped to improve the learning curve related to building and operating within the MUVE.

Findings on self-efficacy indicate that using a MUVE will likely contribute to the improvement of general computer self-efficacy and MUVE self-efficacy. Participants improved their GCSE from the pre-test to the posttest by an average of 6.29 points and they improved their MUVE-SE by an average of 56.47 points over the six-week duration of this investigation. The improvements in GCSE ($t_{(16)} = 2.25$, $p < 0.05$) and MUVE-SE ($t_{(16)} = 7.16$, $p < 0.001$) were statistically significant, and had moderate to large effect sizes as measured by Cohen's d of -0.52

and -1.84, respectively. The high reliability of both the GCSE (*alpha* > .80) and the MUVE-SE (*alpha* > .90) seem to indicate that both may be useful instruments for researchers. Although the MUVE-SE is still under development, its performance shows promise.

Neither measure of self-efficacy predicted pre-test learning efficiency. The Posttest GCSE only predicted posttest learning efficiency on the Chair Task ($F_{(1, 16)} = 4.726$, $p = .046$), and it did so with a moderate R^2 of .24. Interestingly, the posttest MUVE-SE demonstrated predictive ability for the posttest Maze Task ($F_{(1, 16)} = 19.802$, $p < .001$, $R^2 = .569$) and Chair Task ($F_{(1, 16)} = 6.663$, $p = .021$, $R^2 = .308$). General computer self-efficacy does not appear to be specific enough to predict learning efficiency in a MUVE. The MUVE-SE instrument has some predictive ability, however, this is only a pilot test of the instrument which needs to be refined through item analysis and further validation measures before it is ready for dissemination.

Finally, participant impressions of using the Second Life MUVE for the class were generally positive. The majority of participants reported being highly engaged when using the environment and they reported enjoying interacting in the environment. Responses also indicated that their level of planning and thinking related to the building of a simulated educational structure was enhanced by using the environment. Building in the environment took some time for participants to get used to, however. This was supported by the slow rate of improvement in learning efficiency scores. They reported frustration at learning to build due to the three-dimensional nature of the environment and the complexity of manipulating objects with the user interface. Some participants reported technical difficulties related to not having a recommended video card, a robust enough microprocessor, enough RAM, and/or a fast enough Internet connection. A few even reported feelings of nausea and motion sickness, eye strain, blurred vision, and headaches due to prolonged use of the MUVE. The three-dimensional nature of the environment, coupled with the feeling of immersion felt by participants, may result in ill effects for some users. It may be advisable to recommend that users be aware of these potential effects and moderate the amount of time they spend in the environment during an individual session. All of these findings are consistent with earlier findings in the literature (Sanchez, 2009).

Implications

The Second Life MUVE appears to be a promising environment that fosters high levels of engagement of adult learners, supports synchronous online class activities as a distance-education delivery platform, and provides a virtual environment in which educators and teacher educators may build simulated learning environments.

Using this MUVE is likely to improve the general computer self-efficacy of teachers and possibly other adults. However, general computer self-efficacy does not appear to be directly related to using a MUVE. That is why the MUVE-SE instrument is promising; its further development offers the potential of measuring self-efficacy related to the use of a MUVE that may have the ability to predict learning efficiency when using such an environment.

For teacher educators the main implication seems to be that, with scaffolding and time, teachers can be taught to create simple simulated environments in this MUVE. Results from this study support the conclusion that creation in a MUVE can support adult engagement and higher levels of thinking. The creation of advanced simulations, requiring the scripting of object actions, was beyond the scope of this study. However, it can be implied from this study that the creation of advanced simulations will likely require significant training before educators are ready. Second Life, and other MUVEs, may offer educational programming, instructional design, and multimedia classes an environment in which students can create meaningful and engaging simulations.

Additionally, Second Life has the potential to function as an environment in which to hold synchronous classes. Participants have to be taught how to be virtual students with their avatars because the environment presents difficulties related to managing who gets to talk and ask questions. Additionally, as participants reported, there is sometimes a lag between when an event is initiated, such as showing a PowerPoint slide, and when that event actually takes place in the MUVE. This is called lag-time, or just lag. Lag presents a challenge that Second Life has to overcome before it can be used to hold large numbers of avatars in one place representing whole classes of students. Those wishing to teach using a MUVE as a synchronous class space will have to learn to overcome these same challenges as well.

For researchers this study offers a potential contribution in the form of a method with which to study and compare computer user interfaces. Learning efficiency and learning curve seem well suited to this task.

Limitations

This study had several limitations, which if addressed, may provide direction for future research. The sample used for this study was not very large nor was it randomly drawn from a large population of K-12 educators. Therefore, it is not representative and the results of this investigation may not be generalized. The researcher intends to repeat this experiment with larger numbers of participants in the future. In addition, the study did not utilize a comparison group. It is intended that future iterations of this study will be accomplished using some form of control-group design. Finally, the MUVE self-efficacy instrument, created by the researcher, will require more data to be validated.

Suggestions for Further Research

Multi-user virtual environment self-efficacy is a construct that appears to have some usefulness because of its potential predictive abilities. A comparison of MUVE programs should be conducted to identify common features and user interface mechanisms. Based on these similarities, the MUVE-SE should be revised and validated with a much larger sample.

Learning efficiency appears to be a useful framework on which to design comparisons of user-interfaces and programs. This methodology should be validated with a much larger sample of participants. To be generalizeable beyond Second Life, the Maze Task and Chair-Building Task should be examined to determine the MUVE programs to which these tasks apply. Learning efficiency could then be calculated for each MUVE and those results could be applied to choose the environment best suited to the intended user population.

Due to the complexity of building in a MUVE, methods of increasing germane cognitive load should be implemented and studied to determine their ability to increase learning efficiency. Such methods include structuring building lessons that start with a combination of simple partial and whole tasks before progressing to more complex partial and whole task activities. Next, start lessons on building with worked examples, continuing to completion tasks, and ending with whole tasks. Combining this sequence, from simple building to complex building tasks, should increase germane cognitive load and facilitate learning efficiency (van Merrienboer, Kester, & Paas, 2006). It would be useful to know which combinations of methods of teaching simulation construction produce the greatest learning efficiency.

Finally, empirical measures of engagement should be developed and validated for the purpose of measuring the engagement of adults and children when they are using a MUVE and other programs. It would be useful to be able to compare engagement levels for users of a variety of simulations in a MUVE.

References

Bandura, Albert. (1997). *Self-efficacy: The exercise of control*. New York: W.H. Freeman and Company.

Clark, Ruth., Nguyen, Frank., & Sweller, John. (2006). *Efficiency in learning: Evidence-based guidelines to manage cognitive load*. San Francisco: John Wiley & Sons, Inc.

Downey, J. (2007). Measuring general computer self-efficacy: Surprising comparison of three instruments in predicting performance, attitudes, and usage. *Systems Sciences*. Vol. 8, 210

Jackson, Randolph L., and Fagan, Eileen. (2000). Collaboration and learning within immersive virtual reality. *Proceedings of the third international conference on collaborative virtual environments*. San Francisco, CA.

Sanchez, J., Lumbreras, M., & Silva, J. (2001). Virtual reality and learning: Trends and issues. University of Chile. Found online on May 27 at http://www.c5.cl/hh/papers/virtualreality.pdf

Sanchez, J. (2009). Barriers to student learning in Second Life. *Library Technology Reports*, February/March, 29-34.

SimTeach. (2008). *Second Life in Education Wiki: Educational Institutions in Second Life*. Retrieved from the Internet on September 5, 2008 from http://simteach.com/wiki/index.php?title=Institutions_and_Organizations_in_SL#UNIVERSITIES.2C_COLLEGES_.26_SCHOOLS

Teplitz, C. J. (1991). *The learning curve deskbook: A reference guide to theory, calculations, and applications*. New York: Quorum Books van Merrienboer, Joroen, J. G., Kester, Leisbeth, & Paas, Fred. (2006). Teaching complex rather than simple tasks: Balancing intrinsic and germane load to enhance transfer of learning. *Applied Cognitive Psychology*. 20. 343-352.

Yelle, L. E. (1979). The learning curve: Historical review and comprehensive survey. *Decision Sciences*. 10(2). 302-328.