Evaluation of Learning Efficiency and Efficacy in a Multi-User Virtual Environment

Doug Hearrington
Kennesaw State University

Abstract

This study evaluated the multi-user virtual environment (MUVE) known as Second Life, integrated with Moodle and SLOODLE technologies, as an exploratory course delivery platform and for its ability to enable teachers to meet elements of NETS•T. Graduate student participants (N = 17) interacted, constructed simulated schools, and attended classes in the MUVE. The researcher used pre- and posttest measures of self-efficacy and learning efficiency to understand the effects of the MUVE on participants and their rate of learning to make educational use of the environment. Findings imply that the technologies have potential as a distance-learning platform and as a tool to meet elements of NETS•T. Preparing teachers to use the MUVE for these purposes is likely to require a significant amount of scaffolding. (Keywords: MUVE, Second Life, distance education, self-efficacy)

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

Moodle (Modular, Object-Oriented, Dynamic Learning Environment) is a free and open source learning management system similar to Blackboard or WebCT. The researcher tested Moodle in this study because it interfaces with Second Life via the plugin known as SLOODLE (Simulation-Linked, Object-Oriented, Dynamic Learning Environment). Using these two technologies with Second Life allows users to blog, take quizzes, submit three-dimensional objects to drop-boxes, and much more. The ability to integrate the Second Life MUVE with a learning management system (Moodle) via SLOODLE provided an additional reason to test the functionality of these technologies to serve as a distance education platform.

Another compelling reason to use a MUVE in teacher education programs is that the technology may facilitate the implementation of specific program standards. ISTE has, in conjunction with a wide variety of professional education organizations, established the National Educational Technology Standards for Teachers (NETS•T). Teacher-education institutions accredited by the National Council for Accreditation of Teacher Education (NCATE) use the NETS•T as the standards by which to measure teacher readiness to use information and communication technologies (ICTs). A MUVE may help a teacher implement the first two NETS•T standards and several of the performance indicators for those standards. The first NETS•T standard is: "Facilitate and inspire student learning and creativity." This standard contains four performance indicators, each of which may be accomplished using a MUVE:

(a) Promote, support, and model creative, innovative thinking and inventiveness
(b) Engage students in exploring real-world issues and solving authentic problems using digital tools and resources
(c) Promote student reflection using collaboration tools to reveal and clarify students' conceptual understanding and thinking, planning, and creative processes
(d) Model collaborative knowledge construction by engaging in learning with students, colleagues, and others in face-to-face and virtual environments (ISTE, 2008)

The second NETS•T standard is: "Design and develop digital age learning experiences and assessments." Two of this standard's performance indicators clearly lend themselves to using a MUVE:

(a) Design or adapt relevant learning experiences that incorporate digital tools and resources to promote student learning and creativity
(b) Develop technology-enriched learning environments that enable all students to pursue their individual curiosities and become active participants in setting their own educational goals, managing their own learning, and assessing their own progress (ISTE, 2008)

A MUVE is a form of virtual reality (VR), a computer-based technology that provides the visual, aural, and tactile stimuli of a virtual world generated in real time (Sanchez, Lumbreras, & Silva, 2001). A MUVE is an example of interactive multimedia in which experiences are made possible by dynamic elements under the user's control (Rieber, 2005). For example, rather than reading about the surface of the Moon, one can replicate the environment in a MUVE and students can control their own exploration of a virtual Moon, thus providing a form of experiential learning. Additionally, users may interact with objects and events in simulations in MUVEs and similar three-dimensional environments (Cobb & Fraser, 2005).

Users of a MUVE experience three "presence layers" in a 360-degree, three-dimensional environment, with the effect of combining physical and virtual realism 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). The visual and physical proximity of avatars to one another creates the physical presence. Spatially enhanced voice that allows one to sense the direction from which another voice originates, as well as communication via synchronous text chat and asynchronous text communications, such as e-mail and group notices, create the communication presence. Finally, various tools that allow one to know when a friendly avatar is "in-world" or offline create a status presence (Warburton, 2009). Presence helps to establish a first-person experience when one is in a MUVE. This experience has been said to help develop direct, subjective, and personal knowledge (Sanchez, Lumbrera, & Silva, 2001).

Affordances and Barriers to MUVE Use

MUVES may be said to have structural and functional affordances. Structural affordances consist of (a) a collection of objects that model the mathematical/physical properties of the domain, (b) links to numerous representations of the underlying model, (c) opportunities to use the objects in complex ways, and (d) challenges or activities for the student to solve or explore (Edwards, 1995). Functional affordances include the interaction between the user, the software, and the setting in which the environment is used.

Barriers to learning in the Second Live MUVE include a "high learning curve," which is reflected in its complicated user interface, high difficulty level of building new objects for novice users, perceptions that working in the Second Life MUVE was too time-consuming, and technical difficulties (Sanchez, 2009). Some users have expressed disappointment with the environment because of their expectations that it should be more like a playing a game. When Second Life turned out to not be a game, these users became bored and frustrated (Sanchez, 2009). Due to the limitations of monitors, three-dimensional MUVE environments are viewed on a flat screen, creating a lack of peripheral vision. This lack of peripheral vision detracts from the representation of the environment (Cobb & Fraser, 2005). Time delays are often common when using such an environment, and depth perception is difficult for some users to learn when in a three-dimensional environment (Cobb & Fraser, 2005).

Positive aspects of using the Second Life MUVE have also been noted. Users enjoy creating and designing their avatar and 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). 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 an educator's computer self-efficacy 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.

Self-Efficacy

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 (Bandura, 1997). 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, the researcher administered a measure of general computer self-efficacy and created and administered a specific self-efficacy instrument related to using a MUVE.

Cognitive Load Theory

Because 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, which states that 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, which is the mental work imposed by an instructional activity that benefits the

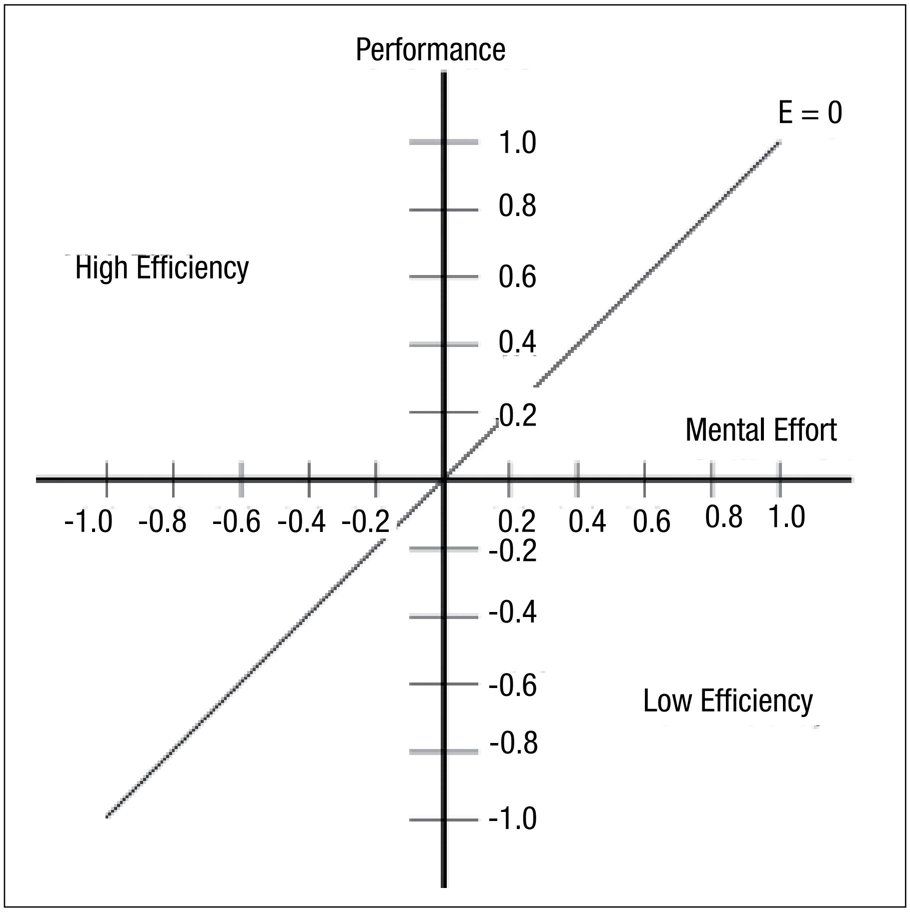

Figure 1. The Learning efficiency grid.

learning of the task; and (c) extraneous (irrelevant) load, which 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 for an example of the learning efficiency grid.

Low learning efficiency occurs when one learns something slowly with great mental effort. High learning efficiency occurs when one learns something quickly and 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, to compare different programs to one another, or to compare learning in a variety of computer or non-computer conditions.

Learning efficiency is a function of achievement and mental effort, but 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 3 minutes to play. Her second attempt took 2.6 minutes. Attempt number three took 2.37 minutes. The fourth attempt took under 2 minutes. The rate of the student's improvement, calculated every other attempt (attempt 1, 2, 4, 8, etc.), is 21%. This means that each doubling results in an improvement of 79%. Although time it takes the student to succeed 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 that require 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, the potential of the environment to be used to meet NETS-T standards, and the increasing use of MUVEs at all levels of education for building and using simulations, the researcher designed this study 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

The researcher collected data from 17 graduate students enrolled in a master's 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. None of the participants had any experience using a MUVE or a similar environment before this study. Table 1 summarizes other demographic characteristics of the participants.

Instruments

Cognitive load was measured for two defined tasks within the MUVE using an established 9-point scale ranging from 1 for very little mental effort to 9 for 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.

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

The Computer User Self-Efficacy Scale (Cassidy & Eachus, 2002) measures general computer self-efficacy (GCSE). 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 or 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 5-point Likert scale ranging from strongly disagree to strongly agree. The researcher worded approximately half of the items positively and half negatively. The researcher reverse-coded the negatively worded items 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).

The researcher created the MUVE Self-Efficacy (MUVE-SE) instrument to measure the multi-user virtual environment self-efficacy. 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 at performing tasks typically required when they use the Second Life MUVE 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 10-point scale, with a 1 representing low confidence and a 10 representing high confidence (Bandura, 1997).

The researcher devised the "Maze Task" as a means of measuring the basic skills commonly required to utilize a MUVE. The researcher designed a maze 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 e-mail it to me to measure mastery of changing the camera (user's) point of view. Next, a sign in the maze directed participants 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. Users commonly perform each of these tasks when using the Second Life MUVE for educational purposes.

Finally, the researcher devised the "Chair-Building Task" as a means of measuring participants' basic building skills. The researcher gave participants a model of a chair and asked them 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 synchronously attended two classes entirely in the virtual environment and met with their teammates outside of regularly scheduled class time to discuss their assignments and work on their building project. During class sessions, the instructor presented information verbally and using a PowerPoint slideshow. The

Table 2: Completion Times (Performance) and Perceived Mental Effort (PME) scores

	Maze Task		Chair-Building Task	
Measurement	M	SD	M	SD
Pretest				
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 pretest and posttest.

Table 3: Pre- and Posttest Learning Efficiency Scores

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

Note. n = 17 for the pretest and posttest.

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 %

instructor also showed students a video during class in the MUVE. Participants conducted small-group discussions in the environment and submitted a simple assignment on a note card into a drop box. To evaluate SLOODLE, participants used the quiz tool and the choice tool (which allows users to vote and see the three-dimensional results of their vote). Finally, participants uploaded a PowerPoint presentation into the MUVE and loaded it into a perpetual slideshow viewer outside of their building. They did this to explain their design and their rationale for their decisions to their classmates and their instructor.

Design and Procedure

This pilot study used a pre-experimental, one-group pretest-posttest design (Campbell & Stanley, 1963). Participants were introduced to the MUVE in class and received a guided 3-hour practice session, during 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, the researcher took measurements for GCSE, MUVE-SE, the Maze Task, and the Chair-Building Task. The researcher recorded the start and end times of the tasks as well as the participants' PME associated with both of these tasks.

The researcher measured learning efficiency 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, the researcher measured performance on the maze and chair tasks based on the amount of time it took participants to complete each task. The researcher measured the difficulty of each task using the PME scale (Clark, Nguyen, & Sweller, 2006). The researcher calculated learning efficiency for each of two performance tasks: (a) navigation through a maze in the MUVE and (b) building a simple chair within the MUVE. The researcher calculated learning efficiency after participants had spent 3 hours in the MUVE (pretest) and then after spending 6 weeks using the MUVE (posttest).

During the 6 weeks in which participants used the MUVE, they worked in teams of three or four to build a simulated digital age classroom that would promote high levels of technology integration based on assigned readings and each group's literature review of technology integration and 21st century or digital age classroom design topics. The instructional strategy used was a combination of problem-based learning and cooperative learning groups, in which each team had the same digital age classroom building assignment. In addition to building a simulated learning environment (SLE), teams were required to itemize the elements they believed were important to the design of such a learning environment and to explain how they addressed each element in their design as part of a presentation to their classmates. Participants uploaded PowerPoint presentations into Second Life and delivered their presentation inside of the MUVE. Each member of the team designed and helped construct the SLE, 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." Each SLE contained 300–500 prims.

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

Finally, the researcher calculated a learning curve for the percentage of improvement in achievement (completion time), PME, and learning efficiency over the 6-week period of this investigation.

Analysis

The researcher analyzed the quantitative data using SPSS and Microsoft Excel.

The researcher transcribed the open-ended questionnaire responses, coded them, entered them into a qualitative research program (HyperResearch), and analyzed them using the constant comparative method to shed light on the quantitative findings.

The researcher used a combination of procedures to answer research question 1, "How efficient are new participants at creating and working in the MUVE, and how does efficiency change over time?" First, the researcher calculated pretest and posttest learning efficiency for both the Chair-Building and Maze Tasks. The researcher conducted paired sample t-tests to determine whether pre- and posttest scores on both tasks (completion time, PME, and learning efficiency) were statistically different from one another. The researcher calculated effect sizes using Cohen's d. Additionally, the researcher calculated the learning curve based on the percentage of improvement participants achieved over the 6-week period of the study.

Research question 2 contained three parts. The researcher performed Pearson correlation analyses and paired sample t-tests to answer part 1 of the question, "What is the relationship between participant general computer self-efficacy and MUVE self-efficacy before and after using the environment?" The researcher used paired sample t-tests were used to answer part 2 of the question, "Does using the MUVE result in increased GCSE and MUVE-SE?" Finally, the researcher used linear regression analyses to answer the third part of question 2, "Does GCSE or MUVE-SE predict learning efficiency?" The researcher calculated effect sizes using Cohen's *d*.

The researcher conducted an analysis of open-ended survey questions for trends and commonalities to answer research question 3, "What were the participants' impressions of their user experience within the Second Life MUVE?"

Results

Moodle and SLOODLE

The integration of Moodle, SLOODLE, and Second Life worked quite well. Following the printed and video set-up instructions was easy, and in about 2.5 hours, the three programs were "talking" together. Once set up was complete, participants were able to use SLOODLE's quiz tool, voting tool, and toolbar when attending a virtual class in Second Life to facilitate asking and answering questions by raising their avatars' virtual hands to get the instructor's attention. The instructor gave one quiz and took one vote during the project, and the technologies worked together perfectly to transmit each participant's score and vote from Second Life, through SLOODLE, and into Moodle for final recording. The combined functionalities of SLOODLE and Moodle seem to offer educators a viable, although not yet fully developed, means of using Second Life as a distance education platform in combination with the learning management system capabilities of Moodle.

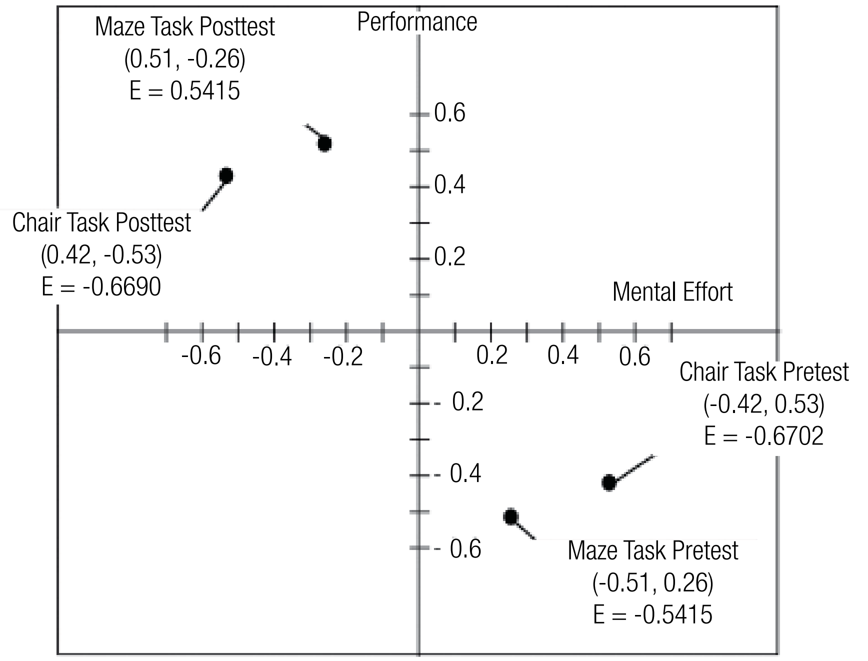

Figure 2. Learning efficiency plot.

Learning Efficiency

The researcher measured learning efficiency for two tasks—the Maze Task and the Chair-Building Task—using completion time as a measure of performance and PME as a measure of cognitive load. The researcher took these measurements for the participants after 3 hours of experience (pretest) in the MUVE and again after 6 weeks of experience (posttest). Table 2 (p. 69) provides means and standard deviations.

The results of two-tailed, paired-sample t-tests comparing the Maze Task pretest 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' PME was not significantly lower on the posttest ($t[16] = 1.62$, $p = 0.12$).

Two-tailed, paired-sample t-tests comparing the Chair Task pretest performance and PME scores with Chair Task posttest performance and PME scores indicated that participants' 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' PME was significantly lower on the posttest ($t[16] = 4.02$, $p < 0.01$) with a large effect size ($d = 1.21$).

Table 3 (p. 69) presents means and standard deviations for learning efficiency. A two-tailed, paired-sample t-test indicated that participants were

Table 5: GCSE and MUVE-SE Means and Standard Deviations

	GCSE		MUVE-SE	
	M	SD	M	SD
Pretest	156.18	12.52	101.29	38.95
Posttest	162.47	11.52	157.76	19.03

Note: $n = 17$

significantly more efficient in the Maze Task ($t[16] = 3.14$, $p < 0.01$) after 6 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 6 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 indicate that a large positive improvement was realized.

Mean learning efficiency pretest 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 pretest and higher achievement based on the posttest scores.

Table 4 (p. 69) presents learning curve values in percentages. These values show the percentage of improvement for the Maze and Chair-Building tasks, as measured by achievement (completion time), PME, and learning efficiency over the 6-week duration of the investigation and as weekly averages.

Self-Efficacy

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

Table 5 provides means and standard deviations for the GCSE and MUVE-SE measures.

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

Linear regression analyses indicate that pretest GCSE scores do not predict pretest 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, pretest MUVE-SE scores do not predict pretest 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 nor the MUVE-SE demonstrated the ability to predict learning efficiency when administered as a pretest.

The researcher performed another set of linear regression analyses to determine whether the GCSE instrument or the MUVE-SE instrument could predict learning efficiency when administered after participants had spent 6 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% 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 more than 56% 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% of the variance in posttest Chair-Building Task learning efficiency. The MUVE-SE, when administered after participants spent 6 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-Building Tasks.

Paired-sample t-tests show that participants' improvement in GCSE was significant ($t[16] = 2.25$, $p < 0.05$) and had a moderate effect size ($p = -0.52$). 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 appears that using the Second Life MUVE for a period of 6 weeks does result in increased GCSE as well as increased self-efficacy related to using the MUVE.

Participant Impressions

The researcher asked the 17 participants 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 accounts to use the MUVE or with logging in the first time. However, designing their personal avatars did take some time, as participants spent up to an hour customizing their avatars' appearances.

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, although it was not a technical problem with the MUVE's software, several participants (who are also inservice teachers) reported not being able to access the environment from their school computers due to Internet filters put in place by their school system.

Participants reported a mixture of impressions regarding their enjoyment of the virtual world experience. The majority were positive ($n = 6$) or neutral ($n = 7$), whereas some 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 that the tools in the MUVE enabled. Mya said, "Very engaging, the researcher 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 loved 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.

Regarding 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 demonstrate that their design was representative of best-practices literature. 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 groups, 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, although she gave no specific reason, 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 6-week investigation on the Maze Task in terms of reduced PME (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, PME was not significantly lower on the posttest for the Maze Task, indicating that after 6 weeks participants still perceived the environment to be challenging to interface with. This is consistent with participants' 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 PME, 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 inservice teachers can become efficient at using and creating in the environment over a 6-week time period. Participants' self-reported engagement in the environment likely played a role in increasing learning efficiency, but because 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 about 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% for the Maze Task, 28% for the Chair-Building Task), but also experienced double-digit reductions in PME (Maze Task = 19%, Chair-Building Task = 33%). Learning efficiency improved at the lowest rate of all (Maze Task = 1.6%, Chair-Building Task = 1.9%). 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 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 GCSE and MUVE self-efficacy. Participants improved their GCSE from the pretest to the posttest by an average of 6.29 points and improved their MUVE-SE by an average of 56.47 points over the 6-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 pretest learning efficiency. The Posttest GCSE only predicted posttest learning efficiency on the Chair-Building 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-Building 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, but 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 enjoying interacting in the environment. Responses also indicated that using the environment enhanced their level of planning and thinking related to the building of a simulated educational structure. 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 with 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 that participants felt, 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 in 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. It works well with both Moodle and SLOODLE.

Using this MUVE is likely to improve the GCSE 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.

Users of Second Life and other MUVEs can build individual objects using the built-in tools of the MUVE itself, prim by prim. They can also build objects outside of the MUVE in a variety of programs, such as Google Sketch-up, that create three-dimensional objects and import the objects into the virtual environment. Additionally, it is possible to build objects within the MUVE and duplicate them so that the user can give or sell the copy to another user within the MUVE. Having the skills to build or create within the MUVE will enable educators to create any simulated environment for any subject they may wish to create for teaching purposes. Having prebuilt objects (created by someone else) would greatly facilitate the creation of virtual simulations by reducing the need to build objects prim by prim, which should reduce the learning curve and may facilitate greater adoption of this type of technology.

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. This capability will enable them to use this type of tool to meet NETS-T standards with their own students should they have the opportunity to do so. 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, this study suggests 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 where 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 begins, 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 MUVE user interfaces and learning through tasks in a MUVE. Learning efficiency and learning curve seem well suited to this task.

Limitations

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 my intention to accomplish future iterations of this study using some form of control-group design. Finally, the MUVE self-efficacy instrument that the researcher created 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.

Author Note

Doug Hearrington is an assistant professor of instructional technology at Kennesaw State University (KSU), near Atlanta, Georgia. KSU is the third largest institution in the University System of Georgia. His research focuses on the design and implementation of virtual and augmented reality to advance learning, metacognitive functioning, self-efficacy, epistemological beliefs, self-determination, as well as technology leadership and integration. He has presented at several state, regional, and national conferences and is the project director of Project OWL (Online World of Learning), a multi-user virtual environment and curriculum for the advancement of third through fifth grade STEM learning and teaching. Correspondence regarding this article should be addressed to Doug Hearrington, EdD, Assistant Professor of Instructional Technology, Department of Instructional Technology, Kennesaw State University, 1000 Chastain Road NW, Box 0127, Kennesaw, Georgia 30144. E-mail: dhearrin@kennesaw.edu

References

Bailey, C. D. (1989). Forgetting and the learning curve: A laboratory study. *Management Science, 35*, 346–352.

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

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research.* Boston: Houghton Mifflin Company.

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

Cobb, S., & Fraser, D. S. (2005). Multimedia learning in virtual reality. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 525–547). New York: Cambridge University Press.

Edwards, L. D. (1995). Microworlds as representations. In A. A. diSessa, C. Hoyles, R. Noss, & L. D. Edwards (Eds.), *Computers and exploratory learning* (pp. 127–154). New York: Springer.

International Society for Technology in Education (ISTE). (2008). *National Educational Technology Standards for Teachers.* Eugene, Oregon.

Rieber, L. P. (2005). Multimedia learning in games, simulations, and microworlds. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 549–567). New York: Cambridge University Press.

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

Sanchez, J., Lumbreras, M., & Silva, J. (2001). *Virtual reality and learning: Trends and issues.* University of Chile. Retrieved May 27, 2009, at http://www.c5.cl/hh/papers/virtualreality.pdf

SimTeach. (2008). *Second Life in education wiki: Educational institutions in Second Life.* Retrieved 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, J. J. G., Kester, L., & Paas, F. (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.