THE EFFECTS OF GAMIFICATION ON SELF-EFFICACY AND PERSISTENCE IN VIRTUAL WORLD FAMILIARIZATION

A Dissertation presented to the

Faculty of the Graduate School

University of Missouri-Columbia

In Partial Fulfillment of the Requirements for the Degree

Doctor of Philosophy

---

by

PETER J. "ASIKAA" COSGROVE

Dr. Joi Moore, Dissertation Supervisor

May 2016

The undersigned, appointed by the Dean of the Graduate School,
have examined the dissertation entitled

The Effects of Gamification on Self-Efficacy
and Persistence in Virtual World Familiarization

Presented by Peter "Asikaa" Cosgrove

A candidate for the degree of Doctor of Philosophy in Information Science and Learning Technologies

And hereby certify that, in their opinion, it is worthy of acceptance.

____

Joi L. Moore, Ph.D.

____

David A. Bergin, Ph.D.

____

Sean P. Goggins, Ph.D.

____

James M. Laffey, Ph.D.

DEDICATION

This dissertation is dedicated to the memory of Derek O. Wells. Your quiet intelligence sparked my curiosity: To always wonder why. Your kindness will forever be a part of my childhood.

My wife, Toni: Without your understanding, compassion, patience and support over the past five years, I would have lost 30% of my body weight as well as a good proportion of whatever sanity I could have laid claim to in the first place. Thank you for always being ready to listen.

My parents, John and Linda, who saw the spark when sometimes I did not
and gave it room to grow. You both struck the perfect balance of belief and encouragement.
I never felt pushed, but I always knew you trusted in how far I might go.

ACKNOWLEDGEMENTS

This dissertation would not have been possible without academic support from the following people. My sincere thanks to them:

Dr. Joi Moore, my advisor, who has always opened my eyes to the possibilities, pushed me to do my best work, and find resilience in the face of frustration.

Dr. David Bergin, who is largely responsible for my interest in educational psychology and has been one of the most significant influences on my own teaching practice. I only wish I could have taken more of your classes.

Drs. Goggins and Laffey, who provided superb feedback during my proposal and have been nothing but supportive throughout the process.

Dr. Nianbo Dong and Dr. Justin Leinaweaver, who fielded questions about obscure analytical problems with a remarkable ability to understand my not-so-adroit attempts to articulate them.

Table of Contents

List of Figures

List of Tables

Abstract

Students who are inexperienced with multi-user virtual environments (MUVEs) such as *Second Life* can experience significant difficulties when first entering the virtual classroom. The unfamiliar paradigm of moving an avatar around a simulated Euclidean space can lead to disorientation, confusion, frustration and ultimately, a defeatist affect that leads to abandonment of the format. Overcoming these barriers-to-entry necessitates a foundational training course in which students can become efficacious in the basic use of virtual worlds. However, researchers have not fully established evidence-based best practices for the design of these training courses. Gamification, the process in which non-game activities are endowed with the typical attributes of a game, is a technique that has been empirically shown to improve motivation and satisfaction in such fields as education and the workplace. This study experimentally tested the effects of gamifying a virtual world familiarization course on self-efficacy, persistence, task skill and satisfaction. The results inform a theoretical model and substantive recommendations for the design of familiarization courses for virtual world learning programs.

Keywords: gamification, motivation, pre-training, self-efficacy, virtual worlds

Chapter 1: Introduction

Background

Virtual worlds and MUVEs. Virtual worlds are computer-generated simulations of three-dimensional space that present the user with an environment in which a representation of the user can be moved around. This virtual embodiment of the user, known as an avatar, is controlled through keyboard, mouse or game controller input. Some virtual worlds are used by multiple users simultaneously, connected to a central server from diverse geographic locations via the Internet. These are known as multi-user virtual environments (MUVEs). Users can see and interact with other users' avatars in real time within the virtual space. Communication is mediated primarily through audio channels: Users wear a headset with earphones and microphone and converse as if they were in close proximity to one another.

Educational institutions, especially those in higher education, have shown increasing interest in the use of MUVEs as online learning environments, providing a flexible and engaging alternative to text-based, asynchronous online courses mediated through course management systems (CMS) such as Blackboard. Classes often replicate "real world" educational formats: lectures where the students' avatars sit in a virtual classroom with the professor's avatar presenting multimedia content on a virtual screen, seminars with real-time discussion mediated by the professor, or group work with students' avatars occupying virtual breakout rooms. A key factor that distinguishes virtual world learning environments from text-based CMS is the ability for students to engage in embodied tasks: "Three-dimensional VEs that allow learners to engage simultaneously in shared tasks and/or produce joint artifacts by operating on the same objects in real time can pave the way for rich and truly collaborative experiences that foster positive interdependence within a learning group" (Dalgarno & Lee, 2010, p. 22).

While educationally focused MUVEs are not intended to be video games, some institutions have reduced development costs by co-opting video games that feature a multi-user virtual environment into ad hoc course hosting systems (Van Eck, 2006), such as *World of Warcraft*, *Everquest*, and *Minecraft*.

Gamification. The term *gamification* refers to the intentional integration of game elements into an activity, task or environment that otherwise possesses no game-like qualities. Point-based scoring systems, rankings, challenges balanced with user abilities and rewards can be considered the *motivational affordances* (Zhang, 2008) of gamification that can increase engagement and encourage users to persist (Barata, Gama, Jorge, & Goncalves, 2013; Raymer, 2011). By increasing the user's enjoyment of the task itself, motivation becomes progressively intrinsic: participation in the activity alone is a reward independent of the end result. According to self-determination theory, intrinsic motivation is more effective for high-quality learning (Ryan & Deci, 2000a) and thus is a desirable characteristic for educational programs. Hence gamification has become a growing area of interest for scholars concerned with the motivational aspects of learning and scholarship (Armstrong, 2013; Hamari, Koivisto, & Sarsa, 2014).

Self-efficacy and persistence. First proposed by Bandura (1977) and later incorporated into his social cognitive theory (1986), self-efficacy is an individual's perception of his or her capacity to perform some action, behavior or task. Self-efficacy is affected positively by successful past experiences and negatively by past failures. It is also influenced by the individual's emotional state, their observation of others' performance of the same action, and how others may have persuaded them about their ability to perform the action (Schunk & Pajares, 2005). This mechanism occurs within a broader system of self-regulation in which an individual's beliefs about their own character, abilities, shortcomings and others' perceptions of them have profound effects on how they control their own behavior and interact with their environment (Bandura, 1986). Critically therefore, self-efficacy does not simply determine one's opinions about one's abilities, but also how one behaves and makes choices. In general, people tend to avoid domains in which they feel less proficient (low self-efficacy) and readily initiate those tasks at which they perceive themselves to be capable (high self-efficacy). Self-efficacy perceptions also predict the amount of engagement with a task, the effort expended, and how long a person will persist at a task once initiated (Pajares, 1997, 2002). Accordingly, a student's perceived self-efficacy for success in the virtual world learning environment influences how the student will behave in that environment, including their motivation to learn and their persistence when faced with difficulties or failure.

Statement of the Problem

Interest in the use of MUVEs as interactive learning environments has increased substantially in recent years. However, a significant concern is that some students face initial problems in becoming familiar with the virtual environment. A lack of experience with the virtual world paradigm can cause spatial disorientation (Burigat & Chittaro, 2007; Darken & Sibert, 1996; Ruddle, Payne, & Jones, 1997) and technical difficulties with the interface (Warburton, 2009). These issues can disrupt effective interaction with the environment, cause frustration, and discourage students from persisting in their use of virtual worlds (Nash, Edwards, Thompson, & Barfield, 2000). In an attempt to address this issue, educational institutions often require students to complete a pre-training familiarization course within the MUVE, presumably to increase self-efficacy prior to attending virtual classes, reduce frustration caused by unfamiliarity and increase the likelihood of persistence. Researchers have found that these are important factors to address: "3-D virtual environments require advanced technology resources, appropriate training and orientation before users can be expected to perform specific tasks, and adequate time for users to become familiar with the environment" (Eschenbrenner, Nah, & Siau, 2008, p. 105). However, the approach to designing these training and orientation courses is inconsistent at best and haphazard at worst. Training often consists of a simple obstacle course that students are left to explore with minimal guidance beyond text cues on virtual signs, and there are no empirically supported recommendations on the design of familiarization training courses for MUVE learners. Studies have suggested specific methods for helping users familiarize themselves with the virtual world paradigm (Van Dijk, Op den Akker, Nijholt, & Zwiers, 2003; Vinson, 1999), but there is little information available on whether these practices have been adopted or if deployed pre-training courses are effective in preparing students for MUVE-based learning. By incorporating motivational affordances, improving self-efficacy and satisfaction, gamification could decrease the probability that students will give up when encountering difficulties during the pre-training course and foster the confidence to choose MUVE-based learning in the future.

Rationale

Learning in MUVEs has been studied extensively, with important findings regarding the pedagogical efficacy of the format (McKerlich & Anderson, 2007), best practices in instructional design (Molka-Danielsen, Richardson, Deutschmann, & Carter, 2007), problems with familiarization (Nash et al., 2000; Vinson, 1999), learner engagement (Gregory & Gregory, 2011), and effects on learner satisfaction (Verhagen, Feldberg, van den Hooff, Meents, & Merikivi, 2011). Researchers have also examined the effect of pre-course training on computer-based learning (Mayer, Mathias, & Wetzell, 2002; Mayer, Mautone, & Prothero, 2002). While many recommendations from these studies may transfer to MUVE-based learning itself, the literature fails to recommend specific and concrete techniques to *prepare* learners for educational activities within the virtual world paradigm.

As an approach that has been shown to improve skill acquisition, user satisfaction and persistence, gamification has the potential to improve learner preparation for MUVE courses by enhancing self-efficacy perceptions and fostering persistence in the face of orientation challenges. However, no research has been undertaken within this specific context. Empirical evidence of improved outcome expectancies, persistence in the face of difficulties, task skill and learner satisfaction would support the significance of gamification for preparing students for MUVE-based learning and provide valuable guidance for instructional designers seeking to create effective familiarization experiences.

Research Question and Hypotheses

RQ: What are the effects of the gamification of familiarization training courses on the factors related to learning in virtual worlds?

H_A: Participants in a gamified virtual world familiarization course will measure significantly higher for the following variables while practicing virtual world familiarization tasks compared to those who participate in a non-gamified course:

H_A1: Self-efficacy

H_A2: Persistence

H_A3: Task skill

H_A4: Satisfaction

Overview of Methodology

Study type and structure. Using an experimental study design, participants were randomly assigned to one of two groups. Both groups completed a virtual world familiarization course. The courses were identical except that the first group's course did not feature any gamification elements while the second group's course was gamified using best practices derived from the literature. All participants were informed that their completion of the course would lead to a certificate of virtual world competence. The gamified treatment group was informed that their certificate would display the actual achievement rewards earned during the course. For each treatment, the familiarization course was divided into sections based on the skillset being developed. These sections were known as *skillset stations*. At each station, the participant familiarized himself or herself with the skillset on a practice substation until he or she felt ready to be assessed (the Learning Phase). When ready, the participant completed an assessment substation for that skillset where their task skill was measured (the Assessment Phase). Participants could attempt the assessment phase as many times as they wished. The self-directed nature of the course allowed for persistence behavior to be observed and measured.

Data collection. Participants in both groups completed pre- and post-tests for virtual world self-efficacy. Participants also completed a post-treatment learner satisfaction survey that measured satisfaction for the familiarization course. Participants' activity in the treatment was video recorded to capture persistence behavior and task skill. A qualitative interview was conducted immediately after completion of the familiarization course.

Analysis. Results of the pre- and post-tests for self-efficacy were quantitatively analyzed to measure changes in participant self-efficacy and to compare the variance of the two groups. Recordings of activity were used to determine skill and persistence scores for each task. These measurements were then quantitatively analyzed to compare participants' skill and persistence at completing tasks within the MUVE. Results of the learner satisfaction survey were quantitatively analyzed to compare levels of student satisfaction for the familiarization course. Post-session interview responses were qualitatively analyzed using grounded theory. Specific analysis techniques are discussed in Chapter 3.

Definition of Terms

The following is an alphabetical list of terms and definitions for how they are used within this report:

Avatar: A humanoid depiction of the user's agent within the MUVE. The avatar performs actions at the user's direction, such as walking, gesturing, and interacting with objects in the virtual world.

Familiarization: The process through which the MUVE user becomes comfortable with the virtual world paradigm of being represented by an avatar, being present in a virtual space, controlling the avatar, and viewing other users as avatars within the shared virtual space.

Gamification: The inclusion of game-like motivational elements in otherwise non-game activities. Typical gamification elements include scores, achievements, bonuses, leaderboards, ranks and levels. Gamification can foster a sense of competition with oneself, peer-based competition, or both. Gamification of pre-training was the independent variable for this study.

MUVE: Multi-User Virtual Environment - A computer-based simulation of a large-scale environment in which several users can be present simultaneously. The virtual world is hosted on a central server and accessed via network (often the Internet).

The MUVE is viewed as a two-dimensional animated rendering of a three-dimensional space through a client program on the user's computer. The user becomes "present" in the virtual world by controlling an avatar via keyboard and mouse input. The MUVE used for this study was *Second Life*, a commercial virtual environment owned and run by Linden Labs, Inc.

Outcome expectancy: The perceived likelihood of a behavior resulting in a specific outcome. As part of self-efficacy theory, it is considered separate and independent from self-efficacy expectancy and appears to positively influence the intention to continue the behavior (Maddux, Sherer, & Rogers, 1982). Therefore, outcome expectancies can be an indicator of future actions.

Persistence: The degree to which an individual is prepared to expend effort at a task in order to achieve a desired result. Persistence was measured by the number of attempts at a given familiarization task.

Presence: A psychological state in which the user of a MUVE has, to some degree, a sense of "being there" – not seated in front of a computer system, but instead present within the virtual space, embodied within the avatar and acting through it: "Presence is […] the extent to which the unification of simulated sensory data and perceptual processing produces a coherent 'place' that you are 'in' and in which there may be the potential for you to act" (Slater, 2003).

Pre-training: A course of study or coaching designed to prepare the participant for an upcoming learning experience. Pre-training is necessary when the learning experience requires skills, knowledge or abilities that the participant does not already possess, and so in order to learn effectively the participant must receive training prior to the main learning experience.

Satisfaction: The perceived degree to which the participant's expectations or needs have been fulfilled by the experience of the virtual world familiarization course. In this study, satisfaction was measured by participant survey.

Self-efficacy: An individual's perceived ability to perform a task successfully. In this study, self-efficacy was a dependent variable that referred to the participant's perceptions of their ability to complete tasks within the *Second Life* virtual world. It was measured by participant survey.

Task skill: The participant's ability to complete a task as part of the familiarization course in *Second Life.* Task skill was a dependent variable in this study. It was calculated using a formula that included the time taken to successfully complete the assessment phase of the skillset station and the number of errors committed while doing so. The task skill score was increased proportionally by a faster time and decreased proportionately to the number of errors.

Chapter 2: Literature Review

Background

With its roots in the earliest of correspondence courses offered in the 18th century and the creation of the United Kingdom's mail-based Open University in 1969, interactive online education is the use of Internet technology for teaching and learning at a distance (Bower & Hardy, 2004). It can feature asynchronous exchange of information using discussion boards (which allow a time delay between posting, reading and replying) and synchronous communication in which all participants are connected in real time and can exchange information almost instantaneously in text-based chat rooms or through video conferencing connections. Since the mid-1990s, these Web-based tools have been integrated into course management systems (CMS), server-based online environments that allow instructors to post course materials and enter grades; and students to converse, collaborate and submit their work (Watson, 2013). Course management systems have also been referred to as virtual learning environments (VLEs). While the term "virtual" may imply something more, these systems remain primarily text-based, asynchronous repositories for course information and mediums for student-student and student-instructor communication that serve as replacements for the traditional classroom-based modality (Bower & Hardy, 2004, p. 8; Browne, Jenkins, & Walker, 2006).

The most recent evolution of the online education paradigm is the employment of virtual worlds as online learning environments. Unlike the CMS or VLE, virtual worlds are simulations of three-dimensional space generated by a computer system. A virtual world can exist on a standalone system with only one user present within the virtual environment, but the multi-user variant of the virtual world, known as a MUVE (multi-user virtual environment) is accessed simultaneously over the Internet by several users who are geographically separate (Duncan, Miller, & Jiang, 2012, p. 950). Users connected to the MUVE control an avatar - a virtual embodiment of themselves - using keyboard input, a computer mouse or a game controller. Through the graphical display, users can experience visual feedback of their actions and interact with other users' avatars in real time. Synchronous audio communication occurs between users who typically wear headsets and converse as if they were in the same physical space.

The range of platforms used for educational MUVEs has been notably diverse, from purpose-built, custom worlds using 3D engines such as *Unity* and *OpenSim* to systems intended as social environments like *Second Life* that were later co-opted as educational destinations. *Second Life* has been a highly popular choice for institutions wishing to offer MUVE classes. In an extensive survey of instructors on *Second Life*, Bowers, Ragas and Neely (2009) found that *Second Life* was used most heavily by instructors from the United States and the United Kingdom, but that 13 other countries were also represented, indicating the cross-cultural appeal of MUVE-based learning. Bowers et al. also found that 25 different academic disciplines were using *Second Life* as a learning medium, demonstrating the potential of MUVEs to support diverse fields of academic study (p.717). Another approach to the provision of MUVE-based learning has been to co-opt certain video games that already feature a multi-user virtual world (Van Eck, 2006). Educators and researchers have used Massively-Multiplayer Online Role Playing Games (MMORPGs) such as *World of Warcraft* and *EverQuest* to reduce the duration and cost of developing a MUVE for hosting learning activities, leverage learner familiarity with the environment as well as utilizing game functions as motivational affordances (Carr, Oliver, & Burn, 2010; Dickey, 2007; Pirius & Creel, 2010; Rankin, Morrison, & Shute, 2009).

The Educational Use of Virtual Worlds

The use of MUVEs for educational activities is becoming increasingly common. *Second Life* has had more than 150 colleges set up instructional programs within its expansive digital realm (Foster, 2007; Jennings & Collins, 2007). Institutions of higher education were early adopters of virtual worlds for learning, but the technology has also been used for learning activities in primary (F. Bailey & Moar, 2001; Ligorio & van Veen, 2006) and secondary schools (O'Connell & Groom, 2010).

Rather than simply replacing text-based course management systems with a more graphically rich, immersive version of the same functionality, there are significant differences in how MUVEs are used for education. According to Duncan et al. (2012), the 3D simulations provided by MUVEs allow students to experience real or imagined world events and sometimes work collaboratively on educational projects in real-time. They also found that collaborative simulations were the most commonly studied of the educational activities undertaken in MUVEs (p. 957).

The literature indicates that MUVE-based learning is perceived by students to be more enjoyable than text-based online learning (Holmberg & Huvila, 2008), and – through the creation of *presence* – has the potential to close the perceived interactional gap between online distance education and face-to-face classroom learning (J. G. Jones, Morales, & Knezek, 2005). Indeed, it has also been proposed that by simulating the everyday interactional learning that occurs in the real world, the use of virtual reality is a powerful tool in producing constructivist learning experiences: "Immersive VR allows us to create from our experiences the kind of knowledge that has hitherto been accessible only through direct experience of the world" (Winn, 1993).

Further research has studied the transferability of effectiveness indicators from text-based online learning environments to the MUVE-based model. In their qualitative observational study, McKerlich and Anderson established that the oft-used Community of Inquiry (CoI) model was both recognizable and applicable in determining the learning effectiveness of MUVE activities (2007, pp. 46–47). The Community of Inquiry model – first systematized for education by Lipman (1991), based on prior work by Addams, Peirce and Dewey (Shields, 1999), and further applied to online learning by Garrison, Anderson and Archer (1999) – holds that the educational experience is informed by three types of "presence" in the learning environment:

- **Social presence**: Interactions that support engagement and discourse between participants and the overall climate of the environment.
- **Cognitive presence**: Interactions that mediate learners' goals, the expression of those goals within social discourse, and the regulation of learning activities.
- **Teaching presence**: The instructor's influence in designing the course of study, setting the learning direction, and regulating the learning climate.

(Garrison et al., 1999)

McKerlich and Anderson observed that the overall sense of presence provided by virtual world avatars supported all three CoI presence domains and appeared to be partially responsible for the high ratings of CoI indicators observed in MUVEs (2007, pp. 46–47).

Experiential learning (Kolb, 1984) and situated learning (Lave & Wenger, 1991) are the predominant theories supporting students' involvement in simulation activities for learning.

David Kolb noticed commonalities between models of learning by Lewin, Dewey and Piaget and combined them to form his experiential learning theory. Kolb built upon Lewin's four-part experiential learning cycle, where (1) a concrete experience is followed by (2) reflection about the experience, from which (3) the learner forms abstract concepts and generalizations which are then (4) tested out in new situational experiences. Kolb suggested that experiential learning could begin at any one of the four points and that the cycle should be more of a cumulative spiral, in which each new experience builds upon the experiences that came earlier and the knowledge that was derived from them. Fundamentally, Kolb observed that human learning approximates scientific inquiry (1984, p. 32), in which the learner observes experiments, reflects upon the results and forms hypotheses that can then be tested in future experiments. Experiential learning theory also notes that learners are not merely passive participants in their experiences but have a degree of agency in directing their own learning. This agency is seen in the transactional relationship between learner and environment. Kolb observed that learners will often use their environment to create situations that cater to their learning objectives (1984, p. 36).

Situated learning theory (Lave & Wenger, 1991) extends experiential learning by proposing that learning occurs through experiencing social interaction while immersed in *communities of practice*: Groups based around some shared activity or knowledge domain. Individuals are members of multiple communities of practice, and for each community can be deeply embedded, on the periphery, or anywhere in between. As members learn and grow in competence through social participation in the community, they can move from the periphery to the center. Learners are thus *situated* within these interactional relationships with others, their shared activities and the world in which they learn. The theory emphasizes the importance of learning in context rather than the abstract: Situatedness (and thus learning) is dependent on context to provide the interactions necessary to the social construction of meaning. Could simulations serve as a medium in which situational context can be created where it might otherwise be difficult? Conventional, non-MUVE simulations have been used in science and engineering education to learn about a set of circumstances and how change occurs over time. However, the learner is inherently removed from the simulation itself, limited to observing from an external viewpoint. Proponents argue that MUVEs can take simulations to a new level of experiential learning: allowing the learner to *enter* the virtual world where the experiment is taking place and become an embodied part of it through their avatar (Gee, 2007, pp. 81–83, 105). Instead of simply observing a simulated phenomenon from the outside, the learner *experiences* the phenomenon and its effect on the virtual environment through his or her in-world avatar. The learner can also interact with the phenomenon in a way that feels more direct, and can immediately experience any effects of the interaction. This is referred to as *situated understanding* (Shaffer, Squire, Halverson, & Gee, 2005, p. 106). Squire (2007) characterizes situated learning in MUVEs as "goal-directed activities in which we are meaningfully engaged and invited to take on the identities of experts" (pp. 52-53). This application of situated learning theory to MUVE-based simulations aligns with Lave and Wenger's view that situatedness is comprised of meaningful interactions between the learner, his/her activities within the world and his/her understanding of the world itself (1991, p. 33).

As with conventional non-MUVE simulations, MUVEs maintain the benefit that students can practice critical skills in a psychologically safe situation without fear of serious consequences when mistakes are made (Boulos, Hetherington, & Wheeler, 2007).

The potential for students to "learn by doing" while fulfilling a situated role in an participatory, immersive MUVE environment is a common theme throughout the literature (Bronack, Riedl, & Tashner, 2006; Chittaro & Ranon, 2007; Winn, 1993), and can therefore be considered part of the rationale for the appropriate use of MUVEs in educational activities.

Difficulties Associated with Using Virtual Worlds

In order to ascertain the emphasis and content of a pre-training course intended to prepare students for learning in a MUVE-based course, the typical difficulties related to MUVE use must be identified.

Interface usability. Inexperienced users often face difficulties with learning how to use the interface of MUVEs such as *Second Life*. If these challenges are experienced at the start of an academic course, learning could be negatively impacted as the student struggles with the controls and is distracted from, or may be unable to engage with, didactic activities. The student must become proficient in the use of the MUVE interface before starting a learning program. Technical competencies crucial to effective use are avatar movement controls, navigating, creating objects and parsing 3D visual elements correctly (Warburton, 2009). A familiarization pre-training course must foster these essential skills in order to overcome interface usability issues.

Dimensional translation and presence. The visual display of a virtual world is a two-dimensional (2D) representation of a three-dimensional (3D) virtual space that exists only in the abstract. Users must disregard the two-dimensional nature of the display and visualize it as three dimensions more akin to the real world. This is achieved by mentally translating 2D perspective on the screen into depth information. At the most basic level this is not cognitively burdensome, as humans are used to performing this dimensional translation with other 2D depictions of 3D space such as paintings, photographs and television. Unlike these formats, however, users of virtual worlds are able to control their movement within the virtual space. This requires users to perform a unique psychological feat: imagining that they are somehow present and oriented *within* the depicted space. When sufficient mental resources are allocated to tasks within the virtual environment, the user's consciousness can shift away from the physical real world space to feel as though they now occupy the virtual world instead (Bystrom, Barfield, & Hendrix, 1999, p. 243). This phenomenon is known as *presence* (Barfield & Weghorst, 1993). Presence is a state of consciousness that requires users to shift most or all of their awareness from the real world to the virtual environment, including cognitively translating between the 2D display and the 3D nature of the virtual world and moving within it. As presence is thought to be a prerequisite to effective performance within virtual environments (Bystrom et al., 1999), difficulties with dimensional translation in a MUVE-based learning environment have the potential to interrupt the state of presence and impair the performance of learning activities in the MUVE. Studies have shown that presence tends to increase as users become more familiar with the virtual environment through practice (Barfield & Weghorst, 1993; Heeter, 1992), providing a rationale for the use of familiarization pre-training courses.

Spatial disorientation. Even when the user is consciously present in the virtual world, disorientation can occur due to the lack of navigational affordance and spatial feedback compared to the real world (Smith & Marsh, 2004). Ongoing disorientation can result in user frustration, dissatisfaction with the experience and eventual withdrawal from the environment altogether (Nash et al., 2000). One solution to disorientation is to provide the user with navigational aids such as maps, signposts, pointers and landmarks (Darken & Peterson, 2002; Parush & Berman, 2004; Vinson, 1999). However, the user's ability to benefit from navigation aids appear to be mediated by the amount of virtual world experience possessed by the user entering the MUVE, especially when the navigation aids require the student to translate a two-dimensional indicator to the three-dimensional Euclidean space represented by the MUVE (Burigat & Chittaro, 2007). The initial disorientation experienced by new MUVE users dissipates within a few hours of experience with virtual world movement and navigation (Ruddle et al., 1997). This indicates that increasing the amount of virtual world experience through a pre-training course will improve users' ability to find their way around the virtual space, reduce their potential for disorientation, and allow them to focus less on navigation and more on salient tasks in the environment.

The Gamification of Non-Game Activities

Over the past decade, researchers have focused increasingly on gamification as a potential tool for improving motivation. Could the challenge and satisfaction of gameplay be fostered in tasks that might otherwise be tiresome, and therefore increase motivation for those tasks?

In a Finnish study designed to assess the psychophysiological responses to gameplay, participants played video games while the researchers took various biological measurements (Ravaja, Saari, Salminen, Laarni, & Kallinen, 2006). By measuring and recording heart rate, skin resistivity and facial muscle activity along with synchronized audio and video signals from the game, the authors were able to correlate in-game events with physiologic reactions. The results included a statistically significant physiologic arousal response to in-game rewards such as scoring bonuses. The participants' heart rate increased in frequency, the skin became more conductive, and the tension in facial muscles associated with positive emotions increased. Studies by Segal and Dietz (1991) and Wang and Perry (2006) found similar heightened physiologic responses to gameplay. In Ravaja, et al. (2006), the level of biological arousal response was more pronounced when the score of the reward was numerically greater, showing a mostly linear proportional relationship (p. 365). The responses observed were those considered to be of a positive nature, even in response to a negative game event such as the player "losing a life.” The authors suggested that this is because the participant knows that it is “only a game” and therefore negative events can still be enjoyed as part of the game experience (pp. 361-362). It can be seen, then, that humans derive significant pleasure from game activities, and not simply from winning at them – the entire experience of playing is an enjoyable activity.

Flatla, Gutwin, Nacke, Bateman and Mandryk (2011) developed several simple video games to experimentally test how system calibration tasks might be made more engaging and less tedious. Tasks included the setting of color parameters for a video display or normalizing a control-to-display ratio. Participants greatly preferred completing the “gamified” calibration tasks to completing the standard versions, and participants put more forth more effort and focused more intensely on the calibration task when it was presented as a game.

As motivation is a crucial component of formal academic learning, educators have understandably become interested in the application of gaming elements to learning processes and environments. Lee Sheldon’s 2011 book *The Multiplayer Classroom* chronicles his experiments with introducing a game structure to courses he taught at Indiana University. Sheldon defined gamification as “the application of game mechanics to non-game activities. Its underlying idea is to increase engagement” (2012, p. 75). In addition to increasing engagement, the results of gamification appear to have a positive effect on academic achievement too. In one of Sheldon’s case studies, 62% of students in a high school biology course were passing the course with a grade of D or higher. After the course was gamified using Sheldon’s methods, 98% of students held a grade of D or higher. In the same course, 10% had a grade of A or B before gamification, rising to 36% after gamification methods were introduced. The teacher noted that the gamified course improved student motivation, attitude and performance (Sheldon, 2012, pp. 55–56).

Barata, Gama, Jorge and Gonçalves (2013) found similar results in their study in which a blended (face-to-face and online) engineering course was used to measure the effect of gamification on engagement and satisfaction. The course had suffered from low attendance, infrequent participation in online activities and a lack of interest in reference materials available for download from the VLE. The study took place over two years, with the course being deployed once each year. The first year of the study, the course was run with no game elements and with researchers collecting data on students' attendance, participation in class activities, and use of course materials. During the second year, the authors gamified the course by adding a points system, challenges, leaderboards, badges and levels of achievement. The study showed several significant differences between the non-gamified and gamified courses. The total number of student downloads of course materials in the gamified course was 47% higher than that in the non-gamified course ($U = 75$, $Z = -2.917$, $p < 0.003$). Discussion post activity was the most significantly impacted metric in the study, with an 845% growth ($U = 36.5$, $Z = -4.314$, $p < 0.00001$) in initiated threads and an increase of 511% in reply posts ($U = 52.5$, $Z = -3.731$, $p < 0.0001$) between the non-gamified course and the gamified version. The authors attributed part of the increase in student discussion activity to small changes in the assessment criteria, but the very large jump in post frequency cannot be fully explained by the minor changes in evaluation, suggesting a positive and significant increase in engagement when the course was gamified (2013, p. 6). The findings were not limited to the online component of the blended course. Attendance at face-to-face class sessions increased 11% ($U = 51$, $Z = -3.654$, $p < 0.001$) when the course was gamified. When surveyed, students reported feeling more motivated and interested in the course when it was gamified. The survey instrument for student motivation and satisfaction was not validated, and the survey was only administered to students in the gamified course. This renders the satisfaction survey results questionable, but does not detract from the other results cited above.

The effect of gamification on online learning was investigated in a Australian study of over 1,000 undergraduate students where an online question and answer system was modified to include a single dimension of gamification: earning achievement "badges" (Denny, 2013). A control group used the system in its usual form (without badges being available) and the experimental group was able to earn badges in proportion to their use of the system. The study demonstrated that the opportunity to earn badges had a significant and positive effect on the number of answers submitted by students, with the experimental group answering 22% more questions than the control ($U = 116386.5$, $Z = 3.45$, $p < 0.001$), and on the number of days that students were active in the system, with the experimental group 13% more active than the control ($U = 116584.5$, $Z = 3.41$, $p < 0.001$). There was no significant difference in the number of correct answers submitted by the two groups, suggesting that gamification did not affect academic achievement. However, this may be attributable to the construction of the badge system: of the 22 available badges, only six were related to submitting correct answers (2013, p. 766).

Another study examining the effects of gamification on learning, participation, and student attitudes provides some important nuance in the potential effects of gamification. The study found that student performance in skill acquisition was significantly and positively affected when e-learning practical assignments were gamified as compared to a non-gamified approach (Domínguez et al., 2013). The results also suggest that while gamification could positively affect skill acquisition, non-gamified approaches were more effective for knowledge acquisition. The motivational aspects of gamification were also examined in this study. The authors considered gamification to have a potentially positive impact on motivation, but saw that it was highly dependent on the choices made by instructional designers. Of particular note is that students were significantly less motivated when gamification elements were perceived to increase the time commitment required of the student (2013, p. 389). Another study of the effects of gamification and social networking on academic performance by some of the same authors confirmed the results of the Domínguez, et al study (de-Marcos, Domínguez, Saenz-de-Navarrete, & Pagés, 2014). It is noted that, as the gamification interventions themselves were focused more on the learning of skills, the study designs could have biased the results in favor of skill acquisition (de-Marcos et al., 2014, pp. 86–87). Another interesting note is that the authors found that student participation (as measured by attendance at lectures and contributions in the e-learning environment (2014, p. 85)) was negatively affected by gamification and postulated that presenting learning activities as a game can counter-productively promote competition in place of collaboration (2014, p. 88). While this finding is salient to the design of collaborative online learning activities, the competitive aspect of gamification is unlikely to negatively affect participation in a familiarization course where the student is not expected to collaborate with others.

A literature review of 24 empirical studies measuring the effectiveness of gamification was conducted by Hamari, Koivisto and Sarsa (2014) and provides a useful classification of game elements considered as *motivational affordances*, such as points, leaderboards, goals, levels and rewards. The review also serves as an excellent summation of the literature's conclusions regarding gamification:

- Gamification appears to produce positive effects on engagement and enjoyment. *This may suggest that gamification has potential for increasing persistence in the face of motivational challenges such as anxiety or boredom.*

- When measured in the learning context, gamification appears to have a largely positive effect on academic outcomes. *This could be an indirect relationship, more attributable to increased engagement (e.g. engaged students participate more, are exposed to more instruction and content and therefore tend to succeed academically), but increased engagement could be an indicator of enhanced motivation.*
- The motivational effects of gamification may be partially attributable to the novelty effect and thus may not persist into the longer term. *This has implications for when gamification may be more contextually appropriate, such as learning a new system or paradigm.*
- In the educational context, gamification can produce negative effects, such as those produced by increased competition and individual participant preferences, such as a dislike of gamified elements. *This also has implications for selecting the appropriate contexts for gamification.*
- Most studies of gamification focus on behavioral outcomes, and of these, only one utilized validated instrumentation. *This suggests that there is a need for more rigorous measurement of behavioral phenomena in gamification studies.*
- Some studies suffered from methodological weaknesses such as small sample sizes, lack of control groups and sole reliance on participant evaluation. *This further suggests that experimentally rigorous studies are needed to fully investigate the potential benefits of gamification.*

Self Efficacy

Theory. Bandura (1986) noted that that an individual's perception of their own competence at a specific task domain – their *self-efficacy* – is a mediating variable in their motivation to attempt a particular task (Schunk, Pintrich, & Meece, 2008, p. 122). Bandura referred to this combination of perceived competence and projected future results as an *outcome expectation*:

> "[p]eople act on their beliefs about what they can do, as well as their beliefs about the likely outcomes of performance. The motivating potential of outcome expectancies is thus partly governed by self-beliefs of capability" (Bandura, 1993, pp. 128–130).

Bandura's social cognitive theory integrates the concept of self-efficacy into a framework of *triadic reciprocality*: a model in which human performance consists of personal, behavioral and environmental factors and the interactions between them (Bandura, 1986). Self-efficacy is considered a personal factor along with cognition and affect. Behavioral factors are how the individual acts in response to events and experiences. Environmental factors are elements external to the person that can affect and modify personal and behavioral factors. The individual's environment in turn can be influenced by behavioral and personal factors (Schunk et al., 2008, pp. 126–128). As the personal factor of self-efficacy impacts behavioral factors such as choice of task, effort and persistence (Schunk, 1989, 1995; Schunk & Pajares, 2002, as cited in Schunk et al., 2008) and these both influence – and are influenced by – the individual's environment, this interactional model may be crucial to understanding the motivation of students in virtual world learning environments.

Impact on performance. In a comprehensive meta-analysis of 114 studies with a collective $N = 21,616$, Stajkovic and Luthans (1998) found an average correlation of $r = .38$ ($p < .01$) between task self-efficacy and task performance. Converting this correlation to a common effect size metric gives a Cohen's $d = .82$, meaning that self-efficacy accounts for a 28% improvement in performance (1998, p. 252). The authors note that this figure may be conservative due to being suppressed by the tendency to avoid tasks that people judge to be beyond their perceived abilities. Another important finding of this analysis is that the complexity of the task has a strong moderating effect on the relationship between self-efficacy and performance, with the magnitude of the correlation strongest when the task was simple and weaker with more complex tasks. Thus self-efficacy perceptions have been shown to significantly and positively predict task performance, especially with simple tasks.

Engaging with technology. Domain-specific self-efficacy perceptions are instrumental in mediating the willingness to engage with computer systems. Computer self-efficacy negatively affects computer anxiety, while the some of the contributing factors of efficacy – prior experience with computers, perceived value of the system interaction and perceived ease-of-use – positively influence usage (Fagan, Neill, & Wooldridge, 2003; Igbaria & Iivari, 1995). Additionally, personal affective traits such as a negative attitude toward technology or innovativeness can significantly influence the magnitude of both computer anxiety and self-efficacy (Thatcher & Perrewé, 2002).

While the effect of self-efficacy on computer usage is established, the link between self-efficacy and learning through online technology is less clear, requiring a more granular examination of possible mediating factors, such as the importance of how closely the user's expectations match their actual experiences in the learning system (Hayashi, Chen, Ryan, Wu, &

others, 2004). This relationship between expectation and experience is more closely associated with satisfaction rather than efficacy, and can predict whether learners will continue the use of online learning.

Efficacy in virtual worlds. A team including Clarke, Dede, Ketelhut, Nelson, and others examined motivation in MUVE-based situated learning repeatedly in several studies between 2002 and 2010. Their research was centered around a MUVE designed to teach scientific inquiry to middle school students. In seeking to determine the most effective learning strategies for MUVE-based learning, they found a significant correlation between the frequency of use of a meta-cognitive guidance system and positive outcomes in subsequent learning assessments (Ketelhut, Dede, Clarke, Nelson, & Bowman, 2007; Nelson, 2007), suggesting that guided learning in MUVEs is likely more effective than unguided, exploratory learning approaches. Further experimentation showed that embedded guidance strategies employed in MUVEs can significantly improve students' self-efficacy perceptions as practitioners of science (Ketelhut, Nelson, Clarke, & Dede, 2010; Nelson & Ketelhut, 2008). These results demonstrate that traditional classroom techniques for boosting self-efficacy can successfully transfer to the virtual environment. For example, a traditional technique for prompting students to reflect on their practice is for the teacher to ask meta-cognitive questions. This can transfer to the virtual environment by allowing students to access system-based questions themselves and thus self-regulate their learning. A key recommendation to emerge from this research is that MUVE-based learning experiences should be designed with meta-cognitive guidance affordances that encourage learners to reflect upon their practice as they build their understanding and skillset. Also important to the current study, these results suggest that users who complete an appropriately designed familiarization course can be expected to increase their self-efficacy for the targeted domain of successful MUVE use.

Measurement. Self-efficacy can be considered at a general level, referring to an individual's perceptions of their overall ability to cope with a broad range of challenges. This general self-efficacy is considered to be an overall personality trait rather than a domain-based belief system and thus conceptually different from task-specific self-efficacy (Bandura, 1986). Accordingly, researchers have developed separate psychometric scales to measure general self-efficacy (Schwarzer & Jerusalem, 1995; Sherer et al., 1982). In a comprehensive guide regarding the formulation of self-efficacy instruments, Bandura (2006) reiterated that an individual's self-efficacy varies across different domains. For example, an experienced teacher might hold high self-efficacy perceptions in teaching schoolchildren but low self-efficacy in riding a motorcycle. Therefore, a general self-efficacy scale is not useful when investigating specific behavioral phenomena. In these cases, self-efficacy should be measured using scales targeted for the specific domain of interest (Bandura, 2006, pp. 307–308; Marakas, Johnson, & Clay, 2007). As self-efficacy refers to perceived capabilities, Bandura emphasized that the wording of self-efficacy scales should focus on what the respondent feels like they *can do* and to what degree they can do it, not what they *should* do or what they *intend* to do: imperatives and intentionality are not relevant psychometrics in pure self-efficacy instrumentation (Bandura, 2006, pp. 308–309). Another important point emphasized by Bandura concerns the multicausality of self-efficacy: several different behavioral factors over which individuals have control can contribute to their efficacy beliefs. Instruments should be constructed with careful attention to all relevant perceived capabilities that contribute to the individual's self-efficacy for the domain, not just those directly related to behaviors in the current study (2006, p. 310).

DeNoyelles, Hornik and Johnson (2014) demonstrated that self-efficacy manifests in a multidimensional manner in virtual world learning environments. Their study of 486 accounting students, conducted using *Second Life*, hypothesized three dimensions of virtual world self-efficacy:

- **Virtual World Environment Self-Efficacy (VWE-SE):** "[S]elf-efficacy for the basic skills that are necessary for students to navigate and complete tasks in the virtual world environment" (2014, p. 263).
- **Virtual World Learning Domain Self-Efficacy (VWLD-SE):** Self-efficacy "to use and manipulate the learning objects embedded in the virtual world to accomplish domain-specific tasks" (2014, p. 264).
- **Learning Domain Self-Efficacy (LD-SE):** "[S]elf-efficacy within the learning domain, which focuses on the particular learning domain skills of interest" (2014, p. 264).

As this study was focused on student familiarization with the MUVE, the first dimension, VWE-SE, is of most interest. This includes non-learning-specific task skills such as walking, navigation, flying, moving around a virtual space such as a classroom, interacting with objects in the MUVE, and communication with other users using voice, text messages and avatar gestures (2014, p. 264). The study found a significant and positive correlation between students' perceptions about their ability to use the virtual environment (VWE-SE) and their self-efficacy beliefs regarding their interactions with objects in the learning domain (VWLD-SE), indicating that student familiarization with the MUVE is an important component supporting subsequent learning activity.

Persistence

Persistence is the refusal to give up on a particular endeavor, task or goal. The persistent individual will endure resistance or opposition and continue to try to succeed toward a goal regardless of difficulties (Bandura, 1986). A key model relating to persistence is expectancy theory (Porter & Lawler, 1968; Vroom, 1964), in which an individual's *motivational force* is determined by three factors in combination:

- *Valence*: The strength of the individual's desire for a particular outcome.
- *Expectancy*: Perceived probability that putting forth effort will result in desired performance goals.
- *Instrumentality*: Perceived probability that performance toward goals will lead to desired outcomes.

(G. R. Jones & George, 2007, pp. 523–525; Koontz, 2010, pp. 292–294)

Hence persistence is part of the effort mediated through expectancy of performance in pursuit of desired outcomes. This can be clearly related to the concept of self-efficacy as an individual's perception of their task competence and projected future results of performance. Indeed, persistence has been empirically linked to self-efficacy, with high perceived efficacy and outcome expectations generally serving as a significant predictor of persistent behavior (Jacobs, Prentice-Dunn, & Rogers, 1984; Lent, Brown, & Larkin, 1984). However, the correlation between efficacy and persistence appears to break down once individuals have attained sufficient skill in the task domain, because the increased skill level negates the need to persist in order to succeed (Schunk, 1989, 1991).

Mindset theory (Dweck, 2006) holds that when people believe that they can improve their intelligence through effort, known as having a *growth mindset*, they become more positively motivated in learning compared to those that believe intelligence is fixed. Research in this area suggests that persistence can be increased by fostering the belief that intelligence is malleable (Blackwell, Trzesniewski, & Dweck, 2007). To promote a growth mindset, positive feedback should focus on acknowledging effort rather than achievement and perseverance rather than intelligence (Dweck, 2007). In educational games, tailoring the in-game rewards in a way that encourages the development of a growth mindset can have a positive effect on task persistence (O'Rourke, Haimovitz, Ballweber, Dweck, & Popović, 2014). This provides important guidance on the most effective reward approach for promoting persistence in gamified learning systems.

Persistence can be observed at differing levels of granularity, and it is important that the distinction between general persistence and task persistence is articulated for the purposes of this study. General persistence refers to an individual's tendency to continue on a course of action toward a general goal, such as persisting to complete a degree program or persisting to learn how to drive a car. Task persistence applies to the lower level of tasks that are components of these overall goals. Therefore a student might be said to persist at learning statistical analysis as part of the completion of her degree, or persist at parallel parking as part of obtaining his driver's license. The current study was concerned with task persistence as applied to the familiarization tasks connected with learning how to interact with a multiuser virtual environment.

The literature provided two behavioral indicators typically used to measure task persistence: amount of time spent attempting a task (Sandelands, Brockner, & Glynn, 1988, p. 211) and the overall number of attempts at a task (Foll, Rascle, & Higgins, 2006, p. 594). For both indicators, greater values equal greater task persistence. These indicators were also found in a meta-analysis of self-efficacy and persistence studies by Multon, Brown and Lent, with the overall number of attempts producing a larger effect size (r = .47) than time spent on task

($r = .17$) (1991, pp. 34–35). It was also noted that measuring persistence using number of attempts is likely to be a more accurate metric, as time spent on task can be affected by high-efficacy individuals completing tasks faster than those with lower efficacy (1991, p. 35). This could confound results by making high-efficacy participants appear to have less persistence. In the Sandelands et al. study, portions of the task sets were deliberately designed to be impossible to test participants' persistence in the face of repeated failure (1988, p. 210).

Self-Efficacy, Persistence and Academic Outcomes

Echoing previous findings of a strong relationship between self-efficacy and work performance, a reliable correlation between high self-efficacy, persistence and improved academic outcomes has been established by prior research. Compared to those that doubt their capabilities, students with high self-efficacy tended to put forth more effort, independently chose more challenging tasks, demonstrated higher levels of persistence and were less emotionally discouraged when encountering adversity (Bandura, 1997; Zimmerman, 2000). These resultant factors consequently led to positive effects on academic performance and achievement.

Schunk's study of elementary schoolchildren used cognitive modeling techniques to strengthen students' self-efficacy in arithmetic tasks. Children that tested with higher self-efficacy persisted longer ($t\ (47) = 5.95$, $p < .01$) and achieved more success at arithmetic problems ($t\ (47) = 11.64$, $p < .01$) than those children who tested with lower efficacy and demonstrated less persistence. Another important factor was identified by this study: similar persistence levels were observed on students' pre-tests and post-tests, suggesting that persistence was likely affected by individual student work-rate preference as well as self-efficacy (Schunk, 1981).

Lent, Brown and Larkin (1984) conducted a study specifically examining the relationship of self-efficacy to persistence and academic achievement. Administering self-efficacy instruments to 42 college students, the researchers measured each participant's efficacy beliefs for their ability to perform the educational requirements and job functions of their chosen college major and related career path. Each participant's ensuing grade point average was then used as a measure of academic achievement. The number of semesters subsequently completed in their chosen major was used as a measure of persistence. The results showed significant mean differences for both achievement and persistence between high efficacy students and low efficacy students. While this study addressed general persistence (in the form of pursuit of an academic major), a later meta-analysis by some of the same authors confirmed the results at both the general- and the task persistence level (Multon et al., 1991). By conducting statistical meta-analyses of 36 prior studies, Multon et al. found that "self-efficacy beliefs account for approximately 14% of the variance in students' academic performance and approximately 12% of the variance in their academic persistence" (1991, p. 34). Effect sizes were $r = .38$ for academic performance and $r = .34$ for persistence.

Motivation, Engagement and Satisfaction in MUVEs

For learning experiences, motivation, engagement and satisfaction can be considered as parts of a reciprocal cycle. Motivation (to which self-efficacy contributes) is an input to the experience that supports initial and ongoing engagement: the drive to engage with an academic program or learning system and persist when challenged. Satisfaction is an output of the learning experience: an affirmative emotional state derived from personal accomplishment and a positive regard for the experience that then contributes to the learner's motivation to continue with similar experiences.

Motivational affordances. What are the design factors that contribute to a person's motivation to use a computer system for learning? Why use computer-based learning instead of the alternatives, such as traditional didactic instruction? While convenience may be a principal factor in initially choosing computer-based learning – such as the use of online courses to preclude the need for travel to a physical classroom – there must be additional elements to foster ongoing engagement and persistence with the learning system. These elements are the system's motivational affordances, and should be integral to the design of the instructional model and the system interface. Without this motivational support, learners may lose interest, become frustrated with technical problems and find it difficult to continue. "*Motivational affordances* comprise the properties of an object that determine whether and how it can support one's motivational needs" (Zhang, 2008, p. 145). Therefore, a system featuring appropriate and effective motivational affordances will make the user feel that it is enjoyable to use, that it assists them in their work, makes life easier and draws them back to use it repeatedly.

Zhang (2008) specified ten design principles for motivational affordances in information and communication technology (ICT) systems, organized into five categories of motivational sources and needs (Table 1):

Table 1

Summary of design principles for motivational affordance (Zhang, 2008, p. 146)

Motivational Sources and Needs	**Design Principles**	**Some Existing Design Examples**
Psychological: Autonomy and the Self	Principle 1. Support autonomy. Principle 2. Promote creation and representation of self-identity.	Desktop skins, cell phone ring tones, online avatars, application toolbar customization.
Cognitive: Competence and Achievement	Principle 3. Design for optimal challenge. Principle 4. Provide timely and positive feedback.	Games and learning systems with various challenge levels and immediate performance feedback.
Social & Psychological: Relatedness	Principle 5. Facilitate human-human interaction. Principle 6. Represent human social bond.	Group based games (e.g. online bridge) with a chat section, visualizations of email exchanges over a period of time to show both tasks and social related messages.
Social & Psychological: Leadership and Followership	Principle 7. Facilitate one's desire to influence others. Principle 8. Facilitate one's desire to be influenced by others.	Blogs (satisfy one's desire to influence by authoring, and to be influenced by reading), virtual communities where leaders sometimes emerge.
Emotional: Affect and Emotion	Principle 9. Induce intended emotions via initial exposure to ICT. Principle 10. Induce intended emotions via intensive interaction with ICT.	Slick/attractive look of iPod or cell phones, engaging games, ICT that induce optimal flow experience.

Principles 3, 4, 9 and 10 are of most relevance to a discussion of gamification for motivational purposes, as they pertain to the level of challenge, reward and emotional investment in the learning system.

Deterding combined the concept of motivational affordances with the gamification of situated, non-game activities: "*Situated motivational affordances* describe the opportunities to satisfy motivational needs provided by the relation between the features of an artifact and the abilities of a subject in a given situation, comprising of the situation itself (situational affordances) and the artifact in its situation-specific meaning and use (artifactual affordances)"

(Deterding, 2011, p. 2). The author goes on to emphasize that motivational affordances should exist to motivate initiation of the desired activity, then provide the participant with the perception of a satisfied need. This will lead to persistent activity in pursuit of continued reward, enjoyment and satisfaction.

At a more granular level of detail, Hamari et al. (2014) reviewed the literature to identify how gamification has been operationalized. As part of this, they listed the most common motivational affordances used in gamified learning systems. In order of highest incidence to lowest, they were:

1. Points system
2. Leaderboards
3. Achievements/badges
4. Levels
5. Story/Theme
6. Clear goals
7. Feedback
8. Rewards
9. Progress
10. Challenge

However, the authors observed that the literature has provided scant insight into which, if any, of these motivational affordances are most effective in fostering engagement and enjoyment in gamified learning systems. Indeed, the literature includes many studies of gamification in a variety of learning contexts, but investigations on its application in MUVEs specifically are rare, possibly due to the benefits of co-opting existing video game platforms rather than attempting to create game mechanisms in non-game environments. At the time of writing, a longitudinal study was under way to trial the use of gamification elements for assessment of learning in virtual environments, but the elements chosen were less common than those identified by Hamari et al. because they were specialized affordances for the learning program, and conclusive results were yet to be published (Wood, Teräs, & Reiners, 2013)

Game Elements for Intrinsic Motivation. Intrinsically motivated behavior refers to the self-determined, autonomous actions taken by an individual as opposed to those driven by the perception of external influences, controls or rewards (extrinsically motivated). The perceived locus of causality of intrinsic motivation is the self, and leads to actions that are performed out of personal interest and in anticipation of inherent satisfaction and enjoyment (Ryan & Deci, 2000b, pp. 72–74). Thus intrinsic motivation is theorized to have the highest degree of self-determination, where a person perceives a sense of agency, control and ownership of their actions. These factors contribute to a positive affect toward the task at hand, are psychologically satisfying, and thus enhance motivation (Ryan & Deci, 2000a, p. 65). Likewise, studies have indicated the factors that foster intrinsic motivation (such as learning in context, personalization, agency and control) appear to have positive effects on learning (Cordova & Lepper, 1996; Fransson, 1977; Vansteenkiste, Simons, Lens, Sheldon, & Deci, 2004).

The motivational affordances discussed previously represent extrinsic rewards, in that they are external to the person using them. One might expect any motivational drive from these game elements to be mostly extrinsic and therefore less self-determined than intrinsic rewards. In an experiment where participants were told that they would be paid an amount based on how many puzzles they solved within a certain timeframe, Deci (1971) found that the promise of financial reward actually decreased intrinsic motivation to continue solving puzzles during unpaid "free choice" time (1971, pp. 109–110). When the financial reward was replaced with positive feedback about the participants' puzzle solving, intrinsic motivation increased as compared to the non-rewarded control group (1971, p. 113). The author postulated that financial remuneration may have caused the participants to reevaluate the exercise as more of a "job" and view the offer of payment as a form of control, reducing his or her perception of autonomous self-determination and consequently, intrinsic motivation (1971, p. 114). These results have been further confirmed and explicated in subsequent studies. A meta-analysis of these (Deci, Koestner, & Ryan, 1999) concluded that intrinsic motivation is undermined by tangible rewards that are contingent upon engagement, completion and performance. However, certain extrinsic rewards can foster intrinsic motivation. For example, intrinsic motivation is enhanced by positive feedback that avoids implying control of participants' behavior. Research also indicated that the expectedness of a reward affects its impact: unexpected rewards tend to have a positive influence on intrinsic motivation, while expected rewards can undermine it.

In light of this research, it can be seen that carefully chosen extrinsic rewards can foster intrinsic motivation. Indeed, this mechanism can be considered the motivational lynchpin of game play: The goals of the game, extrinsic to the player, can become a source of intrinsic satisfaction when they are met and acknowledged through positive feedback, especially if the feedback is intangible (such as points, leaderboard position and player rank) and unexpected (such as point bonuses, ability enhancements and bonus challenge levels).

A 2007 analysis by Dickey examined the effect of the motivational factors typically found in Massively Multiplayer Online Role-Playing Games (MMORPGs). These are a type of MUVE in which users take on the role of a character and play within a themed, shared virtual world simultaneously with other players. In-game goals take the form of quests and exhibit many of the same design principles and motivational affordances identified by Zhang (2008) and Hamari et al (2014) respectively. Dickey (2007) focused on two hallmarks of MMORPGs that may foster intrinsic motivation to engage with the game: the role-playing of a character and the presence of a narrative or storyline. Noting that role-playing within a narrative is used as a form of simulation for learning in such fields as medicine and language, Dickey argues that the choices made by the user as the character is developed over time and through interaction with the storyline enhance intrinsic motivation to continue engaging in the MMORPG (2007, p. 258). This correlates with the findings of the previously noted researchers regarding self-determination, agency, autonomy and contextualization. Additionally, the emotional connection that some players appear to form with their character avatars (2007, p. 258) increases the personal relevance of their interactions: another potential mechanism for improving intrinsic motivation in MUVEs that feature role-playing and an ongoing narrative. As noted previously, these features are typical of situated learning environments where learners take on a role while immersed in a simulation.

Yee (2007) also studied the motivations underlying the use of online role-playing games. A factor analysis was used to derive three components and ten subcomponents of motivation in MMORPGs:

Table 2

The subcomponents revealed by the factor analysis grouped by component (Yee, 2007, p. 773)

Achievement	**Social**	**Immersion**
Advancement Progress, Power, Accumulation, Status	**Socializing** Casual Chat, Helping Others, Making Friends	**Discovery** Exploration, Lore, Finding Hidden Things
Mechanics Numbers, Optimization, Templating, Analysis	**Relationship** Personal, Self-Disclosure, Find and Give Support	**Role-Playing** Story Line, Character History, Roles, Fantasy
Competition Challenging Others, Provocation, Domination	**Teamwork** Collaboration, Groups, Group Achievements	**Customization** Appearances, Accessories, Style, Color Schemes
		Escapism Relax, Escape from Real Life, Avoid Real Life Problems

These appeared to generally agree with self-determination theory and the parts of Dickey's analysis most relevant to the current study: that achievement- and story-based design elements – those types most related to gamification – made up a significant proportion of the motivational affordances of this type of MUVE.

Satisfaction. Satisfaction with a learning experience is derived from the student's perception of their ability to achieve an acceptable outcome (Keller, 1983). It is also related to the degree of correlation between how learners recall their expectations prior to the experience and their perceptions after the experience (Appleton-Knapp & Krentler, 2006). Satisfaction is an important motivational component because satisfaction perceptions are predictors of future behaviors that, in turn, can affect learning outcomes. For example, students that are satisfied with their learning experiences are less likely to drop out of academic programs (B. L. Bailey, Bauman, & Lata, 1998; Levy, 2007).

Anxiety about computer use has a negative effect on learner satisfaction, and this is associated with perceptions of system usefulness and ease-of-use, which are also predictors of satisfaction (Sun, Tsai, Finger, Chen, & Yeh, 2008). Considering these perceptions are psychological drivers of behavior related to self-efficacy, it is not surprising that one of the most important factors related to satisfaction in online learning is self-efficacy. As well as positively predicting future engagement with further online learning, perceived computer self-efficacy has been demonstrated as a consistent predictor of satisfaction for learners using Web-based distance education (Liaw, 2008; Lim, 2001). These findings are highly applicable to the rationale for the current study: students who are satisfied with a MUVE familiarization course and gain a perception of high self-efficacy from its completion are more likely to persist with a subsequent MUVE learning program.

The ARCS Motivation Model (Keller, 1984) proposed four requirements to foster motivation for learning:

- **Attention:** Capturing learners' interest and stimulating the desire to learn.
- **Relevance:** Fulfilling learners' needs to promote a positive attitude.
- **Confidence:** Supporting learners' belief that they will succeed.
- **Satisfaction:** Reinforcing learners' achievements with rewards.

This model provides a useful framework in which satisfaction is positioned within academic motivation, and includes an exposition of the components of satisfaction and strategies for their implementation:

- **Natural consequences:** The learner has the opportunity to use their newly acquired knowledge or skills in "real-world" scenarios.
- **Positive consequences:** The learner receives reinforcing feedback such as verbal praise, symbolic awards or opportunities to exhibit their work.
- **Equity:** The learner is provided with clear performance and assessment expectations that remain consistent from the start to the end of the learning program and are applied to all students equitably.

(Keller, 1987, pp. 5–6)

The ARCS model has been used and validated in multiple studies since its inception in 1984 (see Chang & Lehman, 2002; Keller & Kopp, 1987; Means, Jonassen, & Dwyer, 1997; Small & Gluck, 1994; Visser & Keller, 1990). Keller later formulated recommendations for the application of the ARCS model to computer based instruction and distance learning, suggesting instructional design considerations for each of the four ARCS categories (Keller, 1999). More recently, Keller has acknowledged the necessity of including self-regulatory or *volitional* strategies to the motivational model and has provided expanded recommendations for contemporary instructional modalities such as blended and mobile learning (Keller, 2008).

Verhagen et al. (2011) created and tested a model specifically addressing satisfaction in virtual worlds. Based on expectancy-value theory (Porter & Lawler, 1968; Vroom, 1964), the model assumes the expectancy-value concept that "an individual will be more satisfied with an object or behavior if he or she perceives it as more likely that the object or behavior possesses value" (Verhagen et al., 2011, p. 202). Accordingly, the model contains four sources of "experiential system value," two intrinsic and two extrinsic:

- **Escapism** (intrinsic): The degree to which the user becomes so immersed in the virtual world it fulfills the desire to cognitively and emotionally escape reality.
- **Entertainment value** (intrinsic): The extent to which the user perceives the virtual world experience as fun and enjoyable.
- **Economic value** (extrinsic): The net gain perceived by the user after investing in use of the virtual world.
- **Ease of use** (extrinsic): The user's perception that learning to use the virtual world is uncomplicated and does not represent a barrier to access and enjoyment.

The authors hypothesized that all four sources of experiential system value would positively affect satisfaction, and further, that the two extrinsic sources (economic value and ease of use) would positively influence the two intrinsic sources (escapism and entertainment value).

Verhagen and colleagues conducted a study to test the hypotheses regarding the interaction of motivational sources and satisfaction for experienced users of *Second Life*. Participants ($N = 567$) completed a psychometrically validated questionnaire. Results suggest that each extrinsic source positively influences the intrinsic sources as hypothesized (with the exception of ease of use affecting escapism, which was not significant), and all four sources of experiential system value significantly and positively predict user satisfaction (2011, p. 205).

The measurement and assessment of learner satisfaction in MUVEs specifically has not been well established. However satisfaction in online learning was addressed by Wang (2003), who developed the e-learner satisfaction (ELS) instrument. The ELS instrument contains 17 items categorized into four factors: learner interface, learning community, content and personalization. Wang's ELS instrument was intended for use with text-based, asynchronous virtual learning environments such as Blackboard Learn.

Another study that produced a validated instrument for learner satisfaction in online learning programs was also aimed at Blackboard Learn users. The Domínguez et al. (2013) study measured user satisfaction within a gamified VLE course. It contained items that were somewhat similar to Wang's, but were aimed at the participants' perceived outcomes. The wording of Wang's ELS instrument focused more on perceptions about the system itself. Given that this review has established that satisfaction is a product of how well the learner's perceived outcomes match their memory of their expectations before the experience, the Domínguez et al. instrument appears more applicable for assessing satisfaction in a gamified MUVE experiment.

Synthesis of the Literature

The review of the literature showed that researchers have only recently begun to explore the didactic potential of multi-user virtual environments. Much of the initial research has understandably focused on the core concerns of education, such as pedagogical efficacy, instructional design, student access, and the role of social interaction in constructivist learning. However, there was scarce empirical evidence to guide the formulation of preparatory experiences for sustained MUVE-based learning.

MUVEs have potential as a convenient, engaging and effective option for distance learning and educational simulations, but can suffer from barriers-to-entry when inexperienced students encounter the unfamiliar paradigm and have difficulty moving, navigating and interacting within the virtual world. Rather than simply enrolling students into MUVE-based courses and expecting them to "sink or swim," pre-training students in the basic functions of the MUVE interface can mitigate negative orientation effects by improving their self-efficacy perceptions and enhancing persistence behavior *before* they enter the virtual classroom. Prior research has demonstrated that high self-efficacy – and the task persistence that results from it – also has a positive effect on academic outcomes, so pre-trained students are more likely to succeed once in the academic course after familiarization training. Furthermore, satisfaction has been correlated to overall persistence in learning programs; students who were satisfied with a MUVE-based experience were more likely continue with similar learning formats. Because of these motivational phenomena, any approach that might support effective pre-training, improve self-efficacy perceptions and increase satisfaction in MUVE-based learning is of interest to educators. Studies have already established that gamification can improve self-efficacy, increase persistence and enhance satisfaction in academic courses and have also provided some best practices for achieving these goals. However, the effects of gamification on familiarization pre-training courses have not been investigated. While the interpersonal competition component of gamification has had deleterious effects on synchronous, collaborative learning activities, the positive effects seen in gamification studies makes it a potentially powerful motivational tool for individualized, asynchronous activities such as initial familiarization. Hence the hypothesis that students who complete a gamified MUVE familiarization course would exhibit significantly greater self-efficacy, persistence, and virtual world task skill compared to those who complete a non-gamified course and would therefore be more prepared for MUVE-based learning experiences. It was further hypothesized that the students in the gamified familiarization course would report higher satisfaction than those in the non-gamified course.

Chapter 3: Method

Overview

The objective of the study was to answer the following research question:

> *RQ:* What are the effects of the gamification of familiarization training courses on the factors related to learning in virtual worlds?

This was achieved through testing the following hypotheses:

> Participants in a gamified virtual world familiarization course will exhibit significantly greater measurement for the following variables while practicing virtual world familiarization tasks compared to those who participate in a non-gamified course:
>
> H_A1: Self-efficacy
>
> H_A2: Persistence
>
> H_A3: Task skill
>
> H_A4: Learner satisfaction

A randomized, experimental study was performed by randomly assigning participants to one of two groups. One group completed a non-gamified virtual world familiarization course, and the other group completed a gamified virtual world familiarization course. Each participant session generated one each of the following instruments:

- Participant experience survey (completed by participant)
- Pre- and post-tests for virtual world self-efficacy (completed by participant)
- Persistence worksheet (completed by researcher)
- Task skill worksheet (completed by researcher)
- Post-treatment interview transcript (completed by researcher)

- A post-treatment learner satisfaction survey (completed by participant)

Self-Efficacy. Results of the pre- and post-tests for self-efficacy were quantitatively analyzed to measure self-efficacy as a study variable. As discussed previously, the use of domain-specific scales for measuring self-efficacy is consistent with the recommendations of Bandura (2006).

Persistence. Activity during both treatments was quantitatively analyzed to test persistence by determining the overall number of attempts at each section of the course. The use of this indicator is consistent with the literature mentioned previously (Multon et al., 1991; Sandelands et al., 1988).

Task Skill. Speed and accuracy/error rate during both treatments was used to ascertain participants' task skill scores, which were then quantitatively analyzed to measure task skill as a study variable.

Learner Satisfaction. Results of the satisfaction survey were quantitatively analyzed to determine levels of student satisfaction as a study variable.

See Table 13 for a summary of the study methodology.

Participants

The participants in this study were recruited from the student populations of a large state university and a medium-sized private liberal arts university, both in the Midwestern US. Email invitations to participate were sent to as many students as possible through the use of email lists.

A total of 44 participants completed the main study: 17 males and 27 females. Initial data collected included participants' experience with MUVEs and their self-efficacy perceptions with regard to *Second Life* (Table 3).

Prior experience was measured by surveying participants about their use of MUVEs, video games with MUVE-like components and the various forms of controls used in such systems (see Participant Experience Survey in Appendix A for specific questions). For each question, the level of experience corresponded to a weighted point scale, assigning a numeric value to the answer:

I am a current or recent (less than one year) user: 4 points

I have used it more than twice, but not in the past year: 3 points

I have tried it once or twice: 2 points

I have watched someone else use it: 1 point

I have no experience: 0 points

An overall score was derived from these data: more recent experience with systems similar to the *Second Life* MUVE equaled a higher overall experience score.

To ensure the two treatment groups had a similar distribution of participant experience, experience scores were used to stratify the random assignment of participants to one of two groups: the non-gamified group, who completed the course with no gamification elements; and the gamified group, who completed the course that included gamification elements. The stratified randomization was confirmed through post-hoc quantitative analysis (Tables 15 and 16).

Participants' self-efficacy perceptions for successfully using *Second Life* before the familiarization course were measured using a pre-test (questions are shown in Appendix A).

Table 3

Participant Characteristics

Characteristic		**By treatment group**	**Overall**
Gender	Male	A: 7 B: 10	17
	Female	A: 15 B: 12	27
Mean experience score (0-16 scale)		A: 7.5 B: 9.7	8.6
Mean self-efficacy score (0-1700 scale)		A: 1007.5 B: 923.1	965.3

A Priori Power Analysis

The minimum sample size for this study was determined by an *a priori* statistical power analysis using *G*Power 3* software (Faul, Erdfelder, Lang, & Buchner, 2007). Seven experimental studies investigating the effect of pre-training experiences on learning in game environments were examined by in a meta-analysis by Mayer, who derived a median effect size of $d = 0.77$ (Mayer, 2014, p. 145). This effect size was considered to be large using Cohen's (1992) criteria and approximates an effect size of $f = 0.4$ for a repeated measures, between subjects analysis of variance. If $\alpha = .05$ and power = 0.80, the projected sample size needed with this effect size was approximately $N = 39$ for the between group comparisons in this study. With a final sample size of 40, the study would have power of 81.2% to return a statistically significant result. The final sample size of 44 participants exceeded this *a priori* minimum.

Study Protocol

All participants completed:

- An IRB-approved informed consent form.
- An experience survey consisting of multiple choice questions to assess the participant's previous experience with virtual worlds.
- A pre- and post-test to assess their perceived self efficacy for various tasks involved in becoming familiar with the virtual world paradigm.
- A survey consisting of Likert-scaled responses to assess the level of satisfaction.

Participants in each group were informed that they were to explore the familiarization course alone and help from the researcher would only be available in the learning phases. Each participant's experience was recorded to enable subsequent collection of task skill scores and persistence data.

Timelines

Typical participant timeline.

Day 1:	Participant responds to recruitment materials
+1-3 days:	Participant notified of participation details
+1-7 days:	Participant completes informed consent and experience survey
+7-14 days:	Participant attends treatment session
+7 days:	Participant completes 1st online post-test for self-efficacy
+30 days:	Participant completes 2nd online post-test for self-efficacy

Treatments. Treatment sessions were conducted with one participant at a time and lasted approximately 1 hour:

1. Self-efficacy pre-test (5 minutes)

2. Session introduction/orientation (5 minutes)
3. Virtual world familiarization course (45 minutes)
4. Learner satisfaction survey (5 minutes)

Participants were asked to complete the online post-test for self-efficacy approximately one week after the treatment, and again one month after the treatment.

Materials Required

- Treatment computer (PC or Mac, no special requirements)
- *Second Life* client software
- Virtual world familiarization course with no gamification
- Gamified virtual world familiarization course (otherwise identical to above)
- Screen & audio capture software (i.e. Morae) or video camera and tripod

Study Environment

The familiarization course was designed using the backwards design method, which is highly suitable for learning systems as it allows the design to be driven by the desired educational outcomes (Wiggins & McTighe, 2005). Once designed, the course was custom-built using *Second Life*, a commercial virtual world that is extremely popular with educational institutions. A single virtual world familiarization course was created from the design, and then the sections were duplicated. One set of sections (for the non-gamified group) was left as-is. Gamification was operationalized in the other set of sections (for the gamified group) by adding challenges, clear goals, visual progress indicators, a time-scoring mechanism, a leaderboard, and achievement badges: all prevalent motivational affordances founded upon gamification rationales from the literature. Other common gamification elements such as points, levels and player ranks were impractical to implement in the constraints of the *Second Life* platform and available programming expertise. The inclusion of a story-based theme was not appropriate for the short pre-training context. Participants were informed that they would receive a certificate of virtual world learning proficiency on completion of the course. Non-gamified group certificates did not feature any badges or achievements except course completion. As part of gamification, participants in the gamified group were told that their achievement badges would be printed on their certificate. Badges were available in bronze, silver, gold or platinum depending on how quickly participants completed each section. Unbeknownst to participants, the platinum badges were unobtainable because the system would award only gold badges for the fastest times. Similarly to the Sandelands et al study (1988), this provided a definitive test of persistence as participants might repeatedly try to obtain the platinum badge. The incentive elements of the gamified sections and the performance required to earn them were consistent and equal across all participants in the gamified group.

Familiarization Course Instructional Design

Desired Results.

Broad Goals. Learners entering a MUVE-based learning environment for the first time often experience disorientation with the unfamiliar paradigm of being represented in a virtual space by an avatar. This instructional system design specified a familiarization course as hypothetical training for learners about to begin typical educational activities in MUVEs. The training course was to be used in an experimental study of the effects of gamification on self-efficacy, persistence and satisfaction when becoming familiar with virtual worlds. The course took place within the *Second Life* MUVE in a specially designed area featuring “skillset stations” corresponding to the learning objectives. Broad goals included fostering the following categories of skill:

- Spatial orientation within the MUVE.
- Directing avatar movement using system controls.
- Interacting with virtual objects.
- Communicating with other MUVE users through text message and voice.

The course was self-paced in that the participant spent as much time in a "Learning Phase" at each skillset station as necessary to feel proficient before deciding to proceed to the "Assessment Phase," then moved on to the Learning Phase of the next station. The system and the facilitator provided instructions for the participant to complete the tasks, and the facilitator conducted assessment on the participant's persistence and task skill. Instruments completed by the participant measured self-efficacy and satisfaction.

While the design aimed to familiarize participants with the MUVE paradigm, the primary purpose of the course was to provide a framework in which the study variables could be measured. Suitability for this purpose was evaluated during a pilot study, the results of which were used to inform a final instructional design for the full study.

Learning Objectives.

Table 4

Learning objectives of the familiarization course

Goal	**Learning objectives**
1. Movement and spatial orientation	Objective 1.1: Orientation and movement on the ground Upon completion of the practice course, the participant will demonstrate a sense of spatial orientation within the MUVE and the ability to move purposefully on the ground within virtual space. This will be demonstrated by walking and running the avatar forward and backward, sidestepping left and right, and turning left and right, as measured by following a ground movement course within a specified qualifying time while staying on a path that becomes progressively narrower.

	Objective 1.2: Orientation and movement through the air Upon completion of the practice course, the participant will demonstrate a sense of spatial orientation within the MUVE and the ability to move purposefully by flying within virtual space. This will be demonstrated through activation of the flying mode, flying the avatar forward and backward, and climbing and diving, as measured by executing an series of airborne maneuvers involving flying from the ground, through a series of ring-shaped obstacles, and back to the ground. The participant must successfully fly through all obstacles from the start point to the finish point within a specified qualifying time and without error for this objective to be met.
2. Object interaction	Objective 2.1: Manipulating virtual objects Upon completion of the practice course, the participant will be able to interact with virtual objects (select, interact with functions, pick up and release) as demonstrated by completion of a task involving selecting an object, using a functional menu to change its color, using touch functions to move it into a specified bin, then releasing it. These tasks must be completed within a specified qualifying time and without error for this objective to be met.
3. Communication	Objective 3.1: Communicating with other users Upon completion of the practice course, the participant will be able to communicate to other MUVE users by using text message as demonstrated by successful communication of three sample concepts to an automated listener.

Needs Assessment. The framework shown in Table 5 evaluates the information needs for the initial development of the familiarization course (Rossett, 1987). This was used to determine the aims and format of the pilot study and to help develop the study instruments.

Table 5

Rossett-based framework for needs assessment

Types of Information	What do you need to know?	Why do you need to know this?	Information Sources	Types of Procedures (Instruments)
Optimals	What are the goals for use of the course in the current study?	Alignment with study goals.	Primary investigator	Study design
	What variables need to be measured and using which metrics?	Incorporation of study metrics.	Primary investigator	Study design
	What factors prevent users from using MUVEs effectively?	Determination of training course effectiveness.	Primary investigator	Literature review
	What skills are to be developed by the course?	Instructional focus and scope.	Primary investigator	Study design
Actuals	In which MUVE is the course to be implemented?	Determination of supporting technology.	Primary investigator	Study design
	What experience do learners have with virtual environments in general?	To account for prior experience as a potential confounding variable.	Participants	Participant experience survey
	What support resources will participants need to complete the course?	Determination of support needed without compromising study goals.	Participants	Pre-pilot observational evaluation, pre-pilot participant interview
	How many participants are likely to use the course during the study?	Determination of study scale.	Primary investigator	Study design: *a priori* power analysis
Feelings	Do participants recognize a benefit to the study and their part in it?	Encourage complete participation.	Participants	Pre-pilot participant interview, pilot study participation
	Are participants willing to complete all parts of the study?	Encourage complete participation.	Participants	Pilot study participation

Causes	Does unfamiliarity affect learner motivation and/or outcomes?	Establish need for study	Primary investigator	Literature review
	Does unfamiliarity affect instructor workload?	Establish need for study	Primary investigator	Literature review
	Have familiarization courses been implemented in the past? If so, in what way?	Improve training/ Avoid replication of ineffective training.	Primary investigator	Literature review
Solutions	What form of familiarization course most effectively fosters positive motivational affect in participants and encourages persistence?	Design components of course to address familiarization needs identified from the literature and the MUVE paradigm.	Primary investigator	Data collection from gamified and non-gamified MUVE-based familiarization courses

Task Analysis. Task analyses were completed for the MUVE skills necessary for the goals of movement, spatial orientation, object interaction and communication. These analyses are shown in Appendix B.

Evaluation.

Formative Evaluation. It is important to identify any areas of the design that are not meeting the goals of the study, so a pilot study was specified for this purpose. The pilot study itself served as the formative evaluation for the development of the final familiarization course. The following questions were addressed by the pilot study:

Key Questions.

1. Does the familiarization course meet the primary objective of enabling measurement of the study variables?
    a. Self-efficacy
    b. Persistence
    c. Task skill
    d. Satisfaction

2. Does the familiarization course meet the secondary objective of cultivating the aspects associated with preparing participants to operate in a MUVE-based learning environment?
    a. Self-efficacy
    b. Persistence
    c. Task skill
    d. Satisfaction
3. Are the gamification elements properly designed and deployed to “make a game” of the appropriate course sections?
4. Are there areas the courses do not properly cover or address?
5. Are the courses organized and well structured?
6. Do the courses cover each of the four main learning goals?
7. Do the courses allow appropriate time to cover each of the four main learning goals?
8. What revisions to the pilot familiarization course, if any, are necessary before deployment in the main study?

Approaches to Gain Information. A pre-pilot group completed prototype versions of the familiarization course. Through direct observation, areas that suffered from a lack of clarity or undue difficulty were noted. This information was used to refine the course design and determine support resources. Upon completion of the observation, the observer conducted a brief interview with the participant to obtain their subjective perceptions of the course. The pre-pilot study interview protocol is shown in Appendix A.

Summative Evaluation.

Purpose. The summative evaluation of an educational system is usually intended to assess the effectiveness of the system for supporting learning. However, as this design is aimed at producing a system for an experimental study rather than an academic learning experience, the purpose of this evaluation is to provide a comprehensive review of how the final familiarization course provides effective data collection for the purposes of the study. After recommendations from the formative evaluation were implemented, the pilot group completed the course to provide data for the final course version to be used in the study.

Questions. The following questions were addressed during the summative evaluation:

1. What procedures and support materials are necessary to ensure that participants can complete the familiarization course?
   i. Advance organizers.
   ii. Behavioral guidelines.
   iii. Hardware, software and connectivity requirements.
   iv. Expectations.
2. Does the structure and content of the familiarization course provide participants with the information and situations necessary to complete the study objectives?
3. What skills and knowledge do participants acquire during the course, in terms of:
   i. Movement via walking, running, and flying
   ii. Spatial orientation
   iii. Manipulation of virtual objects
   iv. Communication?

4. Does the course provide adequate opportunity for the effective and accurate collection of data concerning the study variables?
5. How effective is the course in terms of fostering:
    i. Domain self-efficacy
    ii. Task persistence
    iii. Task skills mastery
    iv. Satisfaction?

Summative Methods.

Table 6

Matrix of Summative Questions and Methods

Methods	1. What procedures and support materials are necessary to ensure that participants can complete the course?	2. Does the structure and content of the familiarization course provide participants with the information and situations necessary to complete the study objectives??	3. What skills and knowledge do participants acquire during the course, in terms of movement, spatial orientation, manipulation of objects and communication.	4. Does the course provide adequate opportunity for the effective and accurate collection of data concerning the study variables?	5. How effective is the course in terms of fostering domain self-efficacy, task persistence, task skills mastery and satisfaction?
Observation	X	X	X	X	X
Completion of pilot study instruments		X		X	X

Learning Experiences.

Participant Analysis.

Participant variables were included in Table 7 in the form of questions essential to the instructional design of the familiarization course. The variables were aligned with the contextual analysis through categorization within the orienting, instructional and transfer contexts (Morrison, Ross, Kalman, & Kemp, 2012).

Table 7
Participant Analysis

Orienting Context	**Data Collection**
Participant Background: What prior experience might participants have with *Second Life* and how might that experience affect the study goals and outcomes?	Initial participant survey.
Are there other skills that learners already possess that might affect the study goals and outcomes (first-person video games, etc.)?	Initial participant survey.
Instructional Goals: What should participants know and be able to do after completing the course and what are the benefits?	Refer to study design.
Instructional Context	**Data Collection**
Participant Perceptions: How much support is needed for participants to be comfortable directing their own learning in the course?	Observation Participant satisfaction survey
How challenging will participants perceive the tasks to be?	Observation Participant satisfaction survey
What technological infrastructure will be necessary for the participant to complete the course?	Refer to study design.
Transfer Context	**Data Collection**
Perceived Applicability Are participants cognizant of the value of the course and how the skills learned will improve their abilities in future MUVE-based courses?	Participant satisfaction survey
What level of skill should participants achieve on completion of this course?	Refer to study design

Participant Assumptions.

1. Participants are adults that have provided informed consent.
2. Participants are prepared and able to follow and complete study protocol.
3. Participants have foundational skills in personal computer operation (use of abstracted interface control via keyboard and mouse).

Contextual Analysis.

Table 8

Contextual Analysis

Orienting Context	**Data Collection**
Immediate Environment Factors	
What technological equipment is required, and what is available for the study? What physical space is required, and what is available for study sessions? What is the duration of each study session for which equipment and testing space will be required?	Refer to study design Consult institutional sources
Organizational Factors	
What incentives are available to assist in recruitment of participants? What is the expected number of participants? When will participants complete the study treatments? What types of assessment-based feedback will participants receive, if any?	Refer to study design
Instructional Context	**Data Collection**
Immediate Environment Factors	
How many participants will complete the treatment at once? How many researchers will be needed to conduct the treatment sessions? What difficulties might participants experience in accessing the environment? What backup plans are necessary in case of equipment/software failure?	Refer to study design Observation
Organizational Factors	
What technical support may be needed for treatment sessions, and from what source will it be obtained?	Consult institutional sources
How many participants must be recruited before a set of treatment sessions is booked?	Refer to study design
Transfer Context	**Data Collection**
Immediate Environment Factors	
Is the course design and setting congruent with the type of virtual environments that participants will use in the future?	Literature review Consult institutional sources (what types of MUVE-based learning are currently offered?)

Organizational Factors	
(Optional) Could the course design be adapted and deployed for MUVE-based learning preparation beyond the study?	Consult institutional sources (is there a need for improved preparation?)

Types of Learning Experiences. The course was built within the *Second Life* MUVE, immersing the participant in the environment with which they are being familiarized. Participants completed the course in a controlled setting during study treatment sessions. The study facilitator was present and available to provide directions while the learner was engaged in the course. Although the course was self-paced, the facilitator was able to communicate with the participant according to the study protocol. The course consisted of four skillset sections (ground movement, flying, object interaction and communication), which were completed in one continuous session. Each section consisted of two phases. During the Learning Phase, participants self-paced through a practice course. When the participant felt they had mastered the section objectives, they moved on to the Assessment Phase where they demonstrated their skills pertaining to the section objectives. The following tables relate learning objectives to instructional strategies used in the course.

Table 9

Learning Goal 1: Spatial Orientation and Movement

	Behavioral Objectives	Type of Learning/Type of Learning Performance	Instructional Strategy	Rationale
1.1	Upon self-actuated completion of the Learning Phase of the training unit, the participant will begin the Assessment Phase and demonstrate the ability to move accurately in all directions on the ground, at both walking and running speeds.	Procedure/Application	**Learning phase:** Explanation (initially through virtual instruction signs, facilitator will answer direct questions), and practice. **Assessment phase:** Recall and demonstration	Participants will be able to effectively orient themselves within virtual space and move their avatar purposefully and accurately on the ground in the MUVE.
1.2	Upon self-actuated completion of the Learning Phase of the training unit, the participant will begin the Assessment Phase and demonstrate the ability to complete purposeful and accurate flying maneuvers.	Procedure/Application	**Learning phase:** Explanation (initially through virtual instruction signs, facilitator will answer direct questions), and practice. **Assessment phase:** Recall and demonstration	Participants will be able to effectively orient themselves within virtual space and move their avatar purposefully through the air in the MUVE.

Table 10

Learning Goal 2: Object Interaction.

	Behavioral Objectives	Type of Learning/Type of Learning Performance	Instructional Strategy	Rationale
2.1	Upon self-actuated completion of the Learning Phase of the training unit, the participant will begin the Assessment Phase and demonstrate the ability to manipulate virtual objects.	Procedure/Application	**Learning phase:** Explanation (initially through virtual instruction signs, facilitator will answer direct questions), and practice. **Assessment phase:** Recall and demonstration	Participants will be able to effectively interact with virtual objects located within the MUVE.

Table 11

Learning Goal 3: Communication

	Behavioral Objectives	Type of Learning/Type of Learning Performance	Instructional Strategy	Rationale
4.1	Upon self-actuated completion of the Learning Phase of the training unit, the participant will begin the Assessment Phase and demonstrate the ability to communicate with other users by using text messages and voice chat.	Procedure & Interpersonal/Application	**Learning phase:** Explanation (initially through virtual instruction signs, facilitator will answer direct questions), and practice. **Assessment phase:** Recall and demonstration	Participants will be able to effectively communicate with other users of the MUVE.

Description of Course Design.

Environment and equipment. The course was implemented in the *Second Life* virtual world. A parcel of virtual land measuring 4,096 square meters was leased from an in-world real estate agent as a space in which to build the course. *Second Life* provides simple building tools to create, modify and move virtual objects. These tools were used to build the course elements, such as raised pathways, partitions, and signs (see Figure 1). *Second Life* also features a scripting language known as LSL. LSL scripts can be attached to objects to provide programmatic functionality. This was used to produce interactive functions for course elements, such as timers, announcements and help guides.

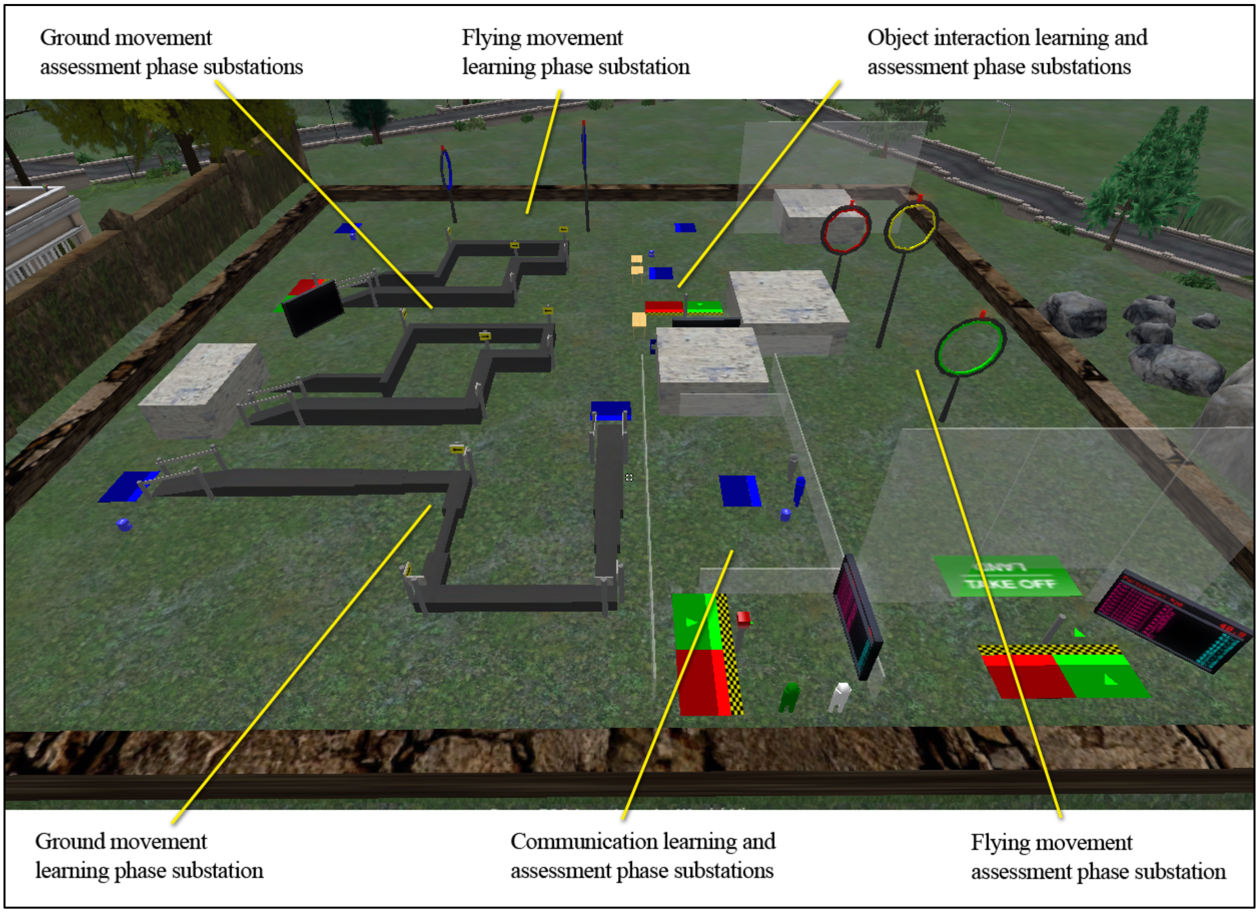

Figure 1. Overview screenshot of the familiarization course in Second Life

Second Life is accessed through client software running on the user's local computer. For the pilot study, the client software was installed on an Apple MacBook Pro running OS X version 10.8. To avoid problems caused by potential unfamiliarity with the Apple trackpad and to provide participants with a consistent interface, a Dell keyboard and mouse were attached to the computer. A video camera was used to capture the activities of the participant's avatar during the treatment session. The same equipment and arrangement were also used for the main study.

Ground movement station. In order for the participant to gain spatial orientation and the resultant ability to purposefully move the avatar on the ground, the course must provide a goal that requires the participant to use visual cues for orientation and the movement controls to achieve an objective. A raised pathway provides both a visual affordance and an incentive for effort. If the avatar falls off the path, the participant will be forced to start over from the beginning. Progressively narrowing the width of the path increases the level of challenge.

Learning phase substation. As seen in Figure 2., the learning phase substation starts as a relatively wide path and becomes narrower in the middle as a challenge to the participant's movement accuracy. Guide rails at the start of the path are a visual and physical affordance for initial alignment of the avatar on the path. To facilitate repetition, the learning phase substation was designed to be bidirectional. On completion in one direction, the participant can turn around and practice again in the other direction. The learning phase substation path is raised above the ground high enough to indicate when the avatar steps off the path, but low enough that the avatar can get back on without starting over. Corner signs at each change of direction provide a visual affordance to assist the participant judge the spatial location of the corner, as well as a physical barrier to help prevent walking off the path as it changes direction.

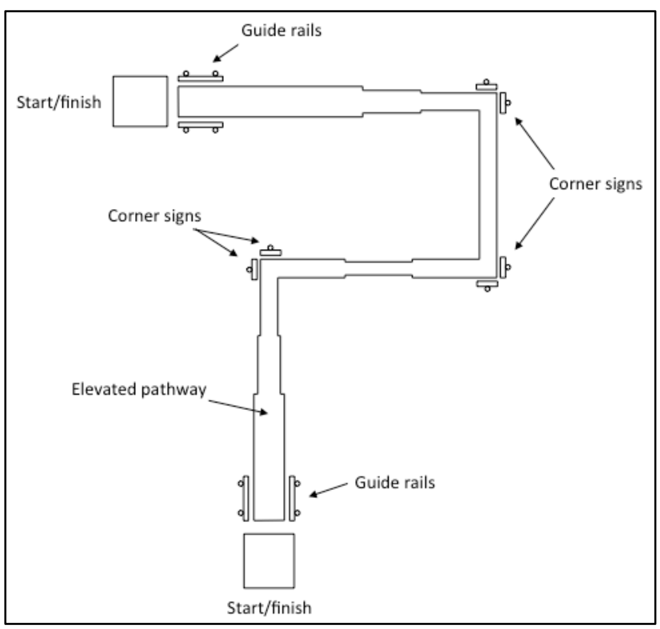

Figure 2: Plan view of ground movement learning phase substation

Assessment phase substation. The assessment phase substation is unidirectional, as each attempt must be the same as the last for consistent assessment of skills. As shown in Figure 3., the path starts wide, becomes progressively narrower, and has more turns than the learning phase substation. The final turn has no corner sign, making it more difficult to judge the corner. A small overrun allows for correctional action if the avatar walks too far. The assessment phase path is raised higher than the learning phase path, meaning the avatar cannot easily get back onto the path after falling off, and must start over. Falling off the path is recorded as an error, which is used when calculating the participant's task skill for this station. The last corner has no turn sign, making it more challenging than others where the sign helps to stop the avatar from walking off the corner. The substation has a start/finish line that programmatically starts and stops the timer when the avatar crosses it. An achievement badge indicator shows the highest achievement earned at this station (not present on non-gamified version).

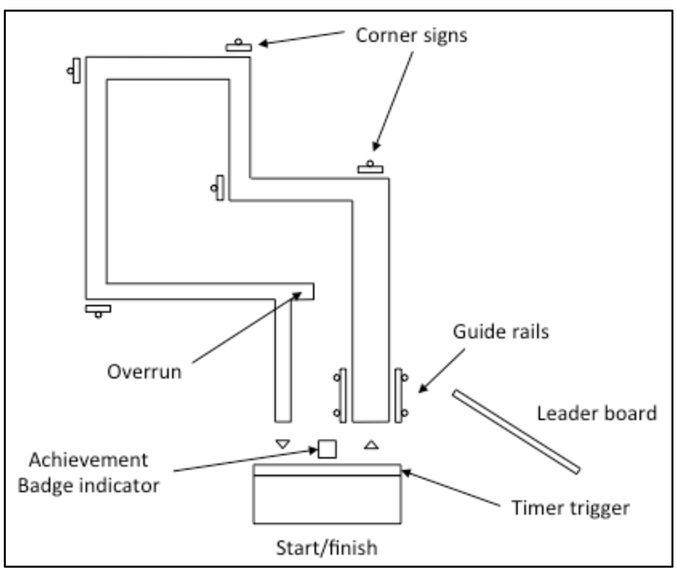

Figure 3: Plan view of ground movement assessment phase substation

Flying movement station. Flying the avatar requires movement in an additional dimension to ground movement. The participant must learn how to move the avatar up and down, as well as forward, back, left and right. This course station must therefore provide a goal in which all three dimensions of movement must be practiced and demonstrated. The concept of flying an avatar through a ring-shaped target has been used as an instructional method in many video games (*Grand Theft Auto: San Andreas*, 2004, *Superman: The New Adventures*, 1999, *Tachyon: The Fringe*, 2000) and would also serve this purpose well in the MUVE.

Learning phase substation. The learning phase substation consists of two large rings mounted at different heights on poles. The participant starts from a takeoff and landing point, attempts to fly the avatar through both rings, turn around, fly back through the rings, then land. Each ring is programmed to emit a chime when the avatar successfully flies through it, and a red light on top of the ring changes to green for ten seconds, further indicating success.

Assessment phase substation. The assessment phase substation is similar to the learning phase substation, but has three rings; a greater height discrepancy exists between rings and the second ring is significantly smaller in diameter, increasing the challenge for the participant (see Figure 4.). The participant must walk or run their avatar over the start/finish line to start the timer, then take off to fly through all three rings in order, then turn around and fly back through all three before landing and walking or running across the start/finish line. Once again, a chime and a green light on top of each ring indicate a successful transit. If the avatar moves past the ring without passing through, it is recorded as an error and the participant must attempt that ring again until successful. If the avatar touches the ground before flying through all three rings in both directions, it is recorded as an error.

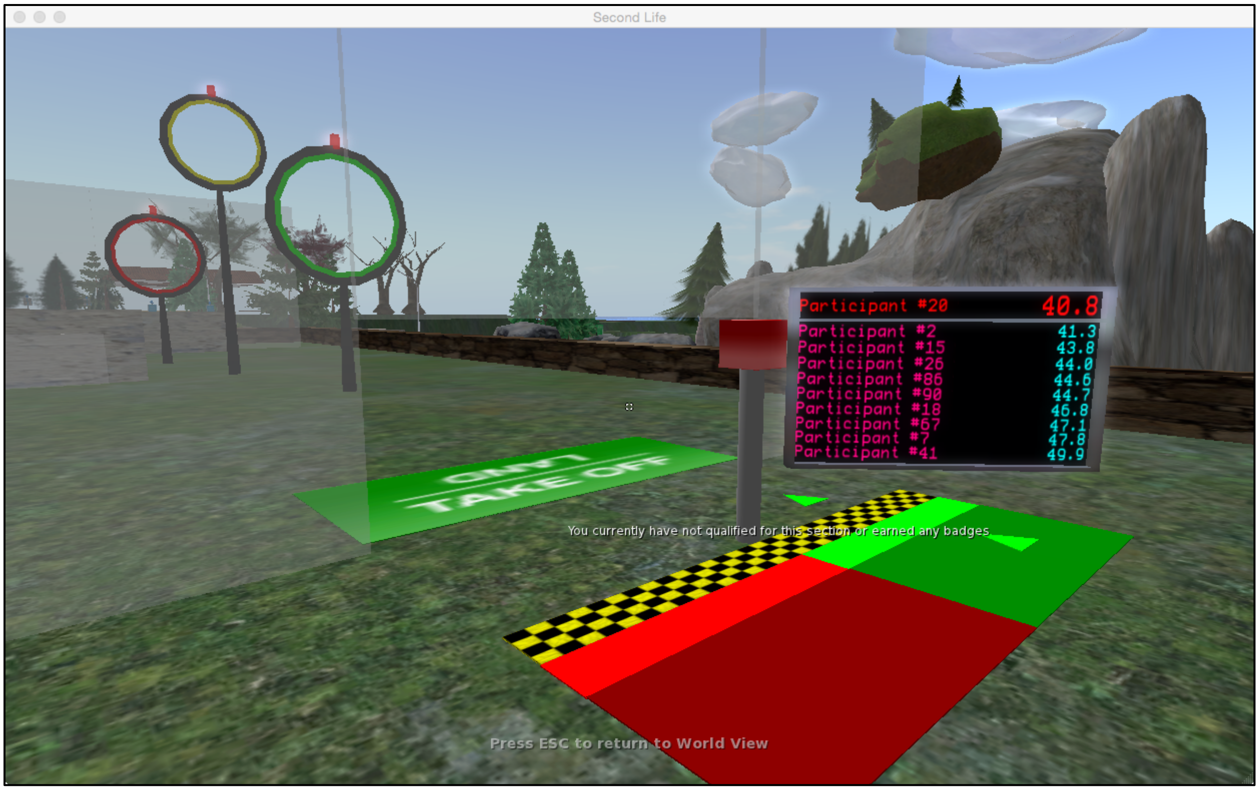

Figure 4: Screenshot of flying movement assessment phase substation

Object interaction station. Objects in *Second Life* have two modes of interaction: moving them around in space and menu-driven functions. For example, a desk lamp can be picked up and moved around a house, and it can be switched on and off. Objects are moved through a process analogous to the fantasy concept of telekinesis: in reality, the participant places the mouse cursor on the object in the client software window, left-clicks and holds while moving the mouse to move the object. In the *Second Life* virtual world, the avatar is seen to point at the object with an outstretched arm and finger, and appears to move the object through telekinesis. An object's function menu (if it has one) is accessed by the participant left-clicking the mouse pointer on the object. This causes the menu to appear on screen and the participant can then select item buttons in the menu. This station of the familiarization course must therefore provide opportunities for both types of interaction.

Learning phase substation. The learning phase substation presents the participant with two tables of differing heights. On one of the tables is a cube-shaped block measuring approximately ten inches to a side. The block's default color is yellow, making it highly visible. The participant is directed to practice moving the block from table to table without dropping it on the ground. The tables provide a spatial target for object movement; the differing heights require the participant to utilize all three dimensions of object movement while practicing. The participant is also directed to practice changing the color of the block through the use of the block's object menu, which provides six different color choices.

Assessment phase substation. The assessment phase substation, as shown in Figure 5., provides the participant with three yellow blocks on a table, with three rectangular bins on the ground to the left of the table. The bins are colored red, blue and green respectively. The participant is directed to walk into the assessment area (crossing the start/finish line and starting the timer), change the color of each block to either red, blue or green, then place each block into the corresponding bin. Dropping a block on the ground or into the wrong bin constitutes an error, which will be recorded as part of the task skill calculation. Once all three blocks are in the correct bins, the participant may walk the avatar back across the start/finish line and stop the timer.

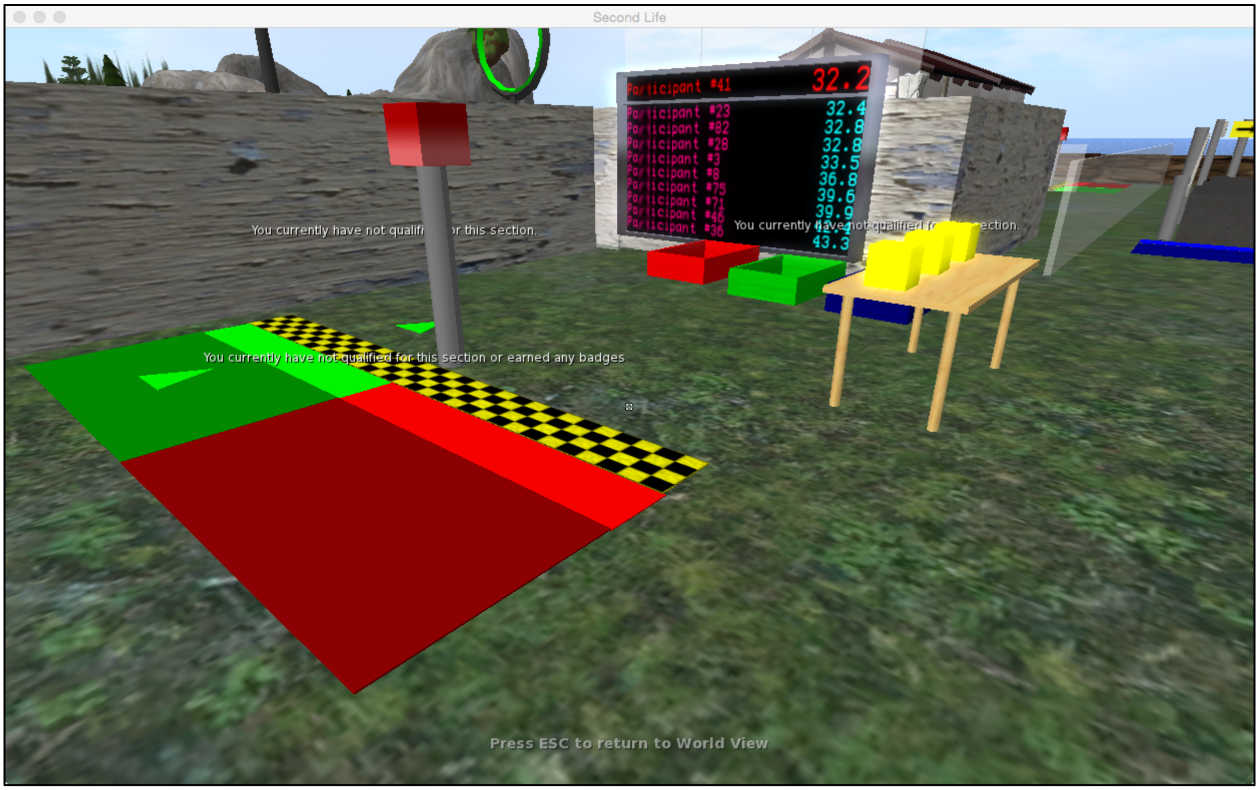

Figure 5: Screenshot of object interaction assessment phase substation

Communication station. There are two modes of communication in *Second Life*: audio chat and text chat. To audio chat, the user simply wears a headset and clicks a button to talk in real time to nearby avatars. The simplicity of this mode renders training unnecessary. Therefore the communication station focuses on the text chat function. This is where the participant opens a chat window (superimposed over the virtual world display) and types a text message. When the Enter key is pressed, the message is displayed to all avatars within a 20-meter radius. The station must provide an opportunity for the participant to learn and practice the steps needed to initiate and complete a text chat quickly and accurately with nearby avatars.

Learning phase substation. The learning phase substation consists of two color-coded robots that are programmed to act as substitutes for avatars. When clicked by the participant, the grey robot will use text chat to provide a phrase and ask the participant to repeat that phrase to the blue robot through text chat. The participant will then open the text chat window, type the phrase and "say" it in the virtual world. As the blue robot is within the 20-meter radius, it will "hear" the phrase and confirm whether the participant typed it correctly into the text window.

Assessment phase substation. This substation has a similar arrangement to the learning phase substation, with two color-coded robots that supply and respond to text chat phrases (see Figure 6.). The assessment phase substation has a start/finish line to start and stop the timer. The participant's avatar must cross the line, click the white robot for a phrase, and then repeat that phrase to the green robot. To demonstrate the skill level for this task, the participant is required to repeat three phrases correctly before crossing back over the start/finish line to stop the timer. If a phrase is repeated incorrectly, it is counted as an error and the participant must try again.

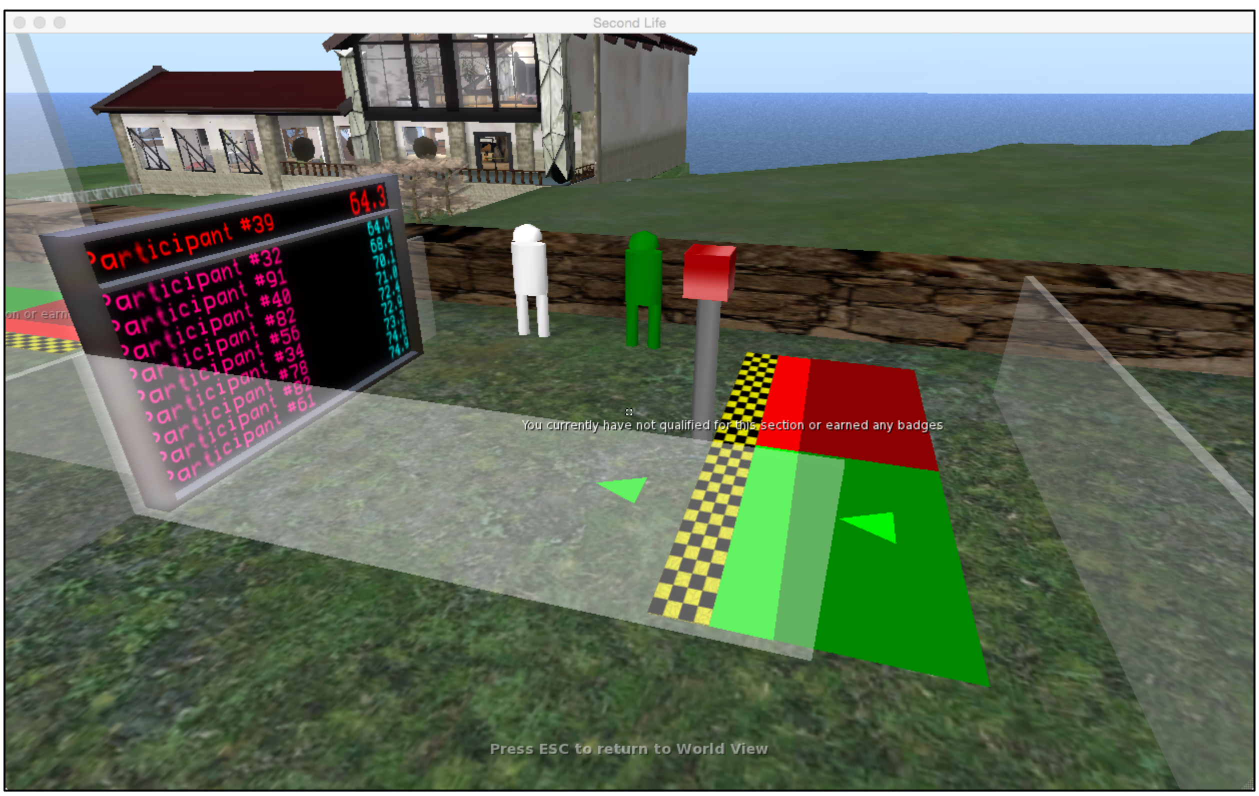

Figure 6: Screenshot of communication assessment phase substation

Gamification Elements

Timer. For every assessment subsection of the course, an automatic timer was programmed to start when the avatar crossed the checkered start/finish line, and stop when the avatar crossed back over the line at the end of the attempt. During the main study, participants in the non-gamified group had no visibility of the timer, but participants in the gamified group were shown their completion time as an on-screen display when they ended their attempt.

Leaderboard. The gamified assessment subsections each featured a large leaderboard showing ten "fastest times" of other participants (due to programming limitations, the fastest times were fictitious, although participants had no way of knowing this). The leaderboards were not present on the non-gamified assessment subsections, and hidden from participants when the course was in non-gamified mode. When a gamified group participant completed the subsection in a time faster than the slowest time on the board, it would update the board with the participant's time, next to the word "You." An on-screen message would tell the participant that they had qualified for a place on the leaderboard in *n*th place.

Qualifying. If a gamified group participant completed an assessment subsection faster than the qualifying threshold for that section, a green cube appeared on a post near the start/finish line with the word "Qualified" in white text. According to the information given to the participant during the pre-session briefing, this would mean that the skill would be checked off as Qualified on their certificate of completion. This functionality was not present on the non-gamified assessment subsections. To qualify, the participant's time had to be under:

- Ground movement: 55 sec
- Flying movement: 65 sec
- Object interaction: 44 sec
- Communication: 75 sec

Participants were not told the time thresholds for qualifying, even if they asked.

Achievement badges. If a gamified group participant completed an assessment subsection in an exceptionally fast time beyond the qualifying time threshold, an on-screen message would congratulate the participant and inform them that they had earned the badge. Additionally, a cube showing a graphic of the badge would appear near the start/finish line. According to the information given to the participant during the pre-session briefing, this event would result in the achievement badge being shown on their certificate of completion. Badges and the time threshold needed to earn them are shown in Table 12. Participants did not have any information on the time thresholds for each badge until the badge had been earned.

Table 12

Achievement badges

Subsection	Badge name	Bronze badge	Silver badge	Gold badge
Ground movement	Prime Mover	 50 sec	 45 sec	 40 sec
Flying movement	Master Aviator	 60 sec	 55 sec	 50 sec
Object interaction	Block Buster	 38 sec	 32 sec	 28 sec
Communication	True Talker	 70 sec	 65 sec	 60 sec

Measures

Self efficacy. Self-efficacy is considered highly domain specific (Schunk et al., 2008, p. 139). For increased accuracy when measuring task efficacy, general self-efficacy scales should be avoided in favor of purposefully targeted scales (Bandura, 2006; Pajares, 1996), so a domain-specific task self-efficacy data collection instrument (shown in Appendix A) was developed using recommendations from *Guide for Constructing Self-Efficacy Scales* (Bandura, 2006). Participants were asked 17 questions about their confidence in their ability to successfully perform a task in *Second Life* on a scale from 0 (least certain) to 100 (most certain). Therefore, the pre- and post-test measures of overall domain self-efficacy for each participant were numeric values in the range 0-1700, with higher values reflecting higher perceived overall self-efficacy for successfully performing tasks in *Second Life*. The self-efficacy instrument was validated prior to the main study by using data from the pilot study (see Pilot Study section).

Persistence. The number of attempts for each learning phase subsection and each assessment phase subsection was obtained from recordings of each participant's session. An "attempt" was defined as the participant's avatar intentionally crossing the start line for the subsection and was counted whether or not the participant was successful in completing the attempt. The persistence worksheet (Appendix A) was used to record participant persistence behavior during the study session. Skillset station persistence was the total number of assessment phase attempts for the skillset station. Overall persistence was measured by adding the total number of attempts within all assessment phases.

Task skill. A task skill worksheet (Appendix A) allowed the facilitator to record the participant's behavior and attainments during the study session. These data were then used to calculate a task skill score for each assessment phase. The task skill score was derived by first calculating the reciprocal of the participant's fastest time in seconds (this caused faster times to equal a higher score). The reciprocal (a decimal between 0 and 1) was then multiplied by 1,000 to produce a positive number within the likely range 10 to 40 (based on typical fastest times). The participant's average number of errors per assessment phase attempt was calculated by dividing the number of assessment phase attempts by the total number of errors during the assessment phase. The average number of errors per assessment phase attempt was converted to a percentage value equivalent to one error per average attempt equaling 10% of the multiplied fastest time reciprocal. This error percentage value was then deducted from the multiplied fastest time reciprocal as a penalty to give the final task skill score. With each average error-per-attempt reducing the final task skill score by 10% of the raw score, this gave a higher task skill score for faster times, and lowered the score in response to a higher average number of errors per assessment phase attempt.

Example: A participant attempts the flying movement assessment phase three times, makes six errors in total and has a fastest time of 90 seconds. The raw task skill score is 11.11 (reciprocal of 90 seconds multiplied by 1000). The average errors-per-attempt is 2, which imposes a penalty of 2.22 (10% of average errors-per-attempt multiplied by the raw task skill score), giving an error-weighted task skill score of 8.9.

Satisfaction. Participation satisfaction was measured through the use of a Likert-scaled survey (shown in Appendix A), administered immediately following the study session. The learner satisfaction survey contained items adapted from a related study where validity and reliability was established and reported. Survey items were adapted with minimal changes of wording from Domínguez, et al (2013, p. 390). The survey contained ten statements related to satisfaction, to which participants could answer 1 (strongly disagree) to 5 (strongly agree). Each participant's responses were totaled to derive an overall satisfaction score. Therefore, the minimum satisfaction score was 10 and the maximum score was 50.

Pre-Pilot Usability Evaluation

To ensure the nascent course design was sufficiently usable, a pre-pilot group completed prototype versions of the familiarization course. The pre-pilot group consisted of a convenience sample of four participants. Prototype course protocols and instructions were used and participants' behaviors observed. Areas of unintended difficulty and course components that lacked clarity were noted and the participants were interviewed about their subjective perceptions regarding the course. This information was used to iteratively refine the course design prior to the pilot study. The pre-pilot instruments are shown in Appendix A.

Pilot Study

To evaluate the instructional design of the familiarization course and the feasibility of the main study, a pilot study was conducted using a convenience sample of 16 participants. The participants were randomly assigned via coin toss to one of two groups, A or B. The non-gamified group completed the non-gamified substations of the familiarization course and the gamified group completed the gamified substations. Two participants did not complete all parts of the pilot study, leading to a final $N = 14$ with 7 participants in each group.

Pilot study session protocol. Each participant met with the facilitator in a quiet, private room in which the treatment computer system was located. The participant was first asked to complete the participant experience survey and self-efficacy pre-test. These were combined into a single online form titled "Initial Participant Survey" on the treatment computer. After the initial instruments were completed, the facilitator read an explanatory statement specific to the treatment group to which the participant was assigned and then asked if the participant had any questions. The participant was then shown the *Second Life* software client with the system already logged in and ready for the first learning phase substation. The facilitator guided the participant through each skill station in turn, moving on only when the participant indicated a desire to do so. When all stations had been completed, the facilitator explained that there would be two online surveys to complete: one in a week's time, and another one month after that. The full protocols are shown in Appendix A.

Validation of self-efficacy pre- and post-test instrument. The self-efficacy instrument was administered in the pilot study to validate it before use in the main study:

Response variability. Items with low response variability were unlikely to add value and so should have been revised or discarded (Bandura, 2006). To determine response variability, the standard deviation across all participants was calculated for each self-efficacy item. A standard deviation below 12.97 would signify that all participant responses for that item were scored within 25% of one another and thus indicate that the item adds little value to the instrument. An analysis of pilot participant responses for both the pre- and post-tests resulted in no items meeting the criterion for low response variability. This indicated that all items had value in determining self-efficacy perceptions. The results of the response variability analysis were included in Appendix C.

Predictive validity. The predictive validity was checked by surveying pilot study participants' prior experience with MUVEs. Experienced MUVE users would be expected to score highly. A correlational analysis indicated that a moderate, positive, and significant ($p < .05$) relationship existed between the participant's experience with similar virtual world systems and domain-specific self-efficacy perceptions prior to the study, $r = .582$, $p = .029$. A simple linear regression analysis was used to regress self-efficacy on participant experience. As indicated by the correlational analysis, experience was a significant predictor of self efficacy, $t(12) = 2.48$, $p = .029$, accounting for approximately 28% of the variance, $R^2_{adj} = .283$.

Internal consistency. Internal consistency was determined using an SPSS reliability analysis, returning Cronbach's alpha $\alpha = .98$ for all 17 self-efficacy items. This was considered an indication of "excellent" internal consistency (George & Mallery, 2003).

Within-subjects reliability. The self-efficacy post-test was administered to participants one week after the pilot study session. After a one-month period, the participant retook the post-test to determine within-subjects reliability. A correlational analysis indicated that a very strong, positive, and significant relationship exists between the participant's first and second post-test scores, $r = .963$, $p < .001$. This provided evidence that the test items were reliable measures of self-efficacy perceptions over time.

Pilot study data analyses. The principal purpose of the pilot study was as a proof-of-concept, intended to test the design, methodology and instruments before undertaking the main study. In-depth data analysis was not an intentional product. However, in the interests of completeness, the results of the pilot study analyses are summarized in Table 13.

Table 13

Summary of Pilot Study Results by Investigative Focus and Gamification Group

Focus	**Gamified**	**Mean**	**Analysis**	**Statistic**
Self-efficacy	N	1453.57	ANCOVA	$F(1, 11) = .184$
(post-test scores)	Y	1424.14		
Persistence	N	11.71	ANOVA	$F(1, 12) = .550$
(total assess. attempts)	Y	13.86		
Task skill	N	70.92	ANOVA	$F(1, 12) = .015$
(task skill scores)	Y	72.81		
Learner satisfaction	N	44.71	ANOVA	$F(1, 12) = .010$
(satisfaction scores)	Y	45.43		

Note: * $p < 0.05$, ** $p < 0.01$

In addition to the small sample size of the pilot study, a flaw in the pilot course design was potentially responsible for the lack of significant differences between the two groups. This issue was addressed below.

Post-Pilot Evaluation of gamification structure. As participants were observed completing the pilot study sessions, it became increasingly clear that the presence of a timed qualification objective in the non-gamified substations was having an unwanted gamification effect. From observed behavior and verbalizations, some participants assigned to the non-gamified group appeared to consider it important that they achieve a qualifying time before moving on. By unintentionally "making a game" of the non-gamified substations, the pilot study design seems to have confounded the principal investigative point of the study. Therefore, all timed qualification goals were removed from the non-gamified substations during the main study, with the goal of more accurately addressing the research question.

Revisions to course design.

The following design revisions were identified as necessary through observations made during the pilot study:

- Removal of qualifying functionality from Treatment A assessment phases.
- Removal any reference to qualifying or competition from Treatment A protocol.
- Reprogramming of timers to be invisible to participants in Treatment A.
- Refining of qualifying times and achievement badge times in Treatment B.
- Simplification of robot correct/incorrect responses in communication subsection.

Revisions to methodology.

Observations from the pilot study guided revisions to the methodology to improve the main study:

Randomization of group assignment. The main study used stratified randomization to assign participants to groups. This produced equal-sized groups and helped balance and control for experience with MUVEs and similar control systems (such as first-person video games), reducing the potential confounding influence of prior experience (Kang, Ragan, & Park, 2008). The stratified randomization procedure consisted of:

- Creation of blocks for the participant experience score.
- Assignation of participants to blocks based on experience qualification.
- Simple random assignment to groups within each block.

Treatment protocols. The treatment session protocols were revised (see Appendix A) to reflect the design changes to Treatment A, where participants no longer qualified or had knowledge of their attempt timings. This was prompted by the unintentional gamification effect observed during the pilot study.

Study Instruments

The main study utilized the same instruments developed for the pilot study, whose validity and reliability were determined during the pilot:

- Pre- and post test for virtual world self-efficacy.
- Persistence worksheet.
- Task skill worksheet.
- Learner satisfaction survey.

A qualitative interview was added to the main study with the aim of determining the underlying motivations of observed behaviors. All instruments are shown in Appendix A.

Table 14

Summary of Methodology

Investigative Focus	**Data collection**	**Measures**	**Primary analysis**
Self-efficacy	Pre- and post self-efficacy tests	Pre- and post-test scores	One-way, between-subjects ANCOVA
Persistence	Screen capture + audio recording	Number of assessment attempts, treatment time	One-way, between-subjects ANOVA
Task skill	Screen capture + audio recording	Skill scores	One-way, between-subjects ANOVA
Learner satisfaction	Learner satisfaction survey	Satisfaction scores	One-way, between-subjects ANOVA
Behavior/motivation	Post-session interview	Verbal responses	Grounded theory coding

Chapter 4: Results

This chapter contains the results of the main study and the interpretation of the results as they relate to the research question and hypotheses. The first section presents the results of the quantitative analyses of data gathered before, during and after the treatment sessions. The second section contains the results of the qualitative analysis of the post-session participant interviews.

Quantitative Analyses

Experience. The degree of experience that a participant had with Second Life, similar MUVEs, MUVE-based learning and control mechanisms could have confounded the study if one group had significantly more experience than the other. To confirm that the stratified randomization of group assignment led to an even distribution of experience between groups, a one-way, between-subjects Analysis of Variance was conducted on the participants' total experience score (TOTEXP) for their group membership. The total experience data passed the Shapiro-Wilk test for normality (S-W = .957, df = 22, p = .425 for the non-gamified group; S-W = .959, df = 22, p = .473 for the gamified group). No violation of the homogeneity of variances assumption was detected in a Levene's test, F = .154, p = .696. The result of the analysis was that there was no significant difference in experience between the treatment groups, $F(1, 42) = 3.43$, p = .071.

Table 15
Descriptive Statistics for Participant Total Experience

Variable	**Gamified**	***N***	***M***	***SD***	**Std. Error**	**Minimum**	**Maximum**
TOTEXP	N	22	7.455	4.03	.86	0	14
	Y	22	9.727	4.11	.88	0	16
Total		44	8.591	4.18	.63	0	16

Table 16

ANOVA: Experience Scores (TOTEXP) by Gamification Group

	Sum of Squares	***df***	**Mean square**	***F***
Between-Groups	56.82	1	56.82	3.43
Within-Groups	695.82	42	16.57	
Total	752.64	43		

Note. * $p < 0.05$, ** $p < 0.01$

Self-efficacy. To determine whether the self-efficacy post-test means, adjusted for initial virtual world self-efficacy, differed significantly between the two treatment groups, a one-way, between-subjects Analysis of Covariance was planned on the participants' post-test self-efficacy scores (SEPOST1) for their group membership and controlling for pre-test scores (SEPRE) as a covariate. However, during preliminary analysis, a Shapiro-Wilk test for normality showed that while the gamified group self-efficacy post-test scores were normally distributed (S-W = .935, df = 22, p = .158), the non-gamified scores were negatively skewed and not normally distributed (S-W = .854, df = 22, p = .004). Therefore, the self-efficacy post-test scores were transformed using a reflected square root transformation, yielding a subsequent Shapiro-Wilk test result indicating normal distribution in both groups (S-W = .974, df = 22, p = .791 for the non-gamified group; S-W = .958, df = 22, p = .45 for the gamified group). A Levene's test detected no violation of the homogeneity of variances assumption, F = .367, p = .548.

While it is unnecessary to transform a covariate to achieve a normal distribution (there is no assumption of normality for covariates in an ANCOVA), the pre-test covariate data (SEPRE) was transformed with the same reflected square root method to allow accurate testing of the homogeneity of regression coefficients assumption after the transformation of SEPOST1.

Further preliminary analysis on the transformed data showed that the correlation between the pre-test covariate and the post-test for self-efficacy was positive, linear and weak, $r = .26$, meaning the inclusion of the covariate in the model was not statistically redundant, but was also not likely to greatly affect the outcome. The covariate was nonetheless included in the model for increased precision, even though the gains may not be substantial. The homogeneity of variance assumption for this model was tested and was found to be not violated, $F = .604$, $p = .442$. The homogeneity of regression coefficients assumption between the transformed DV and transformed covariate was also not violated, $F = 2.33$, $p = .135$, so interaction terms were not included in the final analysis model. The use of ANCOVA compensated for differences in pre-test means between the two groups by adjusting means for the covariate in the analysis model, essentially controlling for the (admittedly minor) effects of the covariate.

The result of the ANCOVA analysis was that the gamified group exhibited significantly lower self-efficacy post-test scores than the non-gamified group, $F(1, 41) = 7.53$, $p = .009$. The effect size of $\eta_p^2 = .155$ ($f = .43$) was considered a large effect according to Cohen's widely-accepted operational definitions (1992, p. 157). Therefore, the null hypothesis for self-efficacy, $H_0 1$ – that participants practicing familiarization tasks on a gamified course will exhibit no difference in measurement of self-efficacy – was rejected. The result was the inverse of the alternative hypothesis $H_A 1$: that the gamified group would test higher for self-efficacy; these results suggested that the gamified group had lower post-test self-efficacy than the non-gamified group.

Table 17

Descriptive Statistics for Self-Efficacy Pre-Test Scores

Variable	**Gamified**	***N***	***M***	**SD**	**SE**	**Minimum**	**Maximum**
SEPRE	N	22	1007.50	404.31	86.20	150	1490
	Y	22	923.14	440.86	93.99	40	1680
Total		44	965.32	420.21	63.35	40	1680

Table 18

Descriptive Statistics for Self-Efficacy Post-Test Scores

Variable	**Gamified**	***N***	***M***	**SD**	**SE**	**Minimum**	**Maximum**
SEPOST1	N	22	1458.09	206.58	44.04	860	1695
	Y	22	1287.41	237.52	50.64	810	1610
Total		44	1372.75	236.32	35.63	810	1695

Table 19

ANCOVA: Self-Efficacy Post-Test Scores by Gamification, With Pre-Test Scores as Covariate (transformed data)

Source	**Sum of Squares**	***df***	**Mean square**	***F***	**Partial Eta Squared**
SE Pre-test	109.57	2	109.57	2.88	.066
Gamification	286.79	1	286.79	7.53**	.155
Error	1562.57	41	38.11		

*Note. * p < 0.05, ** p < 0.01*

Persistence. Analyses were performed on participants’ persistence scores for their group membership. Separate analyses were conducted for overall persistence behavior (total number of learning attempts, TOTPRAC and assessment attempts, TOTASSESS), and for each of the skillset stations (MOVASSESS, FLYASSESS, OBJASSESS and COMASSESS). As a secondary measure of persistence, the total treatment time (TOTTIME) was also analyzed using a one-way, between-groups Analysis of Variance.

Persistence: Total learning attempts. The total learning phase attempts data measured the total number of times the participant attempted the practice subsections during the learning phases. On preliminary analysis, the data for TOTPRAC failed the Shapiro-Wilk test (assuming a significance threshold of $p < .05$) for normality (S-W = .894, df = 22, $p = .023$ for the non-gamified group; S-W = .884, df = 22, $p = .014$ for the gamified group). The data were normalized using a square root transformation. Once transformed, the data passed the normality test and a Levene's test detected no violation of the homogeneity of variances assumption, $F = .279$, $p = .600$. An Analysis of Variance showed that, with a back-transformed mean of 10.76 total learning attempts for the non-gamified group and 14.61 attempts for the gamified group, there was a significant difference between the treatment groups for this dimension, $F(1, 42) = 5.47$, $p = .024$. The effect size, $\eta^2 = .115$ was large. Therefore the null hypothesis for overall persistence, $H_0 2$ – that participants practicing familiarization tasks on a gamified course will exhibit no difference in measurement of persistence – was rejected using total practice attempts as a metric. The gamified group exhibited significantly greater persistence than the non-gamified group.

Table 20

Descriptive Statistics for Total Learning Attempts (untransformed)

Variable	**Gamified**	***N***	***M***	***SD***	**Std. Error**	**Minimum**	**Maximum**
TOTPRAC	N	22	11.27	5.19	1.11	5.0	25.0
	Y	22	15.23	6.36	1.36	7.0	26.0
Total		44	13.25	6.08	.916	5.0	26.0

Table 21

ANOVA: Total Learning Attempts (TOTPRAC_Sqrt) by Gamification Group

	Sum of Squares	*df*	Mean square	*F*
Between-Groups	3.23	1	3.23	5.47*
Within-Groups	24.84	42	.591	
Total	28.08	43		

Note. * $p < 0.05$, ** $p < 0.01$

Persistence: Total assessment attempts. This variable measured the total number of attempts the participant made during the assessment phases. The total assessment attempts persistence data (TOTASSESS) was highly positively skewed and failed the Shapiro-Wilk test for normality (S-W = .781, df = 22, $p < .001$ for the non-gamified group; S-W = .789, df = 22, $p < .001$ for the gamified group). Also, a violation of the homogeneity of variances assumption was detected, $F = 14.23$, $p = .001$. Transformation of the data failed to normalize the distribution, so a Mann-Whitney U test was conducted, as this non-parametric test for dichotomous groups is tolerant of non-normalized distributions (it is worth noting that the Mann-Whitney U test has less statistical power than the parametric alternatives and so there is an increased chance of a Type II error). The gamified group showed significantly higher persistence (Mdn = 12 total attempts) than the non-gamified group for this dimension (Mdn = 5 total attempts). The mean ranks of the non-gamified group and the gamified group were 11.9 and 33.1 respectively, $U = 9$, $Z = -5.52$, $p < .001$. Effect size for the Mann-Whitney U test was calculated as $r = Z$ / sqrt N (Rosenthal, 1991, as cited in Field, 2013, p. 227). Effect size was large at $r = -.83$, or $\eta^2 = .69$. Therefore the null hypothesis for overall persistence, $H_0 2$ – that participants practicing familiarization tasks on a gamified course will exhibit no difference in measurement of persistence – was rejected using number of attempts as a primary metric.

Table 22

Descriptive Statistics for Total Attempts

Variable	Gamified	*N*	*M*	*SD*	Std. Error	Minimum	Maximum
TOTASSESS	N	22	5.23	1.38	.29	4	9
	Y	22	15.0	7.82	1.67	6	43
Total		44	10.11	7.43	1.12	4	43

Table 23

Mann-Whitney U: Total Assessment Attempts (TOTASSESS) by Gamification

Group	Mean rank	Sum of ranks	*U*	*Z*	Sig.
Non-gamif.	11.91	262.00	9	-5.52	.000
Gamified	33.09	728.00			

Persistence: Total treatment time. The total treatment time (TOTTIME) passed tests for normality. No violation of the homogeneity of variances assumption was detected, $F = .389$, $p = .536$. An Analysis of Variance showed that, with mean total treatment times of 2,176.4 seconds for the non-gamified group and 3,056.4 seconds for the gamified group, there was a significant difference between the treatment groups for this dimension, $F(1, 42) = 20.4$, $p < .001$. The effect size, $\eta^2 = .327$ was large. Therefore the null hypothesis for overall persistence, $H_0 2$ – that participants practicing familiarization tasks on a gamified course will exhibit no difference in measurement of persistence – was rejected using total treatment time as a secondary metric. The gamified group exhibited significantly greater persistence than the non-gamified group.

Skillset station attempts. Following the analysis for total attempts and total treatment time, the total assessment attempt data obtained for the individual skillset stations were analyzed. The distribution of the non-gamified group data was skewed due to the consistently low number of attempts for the non-gamified group. Transformation did not normalize any of the distributions. Therefore, each skillset station analysis required a Mann-Whitney test (Table 24).

Table 24

Persistence Results (Total Assessment Attempts) by Skillset Station and Gamification Group

Station	Gamified	Median	Mean rank	*U*	*Z*	Effect Size
Ground movement	N	2.0	14.1	57	-4.44**	$\eta^2 = .45$
	Y	3.5	30.9			
Flying movement	N	1.0	13.8	51	-4.85**	$\eta^2 = .53$
	Y	2.0	31.2			
Object interaction	N	1.0	12.9	30	-5.34**	$\eta^2 = .66$
	Y	3.0	32.1			
Commu-nication	N	1.0	13.5	44	-5.06**	$\eta^2 = .58$
	Y	2.5	31.5			

*Note: * $p < 0.05$, ** $p < 0.01$*

Task skill. Analyses were performed on participants' task skill scores for their group membership. Separate analyses were conducted for overall task skill (total of all task skill scores, TOTSKIL), and for each of the skillset stations (MOVSKILL, FLYSKILL, OBJSKILL and COMSKILL).

Total skill scores. The total task skill data (TOTSKIL) passed tests for normality. No violation of the homogeneity of variances assumption was detected, $F = 1.09$, $p = .303$. The result of the Analysis of Variance was that, with mean total task skill scores of 36.0 for the non-gamified group and 61.2 for the gamified group, there was a significant difference between the treatment groups for this dimension, $F(1, 42) = 33.79$, $p < .001$. The effect size, $\eta^2 = .446$ was large (Cohen, 1992). Therefore the null hypothesis for overall task skill, H_03 – that participants practicing familiarization tasks on a gamified course will exhibit no difference in measurement of task skill – was rejected.

Table 25

Descriptive Statistics for Total Task Skill Scores

Variable	**Gamified**	***N***	***M***	***SD***	**Std. Error**	**Minimum**	**Maximum**
TOTSKIL	N	22	35.97	11.79	2.51	18.9	65.1
	Y	22	61.18	16.58	3.54	30.3	106.4
Total		44	48.57	19.10	2.88	18.9	106.4

Table 26

ANOVA: Task Skill Scores (TOTSKIL) by Gamification Group

	Sum of Squares	***df***	**Mean square**	***F***
Between-Groups	6993.53	1	6993.53	33.79**
Within-Groups	8693.43	42	206.99	
Total	15686.96	43		

Note. * $p < 0.05$, ** $p < 0.01$

Skillset station task skill scores. Following the analysis for total task skill, the data obtained for the individual skillset stations were analyzed. Table 27 shows the summarized results.

Table 27

Task Skill Results (Scores) by Skillset Station and Gamification Group

Station	Gamified	Transformation	Mean	Analysis	Statistic	Effect Size
Ground movement	N	None	8.7	ANOVA	$F = 8.07^{**}$	$\eta^2 = .16$
	Y		13.4			
Flying movement	N	Square root	6.2	ANOVA	$F = 2.43$	
	Y		8.0			
Object interaction	N	None	13.4#	Mann-Whitney	$Z = 4.70^{**}$	$\eta^2 = .50$
	Y		31.6#			
Commu-nication	N	Square root	9.3	ANOVA	$F = 25.1^{**}$	$\eta^2 = .37$
	Y		14.4			

*Note: # Mean rank, * $p < 0.05$, ** $p < 0.01$*

Satisfaction. A one-way, between-subjects Analysis of Variance (ANOVA) was performed on participants' learner satisfaction ratings (SATIS) for their group membership. The analysis passed tests for normality. No violation of the homogeneity of variances assumption was detected, $F = 1.05$, $p = .311$. With mean satisfaction scores of 43.5 for the non-gamified group and 40.2 for the gamified group, the result of the analysis was that there was a significant difference between the treatment groups for this dimension, $F(1, 42) = 4.8$, $p = .034$. Effect size was medium-large, $\eta^2 = .10$. Therefore the null hypothesis for overall task skill, H_04 – that participants practicing familiarization tasks on a gamified course will exhibit no difference in measurement of learner satisfaction – was rejected. This result was the inverse of the alternative hypothesis H_A4: that the gamified group would test higher for learner satisfaction; these results suggested the gamified group were significantly less satisfied than the non-gamified group.

Table 28

Descriptive Statistics for Participant Satisfaction Scores

Variable	Gamified	*N*	*M*	*SD*	Std. Error	Minimum	Maximum
SATIS	N	22	43.50	4.18	.89	34	50
	Y	22	40.18	5.74	1.22	26	49
Total		44	41.84	5.24	.79	26	50

Table 29

ANOVA: Satisfaction Scores (SATIS) by Gamification Group

	Sum of Squares	*df*	Mean square	*F*
Between-Groups	121.11	1	121.11	4.80*
Within-Groups	1058.77	42	25.21	
Total	1179.89	43		

Note. * $p < 0.05$, ** $p < 0.01$

Summary of Results. A significant difference was found between the groups for all study metrics with the exception of task skill scores for the flying movement subsection. Self-efficacy post-test scores were significantly higher for the non-gamified group than those of the gamified group, indicating that the gamified familiarization course resulted in lower self-efficacy perceptions than the non-gamified course. Participants in the gamified group exhibited significantly higher persistence than those in the non-gamified group. Overall task skill was also significantly higher in those who completed the gamified course compared to those who completed the non-gamified course. Finally, participants in the non-gamified group were significantly more satisfied with the experience than those in the gamified group. Table 30 and Figure 7 provide overviews of the analysis results.

Table 30

Summary of Findings

Dependent Variable	**Independent Variable**	**Reference Table(s)**	**Result**	
			Difference*	**Effect**
Self-efficacy post-test score	Group	18, 19	Sig., negative	Large
Total learning attempts	Group	20, 21	Sig., positive	Large
Total assessment attempts	Group	22, 23	Sig., positive	Large
Assessment attempts by skillset:				
Ground movement	Group	24	Sig., positive	Large
Flying movement	Group	24	Sig., positive	Large
Object Interaction	Group	24	Sig., positive	Large
Communication	Group	24	Sig., positive	Large
Total treatment time	Group		Sig., positive	Large
Total task skill score	Group	25, 26	Sig., positive	Large
Task skill score by skillset:				
Ground movement	Group	27	Sig., positive	Large
Flying movement	Group	27	No sig. diff.	-
Object interaction	Group	27	Sig., positive	Large
Communication	Group	27	Sig., positive	Large
Satisfaction score	Group	28, 29	Sig.§, negative	Med-large

Note. * *Positive = gamified group exhibited greater measurement for dependent variable*
Negative = gamified group exhibited lesser measurement for dependent variable
§ *Significant at* $p < 0.05$ *(unmarked = significant at* $p < 0.01$

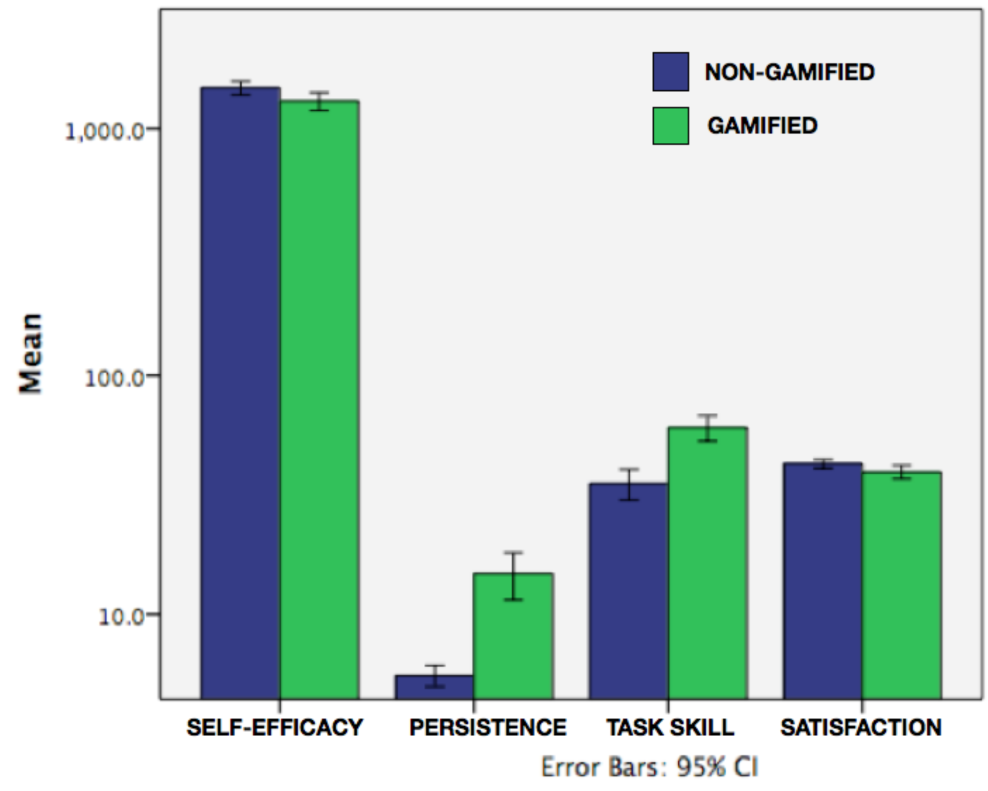

Figure 7: Between-groups comparison of means

Qualitative Analysis

Grounded theory methods of analysis (Glaser & Strauss, 1967) were used to analyze participants' responses to the post-session interview questions. The primary purpose of this analysis was to identify the motivations, perceptions and cognitive processes influencing the behavior observed during the treatment sessions. As a systematic qualitative approach that does not assume any preconceptions or hypotheses (Patton, 2002, p. 489), grounded theory allows for an inductive process of discovery and conceptualization of the data through the iterative examination of interview transcripts. This process begins with detailed, low-level data transcripted directly from the source (characterized as *description* by Corbin & Strauss (2014, p. 59)), and moves toward more general concepts that encompass and characterize emergent themes (*conceptual ordering* (2014, p. 61)) and possible relationships between them that might form the basis for theory development (2014, p. 62).

During *microanalysis* (2014, p. 71) participant responses were annotated with codes identifying the type of data point and some characterization of the data. For example, the participant reply "I knew I wasn't going to do any better," in response to a question regarding moving on in the course could be coded as *perceived maximum competence*, along with other possible codes. This initial examination and classification of data points is known as *open coding*, and generates a relatively large number of codes. The open codes, organized by interview question, are provided in Appendix D. As we were working with participants' viewpoints, values, feelings and expressed thoughts, it was appropriate to include the constructionist contributions to grounded theory put forth by Charmaz (2006), such as keeping the initial codes active (through constant analytical interaction and comparison) and paying particular attention to *in vivo* codes (terms coined natively by participants themselves) and their emergent meanings.

The open codes were considered within their contexts and compared to one another, forming connections and clarifying relationships – a process known as *axial coding*. The question to which the participant was responding often provided the broadest context, but answers to other questions were a critical source of additional information on the participant's mental paradigm and affective influences. Naturally, the treatment group to which the participant was assigned was also of interest. This iterative examination of the open codes resulted in the formation of interrelated categories (Figure 8).

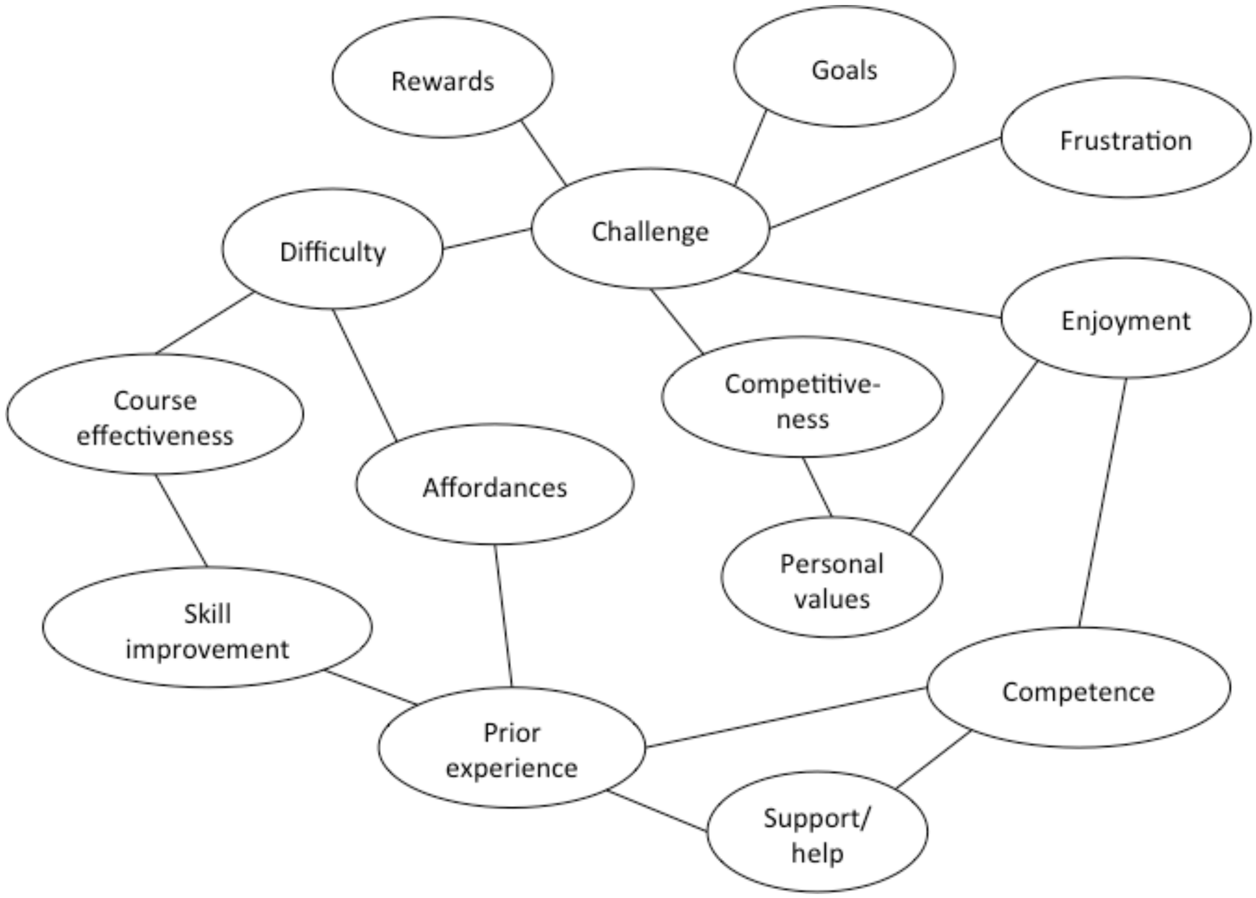

Figure 8: Concept map of categories resulting from axial coding of interview responses

Further analysis resulted in the emergence of four themes in the qualitative data:

- **Perception of course value.** How the participant understands the course in terms of effectiveness, significance and usefulness: A perception of worth.
- **Perception of competence improvement.** The participant's degree of belief that progress through the course is resulting in an improvement in his or her ability to perform in the MUVE.
- **Challenge: Enjoyment vs. frustration.** The participant's affective response to the challenges of the course: Whether they are enjoying the process or are feeling frustrated by the experience.

- **Perceived importance of success.** The level of personal investment in the outcome of the course: The degree to which the participant cares about whether or not he or she is successful.

These themes were then theoretically grouped by the four focal points of the study that the themes appear to be influencing: Self-efficacy, persistence, task skill and satisfaction (Figure 9).

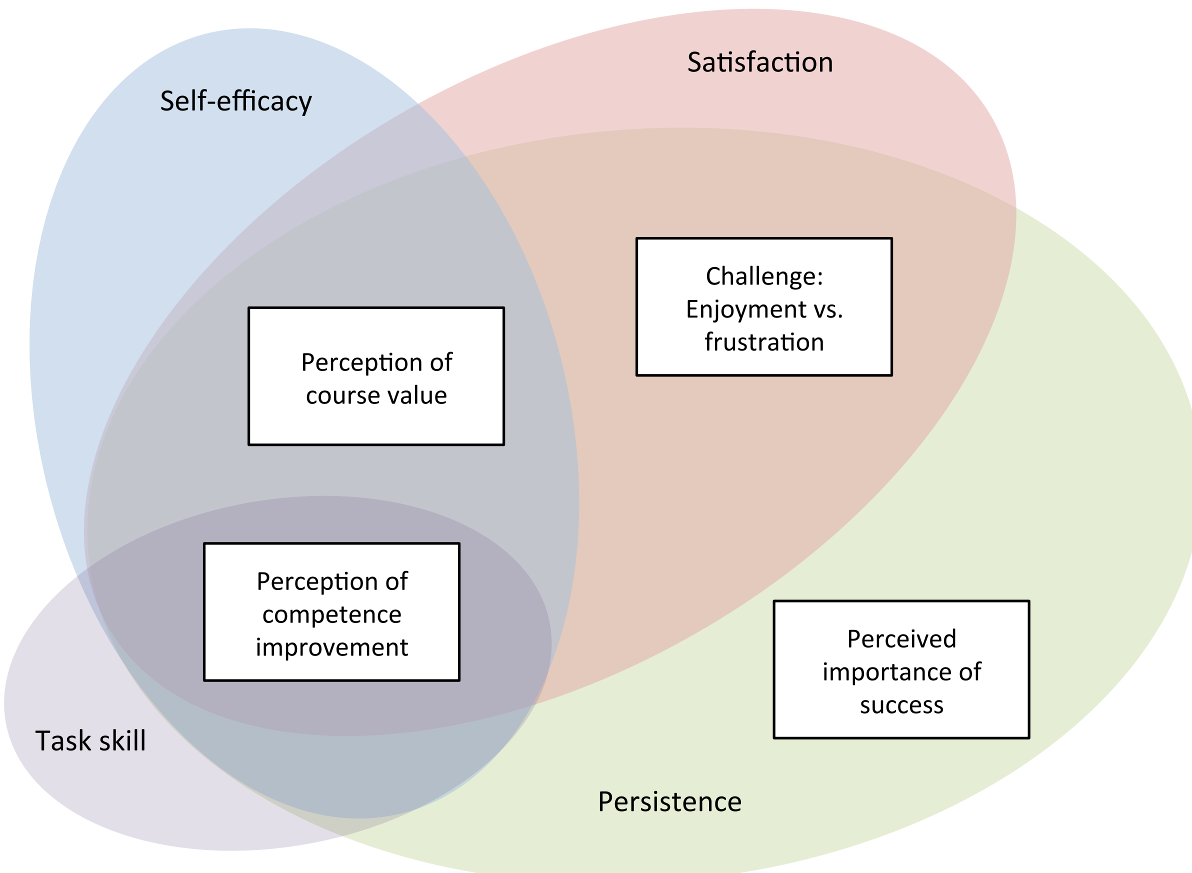

Figure 9: Emergent themes grouped by study focus

The relationships between the themes and study foci helped clarify how the interactions of expectations, lived experiences, and post-session perceptions of participants might have influenced the behavior observed during the study.

The participants' perception of course value appeared to affect three of the four study foci. Participants who considered the course valuable and effective would be likely to have their self-efficacy perceptions positively influenced, as they consider themselves to have been trained using a worthwhile process. Likewise, a positive view of the value of the course should increase satisfaction with the course. Lastly, participants who believed the course to be valuable would be more likely to persist at the tasks compared to those who felt it was a waste of time.

The perception of competence improvement – whether participants felt like they had become more competent – was likely to be a factor in all four quantitative measurements. A participant's self-efficacy might be expected to increase if they perceived that they had successfully completed the course (thus an improvement in perceived competence), and decrease if the participant had reason to perceive themselves as unsuccessful. A participant might be expected to persist less if they developed a perception that they had become competent (thus further practice being unnecessary). Perceived competence, so long as the perception is accurate, should be commensurate with task skill measurements. Satisfaction with the course should have increased when a participant perceived that the course did what it was intended to do (increase competence and familiarity).

Participants' affective response to the challenge of the course might be expected to influence their persistence behavior: Those that become overly frustrated might give up sooner than those that are enjoying the experience. However, this is a factor that could be subject to individual variation in temperament and personal values, as some may have "dug in their heels" as a reaction to frustration, stubbornly refusing to give up. Additionally, perceived time constraints could further confound this dimension: Persistence could have been negatively affected due to the artificial environment of an observed study session and the participant becoming concerned that they were taking too long.

Finally, the degree of importance the participant placed on success in the familiarization course is another factor that could be expected to affect persistence behavior: Those that felt success was very important would be likely to try harder to achieve it.

Chapter 5: Discussion and Conclusion

This chapter presents the findings of the study and discusses how the results relate to the literature. We also discuss how the qualitative data informed the quantitative data and vice versa. The limitations of the study are identified. The implications for future practice and research are also included. Finally, the author's conclusions from the study are provided.

Findings

The research question for the study was formulated with the intent to contribute to the body of knowledge for the design of pre-training courses for MUVE-based educational programs. As supported by the literature review, familiarization courses are necessary to attenuate the counterproductive effects of initial disorientation (Burigat & Chittaro, 2007; Ruddle et al., 1997; Smith & Marsh, 2004) and frustration (Nash et al., 2000) with virtual learning environments amongst new users (Eschenbrenner et al., 2008), and as noted by Jarmon, Traphagan, Mayrath, & Trivedi (2009), there exists little empirical evidence to guide the development of such pre-training courses. It was postulated in the study hypotheses that the gamification of pre-training courses may improve students' self-efficacy with the multi-user virtual environment, encourage them to persist while learning to use it, increase their skill at using it, and finally, improve their feelings of satisfaction with the format. The results of the principal quantitative analyses showed that persistence and task skill of the gamified group were statistically significantly greater compared to the non-gamified group, but self-efficacy and satisfaction were statistically significantly lower.

Self-efficacy. Results of statistical analysis on participants' self-efficacy perceptions after completing the familiarization course showed a significant difference between the group that used the gamified course and the group that used the non-gamified course. Contrary to the hypothesis, the mean self-efficacy score for the gamified group was approximately 12% lower than the mean score for the non-gamified group. Therefore, the non-gamified course resulted in higher self-efficacy perceptions than the gamified course. A possible cause of this phenomenon was the visibility of success and failure indicators for the gamified group, which were not visible to the non-gamified group. The gamified course displayed the timer for each assessment attempt and informed the participant whether or not they had qualified or if an achievement badge had been earned. Conversely, the non-gamified course displayed no such indicators to the participants. Performance feedback for the non-gamified course was simply whether or not the skillset station had been completed, regardless of time. Participants in the non-gamified group could simply go slowly and carefully complete each station without the pressure of a visible timer. This was reflected in the number of errors made during assessment attempts, which were 2.2 mean total errors for the non-gamified group, and 12 mean total errors for the gamified group. The self-efficacy perceptions of the gamified course participants may have been negatively affected by the knowledge that they had made a high number of mistakes and had not performed well enough to earn the rewards of the "game." This theory was supported by the fact that qualifying status and badges were not earned at a high rate, especially on the ground movement station and flying movement station (where only 13% and 9% earned a reward respectively). These were also the stations on which participants expressed most frustration with the difficulty of the tasks. Fogg's work on behavior and persuasive design may provide some further insight here: He found that positive motivation may not be solely sufficient in prompting a change in behavior unless the individual also believes he or she has the ability to perform the behavior (2009). If participants in the gamified group were initially positively motivated by the game elements to increase their persistence in the short-term, the negative performance feedback of failing to earn a reward may have suppressed the long-term self-efficacy perception that they were subsequently more capable of performing the task in the future. Similar impacts on post-test self-efficacy were seen in participants provided with negative performance feedback in research by Daniels and Larson (2001). It is also possible that the negative feedback of repeated failure increased the participants' perceived complexity of the tasks, resulting in negative effects on self-efficacy. This phenomenon of task complexity moderating self-efficacy gains was demonstrated by Stajkovic and Luthans (1998). If the difficulty level of the skillset stations had been lowered, an increased rate of rewards and lower number of errors may have led to a greater perception of task simplicity and a different outcome for this part of the study.

Persistence. Participants in the gamified course group demonstrated significantly higher levels of persistence than those in the non-gamified group. This was seen across all measurements of persistence: the total number of learning phase attempts (36% higher), the total number of assessment phase attempts (187% higher), and the total time spent completing the course (40% higher). This result suggested that participants in the gamified group were incentivized to attempt each skillset station repeatedly while those in the non-gamified group were satisfied with attempting each station once or twice before moving on. Of particular note was the considerably higher number of gamified attempts during assessment phases, where the qualitative data showed participants expressing their desire to perform well. Interview responses also indicated that the motivational sources for repeated attempts could be characterized as a mix of intrinsic ("it is important for me to succeed") and extrinsic ("I really wanted to earn the gold") for different participants. While Deci wrote on the potentially negative effects of extrinsic rewards on intrinsic motivation (1971), later work expounded that intrinsic motivation could be improved by extrinsic feedback when it is perceived to be informational in nature rather than controlling (Deci et al., 1999). Germane to the current study results, Muntean (2011) argued that gamification should include both intrinsic and extrinsic sources in order to increase engagement and motivation. Gamification elements in the current study were constructed in manner that provided extrinsic rewards such as badges and leaderboard scores, while simultaneously supporting intrinsic motivation through encouraging, informational feedback and positive reinforcement that adhered to Deci's concept of social approval (1971, p. 114) and a non-controlling milieu (Deci et al., 1999, p. 629).

The observed increased persistence for the gamified group, both in terms of the number of attempts observed at each station and longer time spent completing the course, were desirable from a design perspective because the purpose of the course was to familiarize the participants with the learning environment. As the course fostered foundational skills for using the MUVE, it was reasonable to expect increased course exposure to equal greater MUVE familiarization and subsequently higher task skill. The results supported this relationship between persistence, exposure, practice, and resultant improvements in skill.

Task skill. The statistical analysis of the overall task skill scores indicated a significant difference between the treatment groups, with a gamified group mean score approximately 70% higher than the non-gamified group's mean score. This aligns with results seen in other gamified studies examining task performance (Li, Grossman, & Fitzmaurice, 2012; Ninaus et al., 2015; Ong, 2013). When the individual skillset station scores were analyzed, all showed a significant difference in favor of the gamified group except the flying movement station, where no significant difference was evident. I postulated that this result was due to the higher difficulty of the flying movement station relative to the other stations, which may have had a compressive effect on task skill scores and reduced the probability of a significant difference. An increased task skill score on the familiarization course was considered desirable because the tasks were designed to represent those skills necessary to interact with the MUVE in a learning context. Consequently, participants who became more skilled at the tasks would be more capable of participating in MUVE-based learning activities without being distracted by the effects of disorientation and frustration.

Satisfaction. Participants in the non-gamified group were more satisfied with the familiarization course than those that completed the gamified course. The difference was small but statistically significant: The non-gamified group reported approximately 8% higher average satisfaction scores than the gamified group. From observations made during the treatment sessions and the qualitative analysis of interview responses, it was postulated that this result, the inverse of the hypothesis, may be due to the "failure mode" present in the gamified group and absent from the non-gamified group. Participants in the gamified group were aware of multiple measures to which they might ascribe success or failure, such as the timer, whether or not they qualified on the station, and whether or not they earned an achievement badge. As discussed previously, many participants had difficulty qualifying or earning a badge, and this could have led to a perception of overall failure and thus feeling less satisfied with the experience. As noted by Sweetser and Wyeth (2005), a delicate balance between level of challenge and user ability must be found for optimum enjoyment of a game-based task, particularly in attaining and maintaining this state of so-called *flow* (Czikszentmihalyi, 1991). Constraints within the *Second Life* platform and on the resources available for this study made it unfeasible to implement sophisticated game mechanics that could react and adapt to changes in the participant's abilities, and this may also have precluded a more satisfying experience for the gamified group.

Theoretical model. The findings contributed to the development of a theoretical model describing the mechanism of how a gamified pre-training system can support familiarization training for MUVE-based learning (Figure 10). The model defines a cumulative cycle of learning and motivational processes: The development of task skill, within the framework of situated learning, combines with performance feedback to increase perceptions of self-efficacy and satisfaction. Positive satisfaction and enjoyment of the experience, together with the challenge and reward of gamification, may encourage improved persistence behavior such as repeated attempts in response to challenges and more time spent learning basic competencies. These behaviors in turn promote familiarity with both the virtual world paradigm and the gamified learning process, allowing the learner to comfortably continue skill development by building on proficiencies acquired during previous iterations of the cycle. The majority of these concepts were supported by the study results, with indications that an adaptive response design may be necessary to optimize the learner experience and support self-efficacy outcomes. Therefore, an adaptive challenge and reward system that assesses and responds to changes in learner performance is recommended to avoid excessive learner frustration and more closely mirror the highly engaging nature of commercial video games.

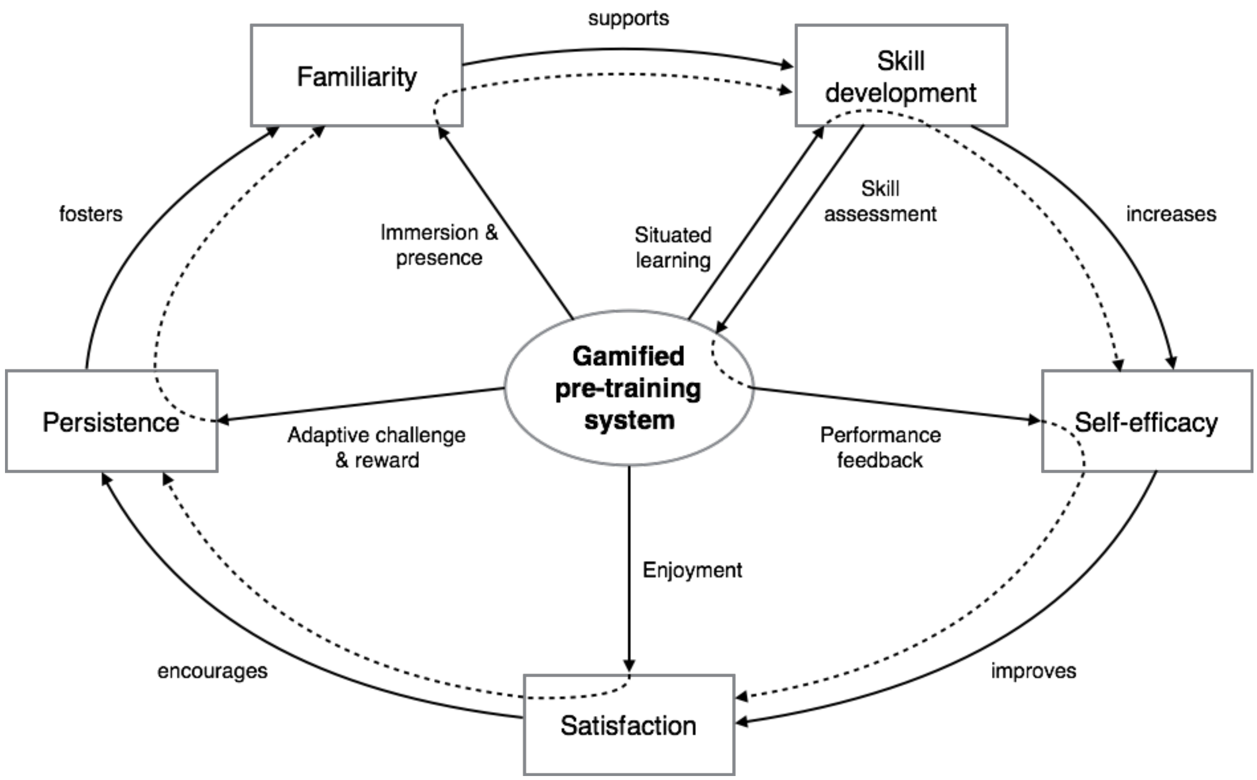

Figure 10: Theoretical model of gamification support for familiarization pre-training

Limitations

Second Life customizability. The primary limitation of this study was that a multi-user virtual environment that fully met all criteria for the planned study with minimal development was unattainable. The development of a custom MUVE was impractical due to cost, time and expertise constraints, leading to the necessity of adopting a Commercial Off-The-Shelf (COTS) solution. The *Second Life* environment was selected as the best COTS option due to its flexibility, low cost, and historically high adoption rate amongst educational institutions. Indeed, *Second Life* offered many opportunities for the creation of custom course components. However, the suitability of *Second Life* for the rapid development of game elements at a low cost was limited, especially with those course components that would have ideally executed complex adaptive behaviors in response to participant changes in ability. These constraints limited the breadth and complexity of gamification components that could be incorporated in the course, and thus narrowed the scope and richness of the techniques that were feasible.

Second Life system responsiveness. The *Second Life* platform suffers from occasional lag, where a noticeable delay is evident between the user's control input and the avatar's subsequent response. While network-induced lag is sometimes unavoidable due to Internet conditions or server load, many commercial video games locate the large graphics files necessary for rendering a rich, immersive three-dimensional world on the local "client" machine, precluding large, on-the-fly downloads. Online video games also use lag compensation techniques to predict player actions and reduce the user's perception of control delays (Bernier, 2001). Due to the communal, user-created nature of *Second Life*, much of its graphical content is stored on the server rather than the client. As an open world without inherent game pathways, *Second Life* is also unable compensate for lag through predictive interpolation. Participants familiar with online video games seemed especially sensitive to how lag affected their ability to control the avatar. These individuals expressed frustration when the responsiveness of the MUVE did not meet the expectations they had developed from playing online video games. This may have had a confounding effect on those participants' self-efficacy and satisfaction with the course.

Task skill score. Due to a concern that participants would be unwilling to commit to multiple study sessions or sessions lasting longer than 90 minutes, the calculation of the task skill score had an inherent weakness: The non-gamified group were not aware of being timed, and so had little incentive to complete the assessment phase substations quickly, but the task skill score was partially dependent on speed of completion. Hence the value of task skill score as a fully accurate comparison of actual task skill between the groups was debatable. A counter-argument can be made that the inclusion of the number of participant errors in the task skill calculation had an offsetting effect on the completion time, as the non-gamified group committed far fewer errors (and thus raised their scores) than the gamified group.

Attempting to eliminate this issue by informing the non-gamified group of the presence of a timer would have perpetuated the unintentional gamification effect observed in the pilot study. The study would then have ceased to be a pure comparison of gamified versus non-gamified courses. A potential solution would have been to add post-course assessments where members of both treatment groups were evaluated using the same process and participants could only attempt each skill assessment once. However, this would have increased the amount of development necessary and greatly extended the time commitment for participants.

Researcher as facilitator. The Primary Investigator was the facilitator for the participant sessions. This may have resulted in the inadvertent introduction of bias as the facilitator interacted with participants. Scripts were used to minimize any differences in the instructions used between and within groups, but some degree of behavioral or linguistic influence may have been possible. Ideally, a third party that was not aware of the investigative focus of the study would have facilitated the study sessions.

Situational authenticity. The study's scenario of students preparing for a MUVE-based learning experience was hypothetical: Participants were not actually about to enter a MUVE-based course in which they enrolled as an actual student. This may have affected the participants' motivation to engage authentically with the familiarization course. Observed behavior in the study may not accurately reflect the behavior of actual students preparing for MUVE-based learning.

Non-representative sample. Participants were all recruited from two higher education institutions within the same US state. Most were undergraduate students within the approximate age range of 18-26 years. The study design did not include measures to ensure that the sample accurately reflected the general population. Hence the results presented here may not be generalizable to other populations and contexts.

Implications

Instructional Design. The results of the current study can inform the future design direction of MUVE-based learning experiences. The study confirmed the value of gamification for improving persistence, which – through increased repetition – led to superior performance in the skills necessary for MUVE-based learning activities. Observation of this phenomenon confirmed the conclusions of previous works that found that increased persistence results in higher performance over time (Lent et al., 1984; Whitehill & McDonald, 1993; Zimmerman, 2000), especially during the earlier phases of learning (Pajares & Schunk, 2001, p. 19). Pre-training exercises are, of course, archetypal of an early phase of learning.

However, while Bandura (1997) theorized that persistence and effort was a product of self-efficacy, the significantly higher persistence measured in the gamified group did not correlate to better post-test self-efficacy in the same group. As postulated earlier, this may have been due to the following factors, or a combination thereof:

1. The presence of more detailed performance feedback from the gamified skillset stations highlighted the participants' errors and potentially led to a perception of inadequate performance.
2. The inability of the familiarization course to dynamically respond and adapt to the skill level of the participant led to a perception that the tasks were complex and that improvements in skill performance were not being realized according to the participant's efficacy expectations.

Instructional designers should consider offsetting feedback that might be considered damaging to self-efficacy with additional support or encouragement, and avoid designing systems with only a fixed level of challenge. To address the results of this study with regard to self-efficacy and satisfaction, gamified pre-training systems should incorporate an adaptive mechanism that modifies feedback and adjusts the difficulty of tasks according to learner performance. The task must remain challenging for motivational purposes, but should allow for learners to achieve some success before raising the difficulty level or lowering support as competence improves. This adaptive challenge mechanism is common in video games, and prior research has confirmed its value in supporting users' self-efficacy and engagement (Backlund, Engström, Johannesson, & Lebram, 2008; Czikszentmihalyi, 1991).

Future Research. This study has provided evidence that the inclusion of game elements in MUVE pre-training courses can have a powerful effect on learner motivation and lead to increased exposure time to familiarization experiences. If MUVEs continue to be used as a delivery modality for educational experiences, there will continue to be a need for new users to become accustomed with the MUVE paradigm before engaging in learning activities. As discussed previously, inadequate empirical information has been available to guide the design of effective pre-training courses for educational MUVEs, some of which have existed for decades. Now that the potential effects of gamification have been demonstrated in this context, the technique might be extended to nascent paradigms in the virtual world domain. At the time of writing, a new wave of virtual reality (VR) technology was emerging. Several leading technology companies were weeks away from launching VR headset systems that promised unprecedented levels of immersion in virtual environments (HTC, 2016; Oculus VR, 2016; Sony Corporation, 2016). These headsets break the model of representing a three-dimensional virtual world on a flat screen desktop display, instead using dynamic tracking systems to synchronize the visor's viewpoint with the wearer's head movements. This leads to the perception of the virtual world as an environment that is sensed in a manner akin to how humans sense reality. This author believes that it is highly probable that VR headset technology will follow a similar path into education as MUVEs (which were also initially driven by video gaming): It will be adopted for synchronous educational activities where learners cannot be geographically together, or the activity is dangerous, impractical or otherwise unfeasible. While the mode of access will be paradigmatically closer to "real life" than desktop screen displays, users are likely to still require some degree of familiarization with the system before becoming competent enough to engage in meaningful learning tasks in the virtual world. Until the technology becomes sophisticated enough to be indistinguishable from the reality with which we are all conversant, pre-training will be required to familiarize the user with the interface protocols, control mechanisms and other affordances of the VR-enabled virtual world. Continued research on the selective use of game elements to reduce motivational barriers and encourage persistence at the familiarization stage will assist instructional designers to produce effective pre-training courses that enable more learners to benefit from these educational opportunities with less frustration.

Conclusion

As a mechanism for improving pre-training course for MUVE-based learning programs, gamification has substantial potential. This study has demonstrated that learners in a gamified pre-training course were more likely to persist in the face of familiarization challenges, spend more time acquainting themselves with the MUVE, and perform at a higher skill level than those using a non-gamified course. However, the study also emphasized the importance of a careful balance between challenge and reward within the gamified design. Combining a high difficulty level with exposure to performance feedback may have been the source of negative perceptions of self-efficacy and reduced satisfaction with the gamified course. Further research on the motivational aspects of game mechanics and their specific effects on self-efficacy would help clarify this hypothesis. As lowered self-efficacy and satisfaction are both counter to the purpose of pre-training courses, these consequences should be attenuated or eliminated wherever possible. Iterative user testing to optimize the challenge-reward experience during development, similar to that used in commercial video game design, is strongly recommended. Where feasible, pre-training courses should also incorporate adaptive tasks whose difficulty adjusts according to user achievement, providing an optimum level of challenge at all stages of familiarization. These enhancements may result in a more satisfying experience that increases self-efficacy perceptions for MUVE-based learning, as well as the positive effects on persistence and task skill evidenced here.

In closing, we suggest that the fundamental design of this investigation may be applicable to familiarization initiatives for new learning technologies that are likely to transpire from the development of emerging virtual reality systems such as the HTC Vive, Oculus Rift and PlayStation VR.

References

Appleton-Knapp, S. L., & Krentler, K. A. (2006). Measuring student expectations and their effects on satisfaction: The importance of managing student expectations. *Journal of Marketing Education*, *28*(3), 254–264.

Armstrong, D. (2013). The new engagement game: the role of gamification in scholarly publishing. *Learned Publishing*, *26*(4), 253–256.

Backlund, P., Engström, H., Johannesson, M., & Lebram, M. (2008). Games for traffic education: An experimental study of a game-based driving simulator. *Simulation & Gaming*.

Bailey, B. L., Bauman, C., & Lata, K. A. (1998). Student retention and satisfaction: The evolution of a predictive model. In *38th Annual Forum of the Association for Institutional Research, 17-20 May, Minneapolis*.

Bailey, F., & Moar, M. (2001). The Vertex Project: Children Creating and Populating 3D Virtual Worlds. *Journal of Art & Design Education*, *20*(1), 19–30.

Bandura, A. (1977). Self-efficacy: Toward a Unifying Theory of Behavioral Change. *Psychological Review*, *84*(2), 191–215.

Bandura, A. (1986). *Social foundations of thoughts and action: A social cognitive theory*. Englewood Cliffs: NJ: Prentice Hall.

Bandura, A. (1993). Perceived Self-Efficacy in Cognitive Development and Functioning. *Educational Psychologist*, *28*(2), 117–148.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W.H. Freeman & Co.

Bandura, A. (2006). Guide to constructing self-efficacy scales. In T. Urdan & F. Pajares (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307–337). Information Age.

Barata, G., Gama, S., Jorge, J., & Goncalves, D. (2013). Engaging Engineering Students with Gamification. In *Games and Virtual Worlds for Serious Applications (VS-GAMES), 2013 5th International Conference on* (pp. 1–8).

Barfield, W., & Weghorst, S. (1993). The sense of presence within virtual environments: A conceptual framework. *Advances in Human Factors Ergonomics*, *19*, 699–699.

Bernier, Y. W. (2001). Latency Compensating Methods in Client/Server In-game Protocol Design and Optimization. Retrieved from https://developer.valvesoftware.com/wiki/Latency_Compensating_Methods_in_Client/Server_In-game_Protocol_Design_and_Optimization#Lag_Compensation

Blackwell, L., Trzesniewski, K., & Dweck, C. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, *78*(1), 246–263.

Boulos, M. N. K., Hetherington, L., & Wheeler, S. (2007). Second Life: an overview of the potential of 3-D virtual worlds in medical and health education. *Health Information & Libraries Journal*, *24*(4), 233–245.

Bower, B. L., & Hardy, K. P. (2004). From Correspondence to Cyberspace: Changes and Challenges in Distance Education. *New Directions for Community Colleges*, (128), 5–12.

Bowers, K. W., Ragas, M. W., & Neely, J. C. (2009). Assessing the Value of Virtual Worlds for Post- Secondary Instructors: A Survey of Innovators, Early Adopters and the Early Majority in Second Life. *World Academy of Science, Engineering and Technology*, *3*, 709–719.

Bronack, S., Riedl, R., & Tashner, J. (2006). Learning in the zone: A social constructivist framework for distance education in a 3-dimensional virtual world. *Interactive Learning Environments*, *14*(3), 219–232.

Browne, T., Jenkins, M., & Walker, R. (2006). A longitudinal perspective regarding the use of VLEs by higher education institutions in the United Kingdom. *Interactive Learning Environments*, *14*(2), 177–192.

Burigat, S., & Chittaro, L. (2007). Navigation in 3D virtual environments: Effects of user experience and location-pointing navigation aids. *International Journal of Human-Computer Studies*, *65*, 945–958.

Bystrom, K.-E., Barfield, W., & Hendrix, C. (1999). A conceptual model of the sense of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, *8*(2), 241–244.

Carr, D., Oliver, M., & Burn, A. (2010). Learning, Teaching and Ambiguity in Virtual Worlds. In A. Peachey, J. Gillen, D. Livingstone, & S. Smith-Robbins (Eds.), *Researching Learning in Virtual Worlds* (pp. 17–30). Springer London.

Chang, M.-M., & Lehman, J. D. (2002). Learning foreign language through an interactive multimedia program: An experimental study on the effects of the relevance component of the ARCS model. *CALICO Journal*, *20*(1), 81–98.

Charmaz, K. (2006). *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis*. SAGE Publications.

Chittaro, L., & Ranon, R. (2007). Web3D technologies in learning, education and training: Motivations, issues, opportunities. *Computers & Education*, *49*(1), 3–18.

Cohen, J. (1992). A Power Primer. *Psychological Bulletin*, *112*(1), 155–159.

Corbin, J., & Strauss, A. (2014). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (4th ed.). Sage publications.

Cordova, D. I., & Lepper, M. R. (1996). Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology*, *88*(4), 715–730.

Czikszentmihalyi, M. (1991). *Flow: The psychology of optimal experience*. Harper Perennial.

Dalgarno, B., & Lee, M. (2010). What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology*, *41*(1), 10–32.

Daniels, J. A., & Larson, L. M. (2001). The Impact of Performance Feedback on Counseling Self-Efficacy and Counselor Anxiety. *Counselor Education and Supervision*, *41*(2), 120–130.

Darken, R. P., & Peterson, B. (2002). Spatial orientation, wayfinding, and representation. *Handbook of Virtual Environments*, 493–518.

Darken, R. P., & Sibert, J. L. (1996). Navigating large virtual spaces. *International Journal of Human-Computer Interaction*, *8*(1), 49–71.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, *18*(1), 105–115.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, *125*(6), 627–668.

de-Marcos, L., Domínguez, A., Saenz-de-Navarrete, J., & Pagés, C. (2014). An empirical study comparing gamification and social networking on e-learning. *Computers & Education*, *75*, 82–91.

Denny, P. (2013). The effect of virtual achievements on student engagement. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 763–772). ACM.

deNoyelles, A., Hornik, S. R., & Johnson, R. D. (2014). Exploring the Dimensions of Self-Efficacy in Virtual World Learning: Environment, Task, and Content. *MERLOT Journal of Online Learning and Teaching*, *10*(2), 255–271.

Deterding, S. (2011). Situated motivational affordances of game elements: A conceptual model. In *Gamification: Using Game Design Elements in Non-Gaming Contexts, a workshop at CHI*.

Dickey, M. D. (2007). Game design and learning: a conjectural analysis of how massively multiple online role-playing games (MMORPGs) foster intrinsic motivation. *Educational Technology Research and Development*, *55*(3), 253–273.

Domínguez, A., Saenz-De-Navarrete, J., De-Marcos, L., Fernández-Sanz, L., Pagés, C., & Martinez-Herráiz, J.-J. (2013). Gamifying Learning Experiences: Practical Implications and Outcomes. *Computers & Education*, *63*, 380–392.

Duncan, I., Miller, A., & Jiang, S. (2012). A taxonomy of virtual worlds usage in education. *British Journal of Educational Technology*, *43*(6), 949–964.

Dweck, C. S. (2006). *Mindset: The new psychology of success.* New York: Random House.

Dweck, C. S. (2007). The perils and promise of praise. *Educational Leadership*, *65*(2), 34–39.

Eschenbrenner, B., Nah, F. F.-H., & Siau, K. (2008). 3-D Virtual Worlds in Education: Applications, Benefits, Issues, and Opportunities. *Journal of Database Management*, *19*(4), 91–110.

Fagan, M. H., Neill, S., & Wooldridge, B. R. (2003). An empirical investigation into the relationship between computer self-efficacy, anxiety, experience, support and usage. *Journal of Computer Information Systems*, *44*(2), 95.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191.

Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics* (4th ed.). SAGE.

Flatla, D., Gutwin, C., Nacke, L., Bateman, S., & Mandryk, R. (2011). Calibration Games: Making Calibration Tasks Enjoyable by Adding Motivating Game Elements. In *UIST '11: Proceedings of the 24th annual ACM symposium on User interface software and technology* (pp. 403–412). Santa Barbara, California, USA.

Fogg, B. (2009). A Behavior Model for Persuasive Design. In *Proceedings of the 4th International Conference on Persuasive Technology* (pp. 40:1–40:7). New York, NY, USA: ACM.

Foll, D. L., Rascle, O., & Higgins, N. (2006). Persistence in a Putting Task During Perceived Failure: Influence of State-attributions and Attributional Style. *Applied Psychology*, *55*(4), 586–605.

Foster, A. L. (2007). Professor Avatar. *Chronicle of Higher Education*, *54*(4), A24–A26.

Fransson, A. (1977). On qualitative differences in learning: IV—Effects of intrinsic motivation and extrinsic test anxiety on process and outcome. *British Journal of Educational Psychology*, *47*(3), 244–257.

Garrison, D. R., Anderson, T., & Archer, W. (1999). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education*, *2*(2), 87–105.

Gee, J. P. (2007). *What Video Games Have to Teach Us About Learning and Literacy*. New York: Palgrave MacMillan.

George, D., & Mallery, M. (2003). Using SPSS for Windows step by step: a simple guide and reference. *Boston, MA: Allyn & Bacon.*

Glaser, B. G., & Strauss, A. L. (1967). The discovery of grouded theory. *Chicago (US): Aldine*.

Grand Theft Auto: San Andreas. (2004). Rockstar Games.

Gregory, S., & Gregory, B. (2011). Do virtual worlds have a role in increasing student engagement as measured by their higher academic grades? In M. Docherty & M. Hitchcock (Eds.), *CreateWorld 2011 Proceedings* (pp. 40–50). Brisbane, Australia.

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does Gamification Work? — A Literature Review of Empirical Studies on Gamification. In *Proceedings of the 47th Hawaii International Conference on System Sciences* (pp. 3025–3034). Hawaii, USA: IEEE Computer Society.

Hayashi, A., Chen, C., Ryan, T., Wu, J., & others. (2004). The role of social presence and moderating role of computer self efficacy in predicting the continuance usage of e-learning systems. *Journal of Information Systems Education*, *15*, 139–154.

Heeter, C. (1992). Being there: The subjective experience of presence. *Presence: Teleoperators and Virtual Environments*, *1*(2), 262–271.

Holmberg, K., & Huvila, I. (2008). Learning together apart: Distance education in a virtual world. *First Monday*, *13*(10).

HTC. (2016, February 21). HTC and Valve bring virtual reality to life with unveiling of consumer edition. Retrieved from https://www.htc.com/us/about/newsroom/2016/htc-and-valave-bring-virtual-reality-to-life/

Igbaria, M., & Iivari, J. (1995). The effects of self-efficacy on computer usage. *Omega*, *23*(6), 587–605.

Jacobs, B., Prentice-Dunn, S., & Rogers, R. W. (1984). Understanding Persistence: An Interface of Control Theory and Self-Efficacy Theory. *Basic and Applied Social Psychology*, *5*(4), 333–347.

Jarmon, L., Traphagan, T., Mayrath, M., & Trivedi, A. (2009). Virtual world teaching, experiential learning, and assessment: An interdisciplinary communication course in Second Life. *Computers & Education*, *53*(1), 169–182.

Jennings, N., & Collins, C. (2007). Virtual or Virtually U: Educational Institutions in Second Life. *World Academy of Science, Engineering and Technology*, *1*, 433–439.

Jones, G. R., & George, J. M. (2007). *Contemporary Management* (5th ed.). McGraw-Hill.

Jones, J. G., Morales, C., & Knezek, G. A. (2005). 3-Dimensional online learning environments: examining attitudes toward information technology between students in Internet-based 3-dimensional and face-to-face classroom instruction. *Educational Media International*, *42*(3), 219–236.

Kang, M., Ragan, B. G., & Park, J.-H. (2008). Issues in Outcomes Research: An Overview of Randomization Techniques for Clinical Trials. *Journal of Athletic Training*, *43*(2), 215–221.

Keller, J. M. (1983). Motivational design of instruction. In C. Reigeluth (Ed.), *Instructional design theories and models: An overview of their current status* (pp. 386–434). Hillsdale, NJ: Erlbaum.

Keller, J. M. (1984). The use of the ARCS model of motivation in teacher training. *Aspects of Educational Technology*, *17*, 140–145.

Keller, J. M. (1987). Strategies for stimulating the motivation to learn. *Performance+ Instruction*, *26*(8), 1–7.

Keller, J. M. (1999). Using the ARCS Motivational Process in Computer-Based Instruction and Distance Education. *New Directions for Teaching and Learning*, *1999*(78), 37–47.

Keller, J. M. (2008). First principles of motivation to learn and e3-learning. *Distance Education*, *29*(2), 175–185.

Keller, J. M., & Kopp, T. W. (1987). An application of the ARCS Model of Motivational Design. In C. Reigeluth (Ed.), *Instructional theories in action: Lessons illustrating selected theories and models* (pp. 289–320). Hillsdale, NJ: Lawrence Erlbaum Associates.

Ketelhut, D. J., Dede, C., Clarke, J., Nelson, B., & Bowman, C. (2007). Studying situated learning in a multi-user virtual environment. *Assessment of Problem Solving Using Simulations*, 37–58.

Ketelhut, D. J., Nelson, B. C., Clarke, J., & Dede, C. (2010). A multi-user virtual environment for building and assessing higher order inquiry skills in science. *British Journal of Educational Technology*, *41*(1), 56–68.

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development* (Vol. 1). Prentice-Hall Englewood Cliffs, NJ.

Koontz, H. (2010). *Essentials Of Management 8E*. McGraw-Hill Education (India) Pvt Limited.

Lave, J., & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press.

Lent, R. W., Brown, S. D., & Larkin, K. C. (1984). Relation of self-efficacy expectations to academic achievement and persistence. *Journal of Counseling Psychology*, *31*(3), 356–362.

Levy, Y. (2007). Comparing dropouts and persistence in e-learning courses. *Computers & Education*, *48*(2), 185–204.

Liaw, S.-S. (2008). Investigating students' perceived satisfaction, behavioral intention, and effectiveness of e-learning: A case study of the Blackboard system. *Computers & Education*, *51*(2), 864–873.

Ligorio, M. B., & van Veen, K. (2006). Strategies to Build a Cross-National Virtual World. *AACE Journal*, *14*(2), 103–128.

Lim, C. K. (2001). Computer self-efficacy, academic self-concept, and other predictors of satisfaction and future participation of adult distance learners. *American Journal of Distance Education*, *15*(2), 41–51.

Lipman, M. (1991). Thinking in education. *Cambridge University, Cambridge*.

Li, W., Grossman, T., & Fitzmaurice, G. (2012). GamiCAD: a gamified tutorial system for first time autocad users. In *Proceedings of the 25th annual ACM symposium on User interface software and technology* (pp. 103–112). ACM.

Maddux, J., Sherer, M., & Rogers, R. (1982). Self-efficacy expectancy and outcome expectancy: Their relationship and their effects on behavioral intentions. *Cognitive Therapy and Research*, *6*(2), 207–211.

Marakas, G., Johnson, R., & Clay, P. F. (2007). The Evolving Nature of the Computer Self-Efficacy Construct: An Empirical Investigation of Measurement Construction, Validity, Reliability and Stability Over Time. *Journal of the Association for Information Systems*, *8*(1), 2.

Mayer, R. E. (2014). *Computer Games for Learning: An Evidence-Based Approach*. MIT Press.

Mayer, R. E., Mathias, A., & Wetzell, K. (2002). Fostering understanding of multimedia messages through pre-training: Evidence for a two-stage theory of mental model construction. *Journal of Experimental Psychology: Applied*, *8*(3), 147–154.

Mayer, R. E., Mautone, P., & Prothero, W. (2002). Pictorial aids for learning by doing in a multimedia geology simulation game. *Journal of Educational Psychology*, *94*(1), 171–185.

McKerlich, R., & Anderson, T. (2007). Community of inquiry and learning in immersive environments. *Journal of Asynchronous Learning Networks*, *11*(4), 35–52.

Means, T. B., Jonassen, D. H., & Dwyer, F. M. (1997). Enhancing relevance: Embedded ARCS strategies vs. purpose. *Educational Technology Research and Development*, *45*(1), 5–17.

Molka-Danielsen, J., Richardson, D., Deutschmann, M., & Carter, B. (2007). Teaching languages in a virtual world. *NOKOBIT Proceedings 2007*.

Morrison, G., Ross, S., Kalman, H., & Kemp, J. (2012). *Designing Effective Instruction* (7th ed.). John Wiley & Sons.

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology*, *38*(1), 30–38.

Muntean, C. I. (2011). Raising engagement in e-learning through gamification. In *Proc. 6th International Conference on Virtual Learning ICVL* (pp. 323–329).

Nash, E. B., Edwards, G. W., Thompson, J. A., & Barfield, W. (2000). A Review of Presence and Performance in Virtual Environments. *International Journal of Human-Computer Interaction*, *12*(1), 1–41.

Nelson, B. C. (2007). Exploring the use of individualized, reflective guidance in an educational multi-user virtual environment. *Journal of Science Education and Technology*, *16*(1), 83–97.

Nelson, B. C., & Ketelhut, D. J. (2008). Exploring embedded guidance and self-efficacy in educational multi-user virtual environments. *International Journal of Computer-Supported Collaborative Learning*, *3*(4), 413–427.

Ninaus, M., Pereira, G., Stefitz, R., Prada, R., Paiva, A., Neuper, C., & Wood, G. (2015). Game elements improve performance in a working memory training task. *Int. J. Ser. Games*, *2*, 3–16.

O'Connell, J., & Groom, D. (2010). *Virtual Worlds: Learning in a Changing World*. Victoria, Australia: Australian Council for Educational Research.

Oculus VR. (2016, January 6). Oculus Rift Pre-Orders Now Open, First Shipments March 28. Retrieved from https://www.oculus.com/en-us/blog/oculus-rift-pre-orders-now-open-first-shipments-march-28/

Ong, M. (2013). *Gamification and its effect on employee engagement and performance in a perceptual diagnosis task.*

O'Rourke, E., Haimovitz, K., Ballweber, C., Dweck, C., & Popović, Z. (2014). Brain points: a growth mindset incentive structure boosts persistence in an educational game. In

Proceedings of the 32nd annual ACM conference on Human factors in computing systems (pp. 3339–3348). ACM.

Pajares, F. (1996). Assessing self-efficacy beliefs and academic outcomes: The case for specificity and correspondence. In *Measuring and mismeasuring self-efficacy: Dimensions, problems, and misconceptions*. New York: AERA.

Pajares, F. (1997). Current directions in self-efficacy research. In M. Maehr & P. R. Pintrich (Eds.), *Advances in motivation and achievement* (Vol. 10, pp. 1–49). JAI.

Pajares, F. (2002). *Overview of social cognitive theory and of self-efficacy*. Retrieved from http://www.uky.edu/~eushe2/Pajares/eff.html

Pajares, F., & Schunk, D. H. (2001). The development of academic self-efficacy. In A. Wigfield & J. Eccles (Eds.), *Development of achievement motivation*.

Parush, A., & Berman, D. (2004). Navigation and orientation in 3D user interfaces: the impact of navigation aids and landmarks. *International Journal of Human-Computer Studies*, *61*(3), 375–395.

Patton, M. Q. (2002). *Qualitative Research and Evaluation Methods* (3rd ed.). SAGE.

Pirius, L. K., & Creel, G. (2010). Reflections on Play, Pedagogy, and “World of Warcraft.” *EDUCAUSE Quarterly*, *33*(3).

Porter, L. W., & Lawler, E. E. (1968). *Managerial attitudes and performance*. Homewood, IL: Irwin-Dorsey.

Rankin, Y., Morrison, D., & Shute, M. (2009). Utilizing Massively Multiplayer Online Games to foster collaboration and learning. In *Science and Innovation Policy, 2009 Atlanta Conference on* (pp. 1–10).

Ravaja, N., Saari, T., Salminen, M., Laarni, J., & Kallinen, K. (2006). Phasic emotional reactions to video game events: A psychophysiological investigation. *Media Psychology*, *8*(4), 343–367.

Raymer, R. (2011, September). Gamification: Using Game Mechanics to Enhance eLearning. *eLearn Magazine*.

Rosenthal, R. (1991). Effect sizes: Pearson's correlation, its display via the BESD, and alternative indices. *American Psychologist*, *46*(10), 1086–1087.

Rossett, A. (1987). *Training needs assessment*. Englewood Cliffs: NJ: Educational Technology Publications.

Ruddle, R. A., Payne, S. J., & Jones, D. M. (1997). Navigating buildings in "desk-top" virtual environments: Experimental investigations using extended navigational experience. *Journal of Experimental Psychology*, *3*(2), 143–159.

Ryan, R. M., & Deci, E. L. (2000a). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, *25*, 54–67.

Ryan, R. M., & Deci, E. L. (2000b). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78.

Sandelands, L. E., Brockner, J., & Glynn, M. A. (1988). If at first you don't succeed, try, try again: Effects of persistence-performance contingencies, ego involvement, and self-esteem on task persistence. *Journal of Applied Psychology*, *73*(2), 208–216.

Schunk, D. H. (1981). Modeling and attributional effects on children's achievement: A self-efficacy analysis. *Journal of Educational Psychology*, *73*(1), 93.

Schunk, D. H. (1989). Self-efficacy and cognitive skill learning. *Research on Motivation in Education*, *3*, 13–44.

Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist*, *26*(3-4), 207–231.

Schunk, D. H., & Pajares, F. (2005). Competence Perceptions and Academic Functioning. In A. Elliot & C. Dweck (Eds.), *Handbook of Competence and Motivation* (pp. 85–104). Guilford Press.

Schunk, D. H., Pintrich, P., & Meece, J. (2008). *Motivation in Education: Theory, Research, and Application* (3rd ed.). Pearson Prentice Hall.

Schwarzer, R., & Jerusalem, M. (1995). Generalized self-efficacy scale. *Measures in Health Psychology: A User's Portfolio. Causal and Control Beliefs*, *1*, 35–37.

Segal, K. R., & Dietz, W. H. (1991). Physiologic responses to playing a video game. *American Journal of Diseases of Children*, *145*(9), 1034–1036.

Shaffer, D. W., Squire, K., Halverson, R., & Gee, J. P. (2005). Video games and the future of learning. *Phi Delta Kappan*, *87*(2), 105–111.

Sheldon, L. (2012). *The multiplayer classroom: Designing coursework as a game*. Cengage Learning.

Sherer, M., Maddux, J. E., Mercandante, B., Prentice-Dunn, S., Jacobs, B., & Rogers, R. W. (1982). The self-efficacy scale: Construction and validation. *Psychological Reports*, *51*(2), 663–671.

Shields, P. (1999). *The community of inquiry: Insights for public administration from Jane Addams, John Dewey and Charles S. Peirce*.

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

Small, R. V., & Gluck, M. (1994). The Relationship of Motivational Conditions to Effective Instructional Attributes: A Magnitude Scaling Approach. *Educational Technology*, *34*(8), 33–40.

Smith, S. P., & Marsh, T. (2004). Evaluating design guidelines for reducing user disorientation in a desktop virtual environment. *Virtual Reality*, *8*(1), 55–62.

Sony Corporation. (2016, March 16). PlayStation VR launches October 2016. Retrieved from https://www.playstation.com/en-us/corporate/press-releases/2016/playstation-vr-launches-october-2016-available-globally-at-44980-yen-399-usd-399-and-349/

Squire, K. (2007). Games, learning, and society: building a field. *Educational Technology*, *4*, 51–55.

Stajkovic, A. D., & Luthans, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin*, *124*(2), 240.

Sun, P.-C., Tsai, R. J., Finger, G., Chen, Y.-Y., & Yeh, D. (2008). What drives a successful e-Learning? An empirical investigation of the critical factors influencing learner satisfaction. *Computers & Education*, *50*(4), 1183–1202.

Superman: The New Adventures. (1999). [Nintendo 64]. Titus Software.

Sweetser, P., & Wyeth, P. (2005). GameFlow: A Model for Evaluating Player Enjoyment in Games. *Computers in Entertainment (CIE) - Theoretical and Practical Computer Applications in Entertainment*, *3*(3), 3–3.

Tachyon: The Fringe. (2000). [Microsoft Windows]. NovaLogic.

Thatcher, J. B., & Perrewé, P. L. (2002). An empirical examination of individual traits as antecedents to computer anxiety and computer self-efficacy. *Mis Quarterly*, 381–396.

Van Dijk, B., Op den Akker, R., Nijholt, A., & Zwiers, J. (2003). Navigation Assistance in Virtual Worlds. *Informing Science*, *6*.

Van Eck, R. (2006). Digital Game-Based Learning: It's Not Just the Digital Natives Who Are Restless. *EDUCAUSE Review*, *41*(2), 16–30.

Vansteenkiste, M., Simons, J., Lens, W., Sheldon, K. M., & Deci, E. L. (2004). Motivating learning, performance, and persistence: the synergistic effects of intrinsic goal contents and autonomy-supportive contexts. *Journal of Personality and Social Psychology*, *87*(2), 246–260.

Verhagen, T., Feldberg, F., van den Hooff, B., Meents, S., & Merikivi, J. (2011). Satisfaction with Virtual Worlds: An Integrated Model of Experiential Value. *Inf. Manage.*, *48*(6), 201–207.

Vinson, N. (1999). Design Guidelines for Landmarks to Support Navigation in Virtual Environments. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 278–185). New York: ACM.

Visser, J., & Keller, J. M. (1990). The clinical use of motivational messages: An inquiry into the validity of the ARCS model of motivational design. *Instructional Science*, *19*(6), 467–500.

Vroom, V. H. (1964). *Work and motivation*. New York: Wiley.

Wang, X., & Perry, A. C. (2006). Metabolic and physiologic responses to video game play in 7- to 10-year-old boys. *Archives of Pediatrics & Adolescent Medicine*, *160*(4), 411–415.

Wang, Y.-S. (2003). Assessment of learner satisfaction with asynchronous electronic learning systems. *Information & Management*, *41*(1), 75–86.

Warburton, S. (2009). Second Life in higher education: Assessing the potential for and the barriers to deploying virtual worlds in learning and teaching. *British Journal of Educational Technology*, *40*(3), 414–426.

Watson, W. R. (2013). An Argument for Clarity: What are Learning Management Systems, What are They Not, and What Should They Become? *TechTrends*, *51*(2), 28–34.

Whitehill, B. V., & McDonald, B. A. (1993). Improving Learning Persistence of Military Personnel by Enhancing Motivation in a Technical Trainiing Program. *Simulation & Gaming*, *24*(3), 294–313.

Wiggins, G. P., & McTighe, J. (2005). *Understanding by design* (2nd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.

Winn, W. (1993). A conceptual basis for educational applications of virtual reality. *Technical Publication R-93-9, Human Interface Technology Laboratory of the Washington Technology Center, Seattle: University of Washington*.

Wood, L. C., Teräs, H., & Reiners, T. (2013). The role of gamification and game-based learning in authentic assessment within virtual environments.

Yee, N. (2007). Motivations of play in online games. *Journal of CyberPsychology and Behavior*, *9*, 772–775.

Zhang, P. (2008). Motivational Affordances: Reasons for ICT Design and Use. *Communications of the ACM*, *51*(11), 145–147.

Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology*, *25*(1), 82–91.

Appendix A: Materials & Instruments

Pre-Pilot Observational Evaluation

1. Can the study variables be effectively and accurately measured from recording participant sessions? If not, how can the course or study protocol be revised to allow measurement?
2. Are participants able to complete the study protocol without confusion? If not, what are the factors causing confusion and how can they be addressed?
3. Are participants able to engage with course aspects pertaining to all four learning goals in the time allowed? If not, how does the course or protocol need to be revised?

Pre-Pilot Participant Interview Protocol

1. Do you feel the course properly familiarized you with the virtual environment to effectively participate in a virtual world learning experience?
2. Do you feel the course prepared you to properly move around in the virtual world?
3. Do you feel the course prepared you to orient yourself within the virtual world?
4. Do you feel the course prepared you to interact with objects in the virtual world?
5. Do you feel the course properly prepared you to communicate in the virtual world?
6. How do you feel about the level of support you received?
7. During the observation, the following events were noted as you attempted to complete the familiarization course. For each, please provide your perceptions of that part of the course [explain each event in turn and obtain the participant's feedback].

Participant Session Protocol: Treatment A (Pilot)

- ☐ Start *Second Life.* Set participant avatar at start pad of ground movement learning phase substation.
- ☐ Open Initial Participant Survey and Learner Satisfaction Survey in browser on study workstation.
- ☐ Seat participant at study workstation and ask them to complete Initial Participant Survey.
- ☐ Once survey is complete, change display to *Second Life* and provide basic controls sheet. Explain: "I would like you to complete a familiarization course in the *Second Life* virtual world. This is a three-dimensional virtual environment in which you control a virtual representation of yourself known as an avatar. You control the avatar using a combination of the keyboard for moving around and the mouse to click on things in the world. The course that you will complete is designed to familiarize you with skills you would need in order to participate in an online learning experience such as a lecture-driven class or a virtual training scenario where you might interact with virtual objects. On successful completion of the course and follow-up study requirements, you will receive a certificate of virtual world competence [show certificate]. The course has four sections that you will complete in this order: Movement on the ground by walking or running, movement through the air by flying your avatar around, interacting with virtual objects, and communicating with other users with text messages. Each course section has a learning phase where you can get used to the controls. You may practice as much as you like. When you feel you are ready, you can move on to the assessment phase where you can have your skill assessed by completing the subsection against the clock. For each section, you will have to complete the assessment subsection within a certain time in order to qualify. You can move on at any point with or without qualifying, but if you do not qualify on any one section, your lack of qualification will be noted on the competence certificate. You can choose to take a break between course sections or stop the familiarization course and quit at any time. There are additional instructions and tips displayed in the virtual world that you can read and use as you go. I will be able to help you during the learning phases but not the assessment phases. I will not be able to make any comments about your performance during the study session."
- ☐ Ask participant if they have any questions.
- ☐ Check video camera on observation workstation and begin recording.
- ☐ Invite participant to begin when they are ready.
- ☐ Assist participant in moving between course sections, but do not assist in completion of assessment subsections.
- ☐ When participant has completed all sections or has decided to stop, switch display to learner satisfaction survey on participant workstation and ask participant to complete it.
- ☐ Inform participant that they will be asked to complete an online survey about their confidence to complete tasks in *Second Life* in one week's time, and again a month later.

Participant Session Protocol: Treatment B (Pilot & Main Study)

- [ ] Start *Second Life.* Set participant avatar at start pad of ground movement learning phase subsection.
- [ ] Open Initial Participant Survey and Learner Satisfaction Survey in browser on study workstation.
- [ ] Seat participant at study workstation and ask them to complete Initial Participant Survey.
- [ ] Once survey is complete, change display to *Second Life* and provide basic controls sheet. Explain: "I would like you to complete a familiarization course in the *Second Life* virtual world. This is a three-dimensional virtual environment in which you control a virtual representation of yourself known as an avatar. You control the avatar using a combination of the keyboard for moving around and the mouse to click on things in the world. The course that you will complete is designed to familiarize you with skills you would need in order to participate in an online learning experience such as a lecture-driven class or a virtual training scenario where you might interact with virtual objects. On successful completion of the course and follow-up study requirements, you will receive a certificate of virtual world competence [show certificate]. The course has four sections that you will complete in this order: Movement on the ground by walking or running, movement through the air by flying your avatar around, interacting with virtual objects, and communicating with other users with text messages. Each course section has a learning subsection where you can get used to the controls. You may practice as much as you like. When you feel you are ready, you can move on to the assessment phase where you can have your skill assessed by completing the subsection against the clock. For each section, you will have to complete the assessment subsection within a certain time in order to qualify. Additionally, you can earn achievement badges for completing the assessments in exceptionally fast times beyond qualifying times. There are bronze, silver, gold and platinum badges available. Any achievement badges you earn will be shown on your certificate. You can move on at any point with or without qualifying or earning badges, but if you do not qualify on any one section, your lack of qualification will be noted on the competence certificate. You can choose to take a break between course sections or stop the familiarization course and quit at any time. There are additional instructions and tips displayed in the virtual world that you can read and use as you go. I will be able to help you during the practice phases but not the assessment phases. I will not be able to make any comments about your performance during the study session."
- [ ] Ask participant if they have any questions.
- [ ] Check video camera on observation workstation and begin recording.
- [ ] Invite participant to begin when they are ready.
- [ ] Assist participant in moving between course sections, but do not assist in completion of assessment subsections.
- [ ] When participant has completed all sections or has decided to stop, switch display to learner satisfaction survey on participant workstation and ask participant to complete it.
- [ ] Inform participant that they will be asked to complete an online survey about their confidence to complete tasks in *Second Life* in one week's time, and again a month later.

Revised Participant Session Protocol for Main Study Treatment A

The following language replaced the fourth checkpoint on the session protocol for Treatment A. This is in response to unintentional gamification of Treatment A in the pilot study.

☐ Once survey is complete, change display to *Second Life* and provide basic controls sheet. Explain: "I would like you to complete a familiarization course in the *Second Life* virtual world. This is a three-dimensional virtual environment in which you control a virtual representation of yourself known as an avatar. You control the avatar using a combination of the keyboard for moving around and the mouse to click on things in the world. The course that you will complete is designed to familiarize you with skills you would need in order to participate in an online learning experience such as a lecture-driven class or a virtual training scenario where you might interact with virtual objects. On successful completion of the course and follow-up study requirements, you will receive a certificate of virtual world competence [show certificate]. The course has four sections that you will complete in this order: Movement on the ground by walking or running, movement through the air by flying your avatar around, interacting with virtual objects, and communicating with other users with text messages. Each course section has a learning phase where you can get used to the controls. You may practice as much as you like. When you feel you are ready, you can move on to a more challenging phase that will give you the opportunity to test your abilities further. You can spend as long as you like on the second phase and move on at any time. You can choose to take a break between course sections or stop the familiarization course and quit at any time. There are additional instructions and tips displayed in the virtual world that you can read and use as you go. I will be able to help you during the learning phases but not the more challenging phases. I will not be able to make any comments about your performance during the study session."

Participant Experience Survey

Please rate your highest level of experience with the following:

1. A school, college or training class held in a multiuser virtual environment where the instructors and students appear as virtual representations of people (avatars) in a three-dimensional world:

   ☐ I am a current or recent (less than one year) user
   ☐ I have used it more than twice, but not in the past year
   ☐ I have tried it once or twice
   ☐ I have watched someone else use it
   ☐ I have no experience

2. The *Second Life* multiuser virtual environment for any purpose:

   ☐ I am a current or recent (less than one year) user
   ☐ I have used it more than twice, but not in the past year
   ☐ I have tried it once or twice
   ☐ I have watched someone else use it
   ☐ I have no experience

3. Any other non-game virtual world environments such as *OpenSim* or *IMVU* for any purpose:

   ☐ I am a current or recent (less than one year) user
   ☐ I have used it more than twice, but not in the past year
   ☐ I have tried it once or twice
   ☐ I have watched someone else use it
   ☐ I have no experience

4. Video games in which you control an on-screen representation of a person (avatar):

   ☐ I am a current or recent (less than one year) user
   ☐ I have used it more than twice, but not in the past year
   ☐ I have tried it once or twice
   ☐ I have watched someone else use it
   ☐ I have no experience

5. If you have experience with any of the items 1 to 4 above where you were controlling the avatar, what was the primary method you used to control the avatar? (if checking more than one, specify which experience for each control method)

   □ I have no experience controlling an avatar
   □ Computer keyboard
   □ Game controller (such as XBox or Playstation controller)
   □ I have used both keyboard and game controller
   □ Other (specify: ______________________________)

6. Experience communicating with other users in a multiuser online game or environment using the keyboard to type text messages:

   □ I am a current or recent (less than one year) user
   □ I have used it more than twice, but not in the past year
   □ I have tried it once or twice
   □ I have watched someone else use it
   □ I have no experience

7. Experience communicating with other users in a multiuser online game or environment using a headset (earphone and microphone):

   □ I am a current or recent (less than one year) user
   □ I have used it more than twice, but not in the past year
   □ I have tried it once or twice
   □ I have watched someone else use it
   □ I have no experience

Pre- and Post-Test for Task Self-Efficacy

If you were asked to perform certain tasks in the virtual environment *Second Life* right now, how certain are you that you could successfully do each task?

Please rate your degree of confidence by recording a number from 0 to 100 using the scale given below:

0	10	20	30	40	50	60	70	80	90	100
Cannot do at all					Moderately can do					Highly certain can do

	Task in *Second Life*	**Confidence (0-100)**
1.	Move your avatar around by walking	_______
2.	Perform other avatar movements like sidestepping	_______
3.	Move your avatar around by running	_______
4.	Move your avatar to a specific place	_______
5.	Move your avatar around by flying	_______
6.	Fly your avatar and land in a specific place	_______
7.	Maintain a good idea of where your avatar is within the virtual space	_______
8.	Pick up a virtual object	_______
9.	Move a virtual object around	_______
10.	Interact with a virtual object's menu	_______
11.	Chat to other users using text messages	_______
12.	Attend a virtual class	_______
13.	Move around in a virtual classroom	_______
14.	Ask questions in a virtual classroom	_______
15.	Interact with a virtual slideshow	_______
16.	Interact with a virtual experiment through objects and menus	_______
17.	Learn successfully in a virtual class	_______

Persistence Worksheet

Participant #: _____________ Gamified treatment? Yes No (circle one)

Count the number of times the participant attempts each task during each phase:

Spatial/movement tasks	Learning phase	Assessment phase	Total attempts
Ground course			
Flying course			

Object interaction tasks	Learning phase	Assessment phase	Total attempts
Select, interact, and move			

Communication tasks	Learning phase	Assessment phase	Total attempts
Text messaging			

Total overall attempts:	

Task Skill Worksheet

Participant #: ______________ Gamified treatment? Yes No (circle one)

Record task skill data for each task during the assessment phase only:

Skill station	**Attempts**	**Errors**	**Fastest Time**	**Highest achievement (if gamif.)**
Ground movement				
Flying movement				
Object interaction				
Communication				

Learner Satisfaction Survey

Please review the following statements. Circle the number that best represents your level of agreement for the familiarization course:

	Strongly disagree	Disagree	Undecided	Agree	Strongly agree
1. Course content was presented effectively	1	2	3	4	5
2. I learned about how to use the system	1	2	3	4	5
3. I enjoyed the experience	1	2	3	4	5
4. Using the system was easy for me	1	2	3	4	5
5. The activities were useful	1	2	3	4	5
6. There was a sufficient number of exercises	1	2	3	4	5
7. There was sufficient time to complete the exercises	1	2	3	4	5
8. My level of involvement was high	1	2	3	4	5
9. I would like to learn more about using the system	1	2	3	4	5
10. This was a worthwhile learning experience	1	2	3	4	5

Adapted from Domínguez et al. (2013)

Post-Session Qualitative Interview Protocol

Once the participant has completed the familiarization course, ask the following questions:

1. Tell me about your experience completing the familiarization course.

2. How important was it to you to be successful in the familiarization course?

3. I noticed that during ____________ part of the course you tried several times. What made you want to keep trying instead of moving on?

4. I noticed that during ____________ part of the course you only tried once/twice before moving on. What made you want to move on instead of keep trying?

5. Did you feel that the experience of going through the course was engaging? What made it engaging/not engaging?

6. Did you feel like the course was a type of game? If so, why? If not, why not?

7. Do you find yourself "making a game" of everyday activities to make them more interesting? If so, please describe some examples and talk about how frequently you might do this.

Use non-leading follow-up questions such as "tell me more about that" and "can you describe that for me?" to prompt more descriptive detail from participants.

Appendix B: MUVE Skill Task Analyses

Spatial Orientation and Movement

Second Life features three redundant sets of movement controls: Keyboard letter keys, keyboard arrow keys, and an on-screen, mouse-driven control panel. To ensure a consistent control modality between participants, only the keyboard letter keys were used.

1. Ground movement
    a. Walking
        i. Forward
            1. Keyboard
                a. Press W key
        ii. Backward
            1. Keyboard
                a. Press S key
    b. Rotating (turning)
        i. Turn right
            1. Keyboard
                a. Press D key
        ii. Turn left
            1. Keyboard
                a. Press A key
    c. Vertical movement
        i. Jump
            1. Keyboard
                a. Press E key
        ii. Crouch
            1. Keyboard
                a. Press and hold C key
    d. Sidestepping
        i. Sidestep right
            1. Keyboard
                a. Press and hold Shift + D key
        ii. Sidestep left
            1. Keyboard
                a. Press and hold Shift + A key
    e. Running
        i. Initiate run
            1. Keyboard
                a. Double-press and hold appropriate walking key
                    i. Double-press & hold W = run forward
                    ii. Double-press & hold S = run backward
                    iii. Double-press & hold D = run right
                    iv. Double-press & hold A = run left
2. Flying

a. Initiating and ending
    i. Keyboard
        1. Press F key to begin flying
        2. Press F key again to end flying
b. Direction
    i. Altitudinal
        1. Keyboard
            a. Press E to ascend
            b. Press C to descend
    ii. Lateral
        1. Keyboard
            a. Press W to fly forward
            b. Press S to fly backward
            c. Press D to turn right
            d. Press A key to turn left
            e. Press and hold Shift + D key to fly right
            f. Press and hold Shift + A key to fly left

Object Interaction

1. Object selection
    a. Click object in the world viewer to select it
2. Object function (when present)
    a. Click object to select
        i. Select function from pop-up menu
3. Object move
    a. Click and hold mouse pointer on movable object
    b. Drag mouse to move object in desired direction on lateral plane
        i. Hold CTRL key while dragging mouse to move object up and down
    c. Release mouse button to let go of object

Communication

1. Text-based chat
    a. Click Chat button at bottom of screen
        i. Type message in Nearby Chat text field in Conversations box
        ii. Press Enter when message is complete

Appendix C: Additional Data Tables

Response Variability Analysis for Self-Efficacy Instrument Validation

Descriptive Statistics for Pilot Pre-Test

	Mean	SD
Move by walking	67.14	29.79
Move by sidestepping, etc.	58.93	29.10
Move by running	60.00	28.82
Move to specific place	59.64	30.73
Move by flying	50.36	27.77
Fly to specific place and land	43.93	26.32
Maintain idea of where you are	60.00	32.76
Pick up virtual object	56.79	29.91
Move a virtual object around	53.57	28.25
Interact with virtual object’s menu	60.00	33.11
Chat using text messages	70.36	31.41
Attend a virtual class	65.00	29.29
Move around in virtual classroom	63.57	29.51
Ask questions in a virtual classroom	67.86	27.78
Interact with virtual slideshow	62.14	24.55
Interact with virtual experiment	58.57	27.76
Learn successfully in a virtual class	71.79	19.28

Descriptive Statistics for Pilot Post-Test (one week after study session)

	Mean	SD
Move by walking	87.86	19.29
Move by sidestepping, etc.	79.29	22.43
Move by running	85.00	22.79
Move to specific place	83.93	21.50
Move by flying	79.29	21.74
Fly to specific place and land	72.50	23.35
Maintain idea of where you are	83.79	15.56
Pick up virtual object	88.93	16.43
Move a virtual object around	87.50	17.18
Interact with virtual object's menu	87.36	19.40
Chat using text messages	88.21	23.66
Attend a virtual class	87.86	20.82
Move around in virtual classroom	87.14	20.91
Ask questions in a virtual classroom	88.21	24.62
Interact with virtual slideshow	83.57	21.07
Interact with virtual experiment	82.36	19.46
Learn successfully in a virtual class	86.07	21.32

Appendix D: Open Coding of Post-Session Interview Responses

Question 1

Tell me about your experience completing the familiarization course.

Open codes generated from responses to this question:

- Gamified group only:
    - *Negative attitude*: Expression of a negative perception of the course.*
    - *Navigation & Pathfinding*: Self-orienting within the MUVE.
    - *Persistence*: Trying again or giving up.
    - *Depth perception*: The participant's ability to judge distance.
    - *Agency*: Decision-making and being in control.
    - *Visual appearance*: The attractiveness of the virtual world.
    - *Skill improvement*: How skills were developed during the course.
- Both groups:
    - *Positive attitude*: The participant expressed a positive perception of the course.
    - *Level of difficulty*: The ease or difficulty of the course.
    - *Usefulness*: Whether or not the course was useful.
    - *Self-evaluation*: Own abilities or performance.
    - *Expectations*: Comparison of actual experience with expected experience.
    - *Affordances*: Perceptions of the controls and instructions.
    - *System responsiveness*: The performance of the system.
    - *Learning*: How the course leads to, or enables, learning.
    - *Familiarity*: Comparison to other systems such as video games.

* Negative perceptions were expressed in the non-gamified group, but only once, compared to ten times in B.

Question 2

How important was it to you to be successful in the familiarization course?

The open codes generated from responses to this question were focused on the perceived importance of success in the course. The expressions used are listed here in descending order of emphasis (the strength of language used):

- *Extremely important.*
- *Very important.*
- *Really important.*
- *Important.*
- *Pretty important.*
- *Moderately/fairly/mildly important. Semi-important.*
- *Not that important.*

There was no apparent difference in the distribution of importance perceptions between the two study groups.

Question 3

I noticed that during ____________ part of the course you tried several times. What made you want to keep trying instead of moving on?

Open codes generated from responses to this question:

- Non-gamified group only:
  - *Seeking greater understanding*: Curiosity about how system or course worked.
  - *Ensuring competence*: Wanting to perform well at assessment phase or next try.
  - *Enjoyment*: Process was fun so did not wish to move on.
- Gamified group only:
  - *Personal values*: Intrinsic perception that tenacity is important/valuable.
  - *Seeking reward*: Anticipation of receiving an award (leaderboard or badge).
- Both groups:
  - *Wanting to improve*: Desire to increase competence and/or correct mistakes.
  - *Responding to progress*: Noticing increase in skills and wanting to build on it.
  - *Avoiding failure or consequences*: Desire to repeat until success was achieved.
  - *Meeting expectations*: Expectation was to do well, which required repetition.

Question 4

I noticed that during ____________ part of the course you only tried once/twice before moving on. What made you want to move on instead of keep trying?

Open codes generated from responses to this question:

- Non-gamified group only:
  - *Uncertainty about expectations*: Did not know whether to try again.
  - *Time concerns*: Did not want to take too long to complete course.
  - *Met expectations*: Performed well enough to make repetition unnecessary.
  - *Curiosity*: Wanted to move on to next section out of curiosity.
  - *Lack of incentive*: No perceived reward for persistence.
- Gamified group only:
  - *Difficulty*: Found it too difficult so gave up.
  - *Lack of challenge*: Found it too easy for repetition to be necessary.
- Both groups:
  - *Reached maximum competence*: Continued repetition perceived as unnecessary.
  - *Satisfied with performance*: Did not need to do any better.
  - *Benefit from continuing*: Did not think further attempts would improve skill.
  - *Success perceptions*: Once successful, repetition considered unnecessary.
  - *Frustration*: Gave up due to feelings of frustration.

Question 5

Did you feel that the experience of going through the course was engaging? What made it engaging/not engaging?

Open codes generated from responses to this question:

- Non-gamified group only:
  - *Anticipation of assessment*: Motivated to focus to gain skill for assessment.
  - *Challenge*: The challenge of the tasks increased engagement.*
  - *Availability of feedback*: Kept engagement level high.
  - *Visual appeal*: MUVE was visually attractive, which helped maintain interest.
- Gamified group only:
  - *Anticipation of reward*: Earning rewards kept engagement level high.
  - *Comparison to real life*: Similarity to real life activities.
  - *Curiosity*: Desire to see the rest of the course maintained engagement.
  - *Presence of goals*: Striving for goals made the course engaging.
- Both groups:
  - *Agency*: Perceptions of decision making and being in control.
  - *Enjoyment*: The course was fun to complete.
  - *Availability of instructions and help*: Contributed to engagement.
  - *Learning process*: The perception of learning skills was engaging.
  - *Wanting to succeed*: Motivated by desire to succeed in the course.
  - *Interest level*: Found course interesting.
  - *Novelty*: Course was new, and therefore engaging.
  - *Involvement/interaction*: Course required high engagement for success.

- *Immersion/presence*: Feeling of “being there” was engaging.

* Challenge was mentioned as a factor once in the gamified group, compared to five times in the non-gamified group.

Question 6

Did you feel like the course was a type of game? If so, why? If not, why not?

Open codes generated from responses to this question:

- Non-gamified group only:
  - *Insufficient challenge*: The tasks undermined feelings that the course was a game.
  - *Training*: Felt more like training to get ready for a game.
- Gamified group only:
  - *Reward system*: The fact that rewards were available made it feel like a game.
  - *Timer*: Presence of timer made it feel like a game.
- Both groups:
  - *Challenging tasks*: The presence of tasks and assessment made it feel like a game.
  - *Similarity to video games*: It looked similar to common video games.
  - *Win/loss*: Presence of win/loss states add to a game-like feel.

Question 7

Do you find yourself "making a game" of everyday activities to make them more interesting? If so, please describe some examples and talk about how frequently you might do this.

Open codes generated from responses to this question:

- Gamified group only:
  - As a child doing chores
  - Out of competitiveness
- Both groups:
  - *With others*: Creating impromptu games to motivate others in mundane tasks.
  - *During chores*: Making a game of cleaning, tidying, etc.
  - *At work*: Setting challenges for self in the workplace.
  - *While exercising*: Using timers and achievement levels while working out.
  - *While cooking*: Incorporating game-like elements to make it more fun.
  - *While studying*: Rewarding writing long papers or memorization exercises.
  - *Boredom relief*: Inventing game mechanisms to complete tedious tasks.
  - *Providing self with reward*: A purchase or food item as reward for success.

There was no apparent difference between the two study groups in the distribution of affirmative and negative responses to whether the participant exhibits this behavior.

Vita

P. "Asikaa" Cosgrove was born in Bedfordshire and raised in Hertfordshire, England. He graduated with distinction from Brunel University in west London and worked as a systems designer for British Telecommunications. After immigrating to the United States of America in 2001, he earned his M.Ed. degree in Instructional Technology (2007) from Drury University in Springfield, Missouri. He earned his Ph.D. in Information Science and Learning Technologies (2016) from the University of Missouri-Columbia, minoring in educational psychology.

Dr. Cosgrove is currently a faculty member and the director of graduate programs in the School of Education and Child Development at Drury University, where he teaches educational technology and psychology. His research interests include teaching and learning in virtual worlds, cognition, and the psychology of motivation.