THE SUBJECTIVE GAMEPLAY EXPERIENCE:
AN EXAMINATION OF THE REVISED GAME ENGAGEMENT MODEL

by

KATELYN C. PROCCI
B.S. Virginia Commonwealth University, 2009
M.S. University of Central Florida, 2013

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Psychology
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Spring Term
2015

Major Professor: Clint Bowers

© 2015 Katelyn C. Procci

ABSTRACT

The study of the subjective gameplay experience spans multiple disciplines, from teachers who want to harness the power of gameplay to enhance instruction to game developers hoping to create the next big hit. Despite decades of interest, little agreement has been found regarding the way constructs—such as immersion, involvement, presence, and flow—are used to describe the subjective gameplay experience. Without the consistent usage of well-defined constructs, it becomes impossible to further scientific understanding of this domain. This dissertation examined the theoretical evolution of the key subjective gameplay experience constructs. From this, definitions for immersion, involvement, presence, and flow were extracted. Based on the prior work of Brockmyer et al. (2009), a revised game engagement model was created that incorporated these definitions. To test the proposed relationships within the revised game engagement model, experienced players of the computer game *Minecraft* were recruited for an experimental study. The participants played the game *Minecraft*, which was manipulated with respect to both level of difficulty and immersive aspects. This allowed for a range of potential game engagement states to be experienced by the participants. Several individual differences hypothesized to influence the different constructs of game engagement also were measured. The results of the study supported many proposed aspects of the revised game engagement model and revealed ways in which the model could be further refined. The theoretically-derived definitions and revised game engagement model resulting from this work, along with the suggested measures for these relevant constructs, provides a framework for future work in this area. This framework will improve the consistency of construct operationalization, benefiting the continued study of the subjective gameplay experience.

To my parents, Kathy and Mike Procci, for their tireless support through the years.

ACKNOWLEDGMENTS

I would like to thank my advisor Clint Bowers for giving me all of the opportunities that got me to where I am today (and for providing equipment, resources, and space!). I also owe thanks to all of those who have served as mentors along the way: Jan Cannon-Bowers, Jessie Chen, and Mike Barnes. I would like to thank my committee members—Florian Jentsch, Valerie Sims, and Rudy McDaniel—for letting me run with my study idea and for providing input. And how could I forget Haydee Cuevas, who only had encouraging words every time I ran into her at HFES?

I could not have completed my dissertation without the help of three dedicated research assistants: Christine Kreutzer, David Garcia, and Cierra Godwin. Special thanks to Jeremy Flynn, who introduced me to the wonderful world of *Minecraft* modding and for providing countless hours of technical support while I built the computer and modified the software I needed for the study.

Thank you to *all* of my undergraduate research assistants and fellow labmates at RETRO for the moral support over the past several years as I completed my journey through graduate school. In particular, I would like to thank Jenny Vogel, Leah Kuffner, Jen Loglia, Skilan Ortiz, Asli Soyler-Akbas, Alen Chao, Shan Lakhmani, Lucas Blair, and Danielle Chelles.

Finally, thank you so much to my parents, Mike and Kathy Procci. I spent more than one Christmas at home, working non-stop on my dissertation—thank you for understanding and providing a quiet space where I could get an enormous amount of work done. I could not have done it without you both.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

ANOVA	analysis of variance
BIP	break in presence
C FSS	CORE Flow State Scale
CAS	Cognitive Absorption Scale
CAVE	Cave Automatic Virtual Environment
CFA	confirmatory factor analysis
CIS	Creative Imagination Scale
DES	Dissociative Experiences Scale
DFS	Dispositional Flow Scale
DFS-2	Dispositional Flow Scale – 2
EFA	exploratory factor analysis
ERF	egocentric reference frame
ESA	Entertainment Software Association
FIVE	Framework for Immersive Virtual Environments
FPS	frames per second
FSS	Flow State Scale
FSS-2	Flow State Scale – 2
GB	gigabyte
GEQ	Game Engagement Questionnaire
GLSL	OpenGL Shading Language
GSR	galvanic skin response

HD	high definition
HDMI	high-definition multimedia interface
HMD	head-mounted display
HP	Hewlett Packard
IBM	International Business Machines
IRB	Institutional Review Board
ITQ	Immersive Tendencies Questionnaire
LCD	liquid crystal display
M	Mean
MC	Minecraft
MPQ	Multidimensional Personality Questionnaire
MSQ	Motion Sickness Questionnaire
PC	personal computer
PERF	primary egocentric reference frame
PII	Personal Involvement Inventory
PQ	Presence Questionnaire
PVR	personal video recorder
R PII	Revised Personal Involvement Inventory
RAM	random access memory
RETRO	Recent and Emerging Technologies Research Organization
R-GEM	revised game engagement model
RPG	role-playing game

S DFS	SHORT Dispositional Flow Scale
S FSS	SHORT Flow State Scale
SD	standard deviation
SDT	Signal Detection Theory
SEM	structural equation modeling
SEUS	Sonic Ether's Unbelievable Shaders
SPSS	Statistical Package for the Social Sciences
SSM	spatial situation model
SSQ	Simulator Sickness Questionnaire
SUS	System Usability Scale
TAM	Technology Acceptance Model
TAS	Tellegen Absorption Scale
TB	terabyte
UCF	University of Central Florida
USB	universal serial bus
UTAUT	Unified Theory of Acceptance and Use of Technology
VE	Virtual Environment
VGSE	video game self-efficacy
VGSES	Video Game Self-Efficacy Scale
VRAM	video random access memory

CHAPTER ONE: INTRODUCTION

Salen and Zimmerman (2004) define a game as "a system in which players engage in an artificial conflict, defined by rules, that results in a quantifiable outcome" (p. 80). Games are similar to simulations, and researchers Garris, Ahlers, and Driskell (2002) summarized the difference between simulations and games thusly: Both simulations and games involve rules and the user's actions result in consequences that are contained within their virtual worlds. According to their definition, simulations represent real-world systems while games do not. While simulations can be game-like, the researchers found that games possessed the following features: "fantasy, rules/goals, sensory stimuli, challenge, mystery, and control" (Garris et al., 2002, p. 443). Marc Prensky (2001) raised an interesting point: A game can *also* represent a real-world system. Prensky (2001) argued that a simulation is any sort of synthesized reality driven by "a mathematical or algorithmic model, combined with a set of initial conditions, that allows prediction and visualization as time unfolds" (p. 211). Also, a simulation cannot be 'won'. A simulation can be transformed into a game as long as the designer has added in certain structural elements, including "fun, play, rules, a goal, winning, competition, etc." (Prensky, 2001, p. 211). Thus, not all simulations are games, but they can become games if competition—either against others or the system itself—for achieving a goal-state is introduced, and all games are simulations. Both are representations of some sort of system with implicit rules with consequences for players' inputs. The focus of this work is on video games, which are electronic games played on a medium such as a dedicated gaming console or a computer. Any future mention of games or gaming is in reference to video games.

Nearly every single journal article, conference proceedings paper, and dissertation about video games opens in the same way: Video gaming is a popular, multi-billion dollar industry. According to the 2012 Entertainment Software Association (ESA) report—which is the gold standard yearly report for information pertaining to the gaming industry—game-related sales, including hardware and accessories, topped $24.75 billion in 2011. The same report also found that of the 2,000 households surveyed, every house owned at least some sort of gaming device, such as a smart phone or a personal computer. Furthermore, nearly half of the respondents reported that they owned a dedicated gaming console. The ages of the game players in those households were well-distributed, where 32% of the players were under the age of 18, 31% of the players were between 18 and 35, and 37% of the players were 36 years old or older. Even gender was nearly evenly divided, where players were 53% male and 47% female. Given the widespread presence of gaming equipment in households and the diverse nature of gamers themselves, the findings of the ESA Report supported the notion that gaming in the United States is indeed ubiquitous. Furthermore, this phenomenon is not isolated to this country. The gaming industry was estimated to be worth nearly $67 billion worldwide in 2012, and is expected to reach $82 billion by 2017 (Gaudiosi, 2012).

Aside from being a part of a profitable entertainment industry, the use of games for purposes other than enjoyment, such as for training and education, advertising, and raising awareness, has risen dramatically in recent years. Effective examples of these so-called *serious games* are found across many diverse fields to achieve a variety of goals, from promoting positive self-care behaviors in pediatric oncology patients (Kato, Cole, Bradlyn, & Pollock, 2008) to raising awareness of world events, such as the war in Darfur (Brown, 2007). There is some disagreement as to the definition of a serious game. For example, as summarized in Girard, Ecalle, and Magnan (2012), some consider any game to be a serious game as long as it can be used for some sort of purpose other than enjoyment alone, while others (Girard et al., 2012 included) staunchly believe that a serious game is only one that was designed with the specific intent to achieve some goal aside from pure entertainment. Regardless, serious games have been found to be very effective training tools, particularly when they augment existing instruction (Sitzmann, 2011).

Given the popularity of entertainment games as ascertained by a booming gaming industry, alongside the evident utility of serious games to achieve non-entertainment ends, research in both areas has expanded dramatically. Research regarding entertainment games focuses primarily on the underlying motivations of gamers as well as on the subjective experience of play (Boyle, Connolly, Hainey, & Boyle, 2012). Research on serious games focuses on whether games can be used to achieve a range of outcomes, from declarative knowledge gains to motor skill acquisition, as well as best design practices regarding both pedagogy and affective responses (Connolly, Boyle, MacArthur, Hainey, & Boyle, 2012). Both streams of research are similar in that they explore how high-quality, enjoyable experiences can improve games, either for financial reasons as good games make for a better bottom-line in the entertainment industry, or to better understand how gameplay can be shaped to result in improved instructional outcomes.

But what exactly is this sought-after game experience? The subjective gameplay experience has been appraised using several different constructs, such as immersion, presence, and flow (Boyle et al., 2012). Flow, for example, is quickly rising to be one of the favored constructs used to describe the gaming experience. Flow describes the subjective, enjoyable experience in which one becomes fully absorbed in an activity to the point where s/he loses awareness of both the self and the passage of time (Csikszentmihalyi, 1975a, 1975b, 1990). Game researchers are particularly interested in flow, as the typical description of the gamer, whose one-more-level mentality often keeps them playing far later into the night than they had originally intended, aligns very well with what Csikszentmihalyi (1975a, 1975b, 1990) described as the flow state. Still, Boyle et al.'s (2012) review found that "there is still a lack of consensus about how best to characterise subjective experience in games" (p. 778). This has been reflected in the measures used to quantify the game experience:

> Therein lies the potential problem; there seem to be many different words used to describe what might be the same construct. For example, one flow scale for gaming combines the time transformation and loss of self-consciousness aspect of flow into one subscale labeled "immersion" (e.g., "I forget about time passing while playing the game" or "I become unaware of my surroundings while playing the game," Fu, Su, & Yu, 2009), while another makes a distinction between their flow items and their single item for immersion ("I really get into the game," Brockmyer et al., 2009). (Procci & Bowers, 2011, p. 2183)

Definitional issues plague many constructs, resulting in a number of dense theoretical papers whose purpose is to sort out the issue. Sometimes those papers are successful: Mica Endsley tackled the problem of defining situation awareness because "unless researchers stick to a clear, consistent meaning for the term, the problem will present a significant handicap to progress" (Endsley, 1995, p. 36). She was successful in her goal of setting a strict yet robust definition of situation awareness that has since become generally accepted. Such progress has yet to be made regarding the subjective state of gameplay (Brockmyer, Fox, Curtiss, McBroom, Burkhart, & Pidruzny, 2009). It is necessary to establish clear, consistent definitions for the aspects of the subjective gameplay experience in order to advance the science of games.

The goals of this work are to: (1) progress chronologically through the literature examining constructs relevant to the subjective gameplay experience, specifically immersion, presence, flow, and absorption; (2) extract the most well-founded and validated definitions for these constructs; (3) develop a simplified model for the subjective gameplay experience that consolidates these concomitant theoretical models; (4) identify potential methods in which each of these constructs may be accurately measured; and, (5) empirically examine whether that model encapsulates the subjective gameplay experience. This is a unique contribution to the science of games in that it sets clear definitions and defines relationships between these distinct—but often confused—constructs relevant to the subjective gameplay experience, in an effort to provide a unified understanding of these aspects, thus paving the way for congruous future research.

CHAPTER TWO: LITERATURE REVIEW

The first task is to determine the relevant constructs that the subjective gameplay experience entails. Immersion seems to be an important concept to the gaming community, as it is often used in game reviews to describe realism and is linked to enjoyment and quality in this context (Brown & Cairns, 2004). Flow and absorption also are candidates worth examining based on the work of other gaming researchers (Brockmyer et al., 2009).

The scientific literature regarding both immersion and its sister concept, presence, has been devoted to frustratingly circular definitional issues for the previous two decades. There is a general consensus that presence in a virtual environment (VE) is the feeling of “being there” (Sheridan, 1994, p. 1073), to the extent that the virtual world becomes more salient and realistic than the current physical reality. Still, this general definition struggles because it fails to be explicit from other similar concepts (Wirth et al., 2007). There is even less clarity regarding the concept of immersion. Some have purported that immersion is a term that should be used to describe the VE technology and its propensity to encourage the sensation of presence (e.g., Slater & Wilbur, 1997), but immersion also has been described as a subjective experience of being enveloped by a VE’s stimuli (Witmer & Singer, 1998).

From the early theoretical work of the 1980s, followed by the initial definitions proposed in the 1990s, the usage of both immersion and presence has slowly mutated over time. The terms, along with many others that describe engaging media experiences such as flow, have been used interchangeably and with total disregard of their original definitions (Procci, Singer, Levy, & Bowers, 2012). This trend of word usage does not reflect the advancement of theory, but rather a misunderstanding of constructs.

To disentangle these definitional issues, it is best to start at the beginning. The following review highlights the major theoretical works in the immersion and presence literature by directly quoting their operationalizations and consolidating their definitions. In addition to immersion and presence, other relevant constructs, specifically flow and absorption as suggested by Brockmyer et al. (2009) as well as involvement (Witmer & Singer, 1998), will be discussed as they apply to the subjective gameplay experience. Definitions for each construct are provided that are clear and distinct. Measurement methods for each construct are also discussed. This understanding culminates in a testable model of the subjective gameplay experience, known as the revised game engagement model (R-GEM).

The 1990s: Defining immersion and presence

While much of the research regarding presence occurred in the 1990s, the concept has its roots in the preceding decade. In 1980, Marvin Minsky introduced the idea of telepresence, in which a teleoperator feels as if they are physically present at another location than where s/he currently is, which may have implications for his or her performance (Draper, Kaber, & Usher, 1998). Thomas B. Sheridan (1992) expanded on this notion and set the stage for early theoretical work regarding presence. He formally defined the term *telepresence* as the “sense of being physically present with virtual object(s) at the remote teleoperator site” (Sheridan, 1992, p. 120) and *virtual presence* as the “sense of being physically present with visual, auditory, or force displays generated by a computer” (Sheridan, 1992, p. 120). He proposed that both telepresence and virtual presence are subjective states that occur when a user is provided with a high-fidelity display, a willingness to devote attentional resources to the task environment, and a level of motor-based interaction with that environment.

He pointed out that there was a dearth of research regarding presence in the early 1990s: no real theoretical work regarding the presence construct had been conducted, a valid measure of presence had not been established, and there had been no empirical examination as to how presence affected performance and training. For the study of presence to move forward, he proposed that a strong operational definition and a reliable, useful method of measurement must be developed. Sheridan (1992) proposed that since presence is subjective, the measure should incorporate subjective assessment. He also stated that an additional objective measurement will serve to increase the strength of the measure, citing Held and Durlach (1987, as cited in Sheridan, 1992), who suggested that including stimuli within the VE that would cause the user to react, such as a virtual ball flying toward the user's head, would be a way to obtain objective behavioral responses as an indication of the experience of presence. Furthermore, similar to the constructs of workload and mental models, he suggested that presence is likely not unidimensional.

Sheridan (1992) proposed three dimensions the promote presence: That the environment provides ample sensory information that is perceived by the user; that the user is able to control and use senses within the environment; and that the user's actions are able to modify the environment. He also believed that elements of the task itself influenced presence, including task difficulty as quantified by Fitts' index (1954, as cited in Sheridan, 1992), and degree of automation (e.g., is walking in the environment automated in such a way that the user cannot control where they go?). He proposed that presence has three outcomes, including a measureable sense of presence that is both subjective and objective, training efficiency, and improved task performance.

Sheridan (1994) expanded on his earlier work by proposing more detailed ways as to how presence could be measured. He believed that while telepresence and virtual presence are part of different contexts, the subjective experience is essentially the same. Given this inherent equivalency, future references to the term *presence* in this work refer to Sheridan's (1992) notion of virtual presence. First, Sheridan (1994) proposed that presence should make use of objective measurement through the use of reflexes, as previously suggested (Sheridan, 1992). Second, presence should be measured using a subjective rating scale that is multi-dimensional. He also proposed that his three factors that contribute to presence (sensory information, control, and interactivity) are orthogonal, and that three to five items should be created for the user to rate for each factor. As an important note, Draper et al. (1998) later argued that these three factors are not orthogonal given that all of the factors will interact with one-another in a VE (e.g., how a user will continue to interact with the environment depends on the sensory feedback they receive). Third, presence might be able to be assessed using an image discrimination task to determine whether degraded images are perceived as being part of a real or virtual environment. This is later argued by IJsselsteijn, Ridder, Freeman, and Avons (2000) to not be a viable measure of presence as it likely only measures image discriminability.

Sheridan (1994) also expanded on the ways that each of the three orthogonal factors could be manipulated experimentally regarding magnitude, time, and space. For example, visual perception information can be manipulated with respect to resolution (space), frame rate (time), and color bits per pixel (magnitude), which should have an effect on presence, where *more* is essentially more realistic, and thus more likely to result in the experience of presence. For sensory control, introducing latency, such as a few seconds passing between the user's input and the response in the VE, or increasing video jerkiness, will decrease presence. Finally, regarding interactivity, the environment's reactions should be realistic and of the appropriate magnitude for the user's input.

It is from Sheridan's groundwork that Mel Slater and his colleagues focused their initial research efforts, which would set the stage for future definitions and debate about immersion and presence. Slater, Usoh, and Steed (1994) asserted that immersion is a descriptive quality of VE technology that promotes a sense of immersing a user. Immersion, they suggested, may lead to presence, which is the sense of physically existing within the VE. In this early work, they proposed several external characteristics that would make a system more immersive, such as expanding the field of view, incorporating auditory stimuli, increasing the level of interactivity, modeling realistic behavior of objects in the VE, the inclusion of a virtual representation of the player as an avatar, which they referred to as a "virtual body" (Slater et al., 1994, p. 131), and some amount of matching between the user's input and the avatar's actions within the VE.

Slater et al. (1994) also conducted a study in an effort to determine how several subjective factors, including perception of the visual, auditory, and kinesthetic elements of a VE, influenced presence while also manipulating several system characteristics purported to increase immersion. These system characteristics included whether the VE had gravity, the existence of a virtual cliff, whether there was another virtual actor within the environment that followed the participant, and stacking depth. Stacking depth refers to an increasing depth of presence as a result of how many scenarios the participant experienced within the VE. Each scenario required that the participant enter a *new* virtual environment while already in a virtual environment (essentially, a VE within a VE – a concept not dissimilar to the 2010 film *Inception*). These transitions were manipulated in an effort to increase presence in that the participant either entered the new scenario by walking through a virtual door, which is essentially like moving to a new room, or donning a virtual head-mounted display (HMD), which is a display device that is worn by real-world VE users that fits a display over the eyes and blocks out the physical world entirely. The researchers hypothesized that including or increasing these particular system characteristics would increase presence. Ultimately, the goal of their work was to create a mathematical equation involving these elements to predict the amount of presence experienced. The results suggested that the amount of presence experienced was positively associated with subjective ratings of the visual and kinesthetic aspects of the system, yet negatively associated with auditory ratings. They also found that the number of transitions to deeper VEs was positively associated with presence when a virtual HMD was used, but negatively associated with presence when the transition occurred between doors. None of the other factors were found to be significant. This work suggested that elements of the VE system have an impact on presence. In particular, visual and kinesthetic experiences were most influential in creating presence.

Draper et al. (1998) created a summary of the theoretical work for the construct of telepresence, mainly focusing on the developments of the 1990s. Similar to Sheridan (1994), they noted that presence and telepresence apply to different contexts, but are the same construct related to the "displacement of the user's self-perception into a computer-mediated environment" (Draper et al., 1998, p. 354). They defined immersion, by citing Biocca and Delaney (1995), as "the degree to which a virtual environment submerges the perceptual systems of the user in computer-generated stimuli" (p. 57, as cited in Draper et al., 1998, p. 356). The authors organized all previous work regarding telepresence into two categories: The technological approach, termed “cybernetic presence”, and the psychological approach, termed “experiential presence” (Draper et al., 1998, p. 356).

The technological approach largely focused on how characteristics of the system affect telepresence and operator performance in remote environments. Draper et al. (1998) summarized that the common technological factors to improve telepresence include the fidelity of the synthetic operating environment as well as the operator’s ability to control and interact with that environment. They also stated that telepresence is likely influenced by the user’s individual differences when using the technology and highlights the importance of naturalistic, spatially-correct, and easy-to-use input devices, as well as high-quality feedback to keep the operator informed on their progress in the remote environment. Slater et al.’s (1994) research would be considered as having adopted the technological approach to presence.

Regarding the psychological approach, this stream of research was in its infancy at the time, although it would later dominate discourse in the scientific literature about the construct. Draper et al. (1998) suggested that it would be most pertinent to ground further examination of telepresence within established psychological theories. They suggested that research efforts should focus on: the “psychocybernetic” (Draper et al., 1998, p. 362), which involved operator-environment feedback loops and stressed the importance of easy-to-use and predictable controls; flow theory, which is the idea of the optimal experience introduced by Csikszentmihalyi (1975a); distal attribution (Loomis, 1992, as cited in Draper et al., 1998), which is the psychological phenomena of taking external objects and associating them with oneself; and situation awareness, which refers to an operator’s ability to attain a “state of knowledge” in a dynamic environment due to the ability to perceive important elements and critical cues, comprehend what those elements and cues mean when taken together, and predict near-term future states of the system in order to make effective decisions (Endsley, 1995, p. 36). The authors even go as far as to say that situation awareness to the fullest extent is telepresence. This may not be exactly true. The fullest extent of situation awareness involves the ability to predict the future state of a *complex* system. It is not easily achievable and requires a large amount of mental resources to process the interactions of several variables in order to predict an outcome. Feeling present in a VE may be difficult to achieve, however it does not require this amount of mental effort and is often not couched in a decision-making context, although the ability to successfully predict what will happen next does was proposed as an element that promotes presence (e.g., Witmer & Singer, 1998). This, however, is more related to whether an environment is logical and coherent in that the VE behaves in a rational way that does not violate expectations. Being able to predict that a dropped object will fall in a terrestrially-based VE is not the highest level of situation awareness: it is instead the implementation of the relevant physics that makes the environment seem natural, thus supporting the experience of presence.

Finally, citing Psotka and Davison (1993), Draper et al. (1998) made it clear that the user's individual differences must be accounted for, including the likelihood to experience telepresence (which is later a requirement highlighted by both Witmer & Singer, 1998, and Wirth et al., 2007) and their perception of the immersive qualities of the system, such as whether they experienced simulator sickness due to the system characteristics.

Two of the authors included in the Draper et al. (1998) summary would become the two major contenders in the 1990s regarding the definition of immersion. Slater and Wilbur (1997) would come to champion the cybernetic (technological) approach while Witmer and Singer (1998) would instead focus on the experiential (psychological) approach. Both camps focused on their respective approaches exclusively, despite Draper et al.'s (1998) insistence that the two should be considered together.

Cybernetic approach: Slater & Wilbur, 1997. Slater and Wilbur (1997), in association with the framework for immersive virtual environments (FIVE) working group whose purpose was to research presence, built upon the previous work of Slater et al. (1994). The researchers reiterated that immersion describes the characteristics of VE technology that can be manipulated to create a sense of presence in the users. They explicitly defined immersion and presence as such:

- Immersion is "a description of a technology, and describes the extent to which the computer displays are capable of delivering an *inclusive*, *extensive*, *surrounding* and *vivid* illusion of reality to the senses of a human participant" (Slater & Wilbur, 1997, pp. 604-605).
- Presence is "a state of consciousness that may be concomitant with immersion, and is related to a sense of being in a place" (Slater & Wilbur, 1997, p. 603).

Slater and Wilbur (1997) further explicated on the system characteristics that are the factors of immersion as stated in their definition:

- Inclusive "indicates the extent to which physical reality is shut out" (Slater & Wilbur, 1997, p. 605).

Essentially, increasing the inclusive nature of the VE is achieved by eliminating outside distractions, which in turn increases immersion. Citing a previous study conducted by Slater and

Usoh (1992, as cited in Slater & Wilbur, 1997), outside distractions, such as the experimenters' voices, reduced presence. Patel (1994, as cited in Slater & Wilbur, 1997) also found that playing white noise, as opposed to no noise, while in a VE sufficiently blocked out the real world and more presence was experienced. They also stated that it is important to reduce the cues that the technology responsible for generating the VE exists. For example, a heavy HMD will be more disruptive to presence than a lighter, less obtrusive one. Another consideration to minimize cues concerns the capabilities of the system itself. For example, Slater and Usoh (1992, as cited in Slater & Wilbur, 1997) found that poor resolution made users feel less present. Additionally, Barfield and Hendrix (1995, as cited in Slater & Wilbur, 1997) found that higher frame rates resulted in higher reports of presence, and that there seemed to be a critical point where the frame rate should be at least 15 to 20 Hz for presence to occur.

- Extensive "indicates the range of sensory modalities accommodated" (Slater & Wilbur, 1997, p. 605).

In theory, increasing the number of the senses modeled should increase immersion (e.g., a VE that includes visual and haptic aspects should be more immersive than one that only had visual aspects). For example, Hendrix and Barfield (1995; 1996b, as cited in Slater & Wilbur, 1997) found that including spatialized audio in a VE increased presence.

- Surrounding "indicates the extent to which this virtual reality is panoramic rather than limited to a narrow field" (Slater & Wilbur, 1997, p. 605).

Essentially, a wider field of view is more immersive. Hendrix and Barfield (1995; 1996a, as cited in Slater & Wilbur, 1997) found that increasing field of view to 50° and 90° from 10° increased presence. One can imagine that incorporating a HMD or Cave Automatic Virtual

Environment (CAVE) system—which is simulator that completely surrounds an individual, often with screens that fill the space from floor to ceiling (Cruz-Neira, Sandin, DeFanti, Kenyon, & Hart, 1992) —would result in more presence than a VE system displayed on a single monitor.

- Vivid "indicates the resolution, fidelity, and variety of energy simulated within a particular modality (for example, the visual and color resolution). Vividness is concerned with the richness, information content, resolution and quality of the displays" (Slater & Wilbur, 1997, p. 605).

The concept of vividness suggests that realism promotes immersion. For example, Welch et al. (1996, as cited in Slater & Wilbur, 1997) found that increasing realism in a driving simulator increased presence. Also, including more realistic visual aspects, such as dynamic shadows, also has been found to increase presence (Slater, Usoh, & Chrysanthou, 1995, as cited in Slater & Wilbur, 1997). The inclusion of realistic physics may also contribute to presence. Uno and Slater (1997, as cited in Slater & Wilbur, 1997) found that realistic physics were significantly and positively correlated with presence for friction, but elasticity and collision were not found to have a significant association. Although it seems that a large emphasis from the above studies was placed on realism, Slater and Wilbur (1997) were very clear in stating that "presence does not imply realism" (p. 609), meaning it may not be so important that the vivid environment looks as real as possible, but that its realistically complex and cohesive in its own right. For example, a cartoon world VE would not look photorealistic, but as long as the world was vivid, perhaps through high-resolution textures and careful attention to detail, and all aspects were coherent, it would be realistic in its own right and could still support the experience of presence.

Furthermore, according to Slater and Wilbur (1997), there are two fundamental aspects that the VE must have for immersion to occur, which are self-representation and matching. Self-representation is the capacity in which the individual is represented within the VE by a player-controlled avatar, and that the user can use the avatar's senses and body to interact with the VE.

Matching "requires that there is match (sic) between the participant's proprioceptive feedback about body movements, and the information generated on the displays" (Slater & Wilbur, 1997, p. 605). For example, if the user turns their head, their avatar's head should turn in concert, and that the VE's display should match what the user would expect to see when turning their head at that speed and in that direction. This is further improved when the individual's actual movements match that which occurs in the VE. For example, Hendrick and Bartfield (1996a, as cited in Slater & Wilbur, 1997) found that when head tracking was used to control the user's avatar, presence was increased. Similarly, Slater, Usoh, and Steed (1995, as cited in Slater & Wilbur, 1997) found that walking in place to move the user's avatar resulted in more presence than using a button-based input system. Matching is also promoted when there is not a prolonged latency between the user's input and the avatar's reaction. In a study by Meehan et al. (2003, as cited in Slater, n.d.), a latency of 90ms decreased presence when compared to a group with a 50ms delay, as measured by increases in heart rate in a visual cliff task, where a higher heart rate was said to be indicative of more presence.

As noted above, the visual display should visualize that which is perceptually realistic for the user's actions (Slater & Wilbur, 1997). This can include several visual cues such as optic flow, motion parallax, and stereopsis (Bowman & McMahan, 2007; Ling, Brinkman, Nefs, Qu, & Heynderickx, 2012). Optic flow refers to the perception of movement relative to the self as objects move over the retina (Gibson, 1950). Motion parallax is a depth clue where closer objects will move across the retina faster than objects that are more distant (Gibson, 1950; Gibson, Gibson, Smith, & Flock, 1959). Stereopsis is also a depth cue where binocular images are blended together to produce a three-dimensional effect (Ling et al., 2012). While all increase perceptual realism, it may be that motion parallax is a stronger depth cue, and that this may be more influential in engendering presence than stereopsis (Ling et al., 2012).

The plot of the VE also likely plays a factor in that, “the more the ‘plot’ line potentially removes a person from everyday reality, and presents an alternate self-contained world, the greater the chance for presence” (Slater & Wilbur, 1997, p. 5). For example, a study by Welch et al. (1996, as cited in Slater & Wilbur, 1997) found that VEs which promoted the user to interact resulted in more presence experienced by users than those who were simply observing a simulation running within the VE. It seems that increasing interaction and autonomy as part of a cohesive plot in a self-contained world increases presence. Later work by Witmer and Singer (1998) will suggest that this notion is captured by the construct of involvement, which acts to reciprocally increase immersion, and thus improves the chances of experiencing presence.

Ultimately, all of these factors work together to influence the overall level of immersion, and that immersion is then predictive of presence. Slater and Wilbur (1997) stated that immersion is objective and quantifiable, while presence is a state of consciousness that can be assessed subjectively through ratings of feeling physically present in the VE, as well as objectively through behaviors exhibited by the participant. In line with Held and Durlach (1987), Slater and Wilbur (1997) believed that if a virtual object suddenly appeared and flew directly at the user in the VE, the user is likely to actually duck in real-life, even if their body movements were not controlling the movement of their avatar in the VE. Thus, the researchers believed that "the subjective [measure of presence] may be correlated with the higher levels of immersion", which would be the inclusive, extensive, surrounding, and vivid aspects, while "the objective [behaviors of presence] may be correlated with more fundamental aspects of immersion" (Slater & Wilbur, 1997, p. 606), which included self-representation, matching, and possibly plot. As a note, Slater and Wilbur (1997) did not explicitly classify the different factors as such, and the rest of the original article seems to suggest that self-representation, matching, and plot are still linked to immersion, but are not directly associated with presence.

Slater and Wilbur (1997) also proposed that these immersive factors are likely mediated by two additional aspects: the context and the individual's own sensory modality preferences. Regarding context, the goal of the VE will alter the relative importance of the different factors, providing the example that the auditory elements will be far more relevant and important than visual ones in an orchestra VE. Regarding individual differences, they also suggested that different senses are more important to some individuals (e.g., a blind person will consider auditory and haptic cues as more important to immersion in a VE than visual cues). The complete Slater and Wilbur (1997) model of immersion and presence is represented below in Figure 1, while potential experimental manipulations to the factors and aspects influencing immersion are found in Table 1.

Despite these clear, objective guidelines for quantifying immersion, which in turn is said to allow for presence, measuring the subjective experience of presence proved difficult. In a later review, Slater (n.d.) examined the many different ways presence could be measured. Citing Freeman et al. (1999), measuring presence with questionnaires is unreliable, and, citing himself, "the very asking of questions about 'presence' may bring into being, post-hoc, the phenomenon that the questionnaire is supposed to be measuring" (Slater, 2004, as cited in Slater, n.d., p. 3). Slater suggested that objective measures should be used, such as behavioral outcomes like the user reacting to stimuli in the VE as Held and Durlach (1987) had suggested. Slater also suggested that physiological measures, such as heart rate and eye tracking, could be utilized to objectively measure presence.

Table 1. Manipulations to increase immersion based on the Slater and Wibur (1997) model.

Factor / Aspect	**Manipulation to Increase Immersion**

Inclusive	• Reduce outside distractions
	  ○ e.g. loud external noise and voices vs. no external noise
	• Unobtrusive controls and peripherals
	  ○ e.g. a heavy HMD vs. light HMD
	• High resolution (example resolution not provided)
	  ○ e.g. high resolution vs. low resolution
	• High frame rate (more than 15 to 20 Hz)*
	  ○ e.g. frame rate at 10, 15, … 30 Hz, and 60 Hz

Extensive	• Increase number of sensory modalities ○ e.g. whether haptics, spatialized audio*, etc., are included
Surrounding	• Increase field of view thorough immersive technologies (50 – 90° is better than 10°)* ○ e.g. comparing CAVE to an HMD to a monitor

Vivid	• Increase scene detail
	  ○ e.g. detailed, cohesive graphics vs. incomplete wireframes
	• Increase realistic physics for the context
	  ○ e.g. turning on and off physics (friction**)

Self-Representation	• Have self-representation that allows for control and interaction ○ e.g. first-person perspective vs. third-person perspective

Plot	• Have an interactive, cohesive purpose to the VE
	  ○ e.g. interact with the VE or watch someone do it
	  ○ e.g. goal-driven activity vs. free-explore with no purpose

Matching	• Movement results in on-screen, realistic perceptual feedback
	  ○ e.g. realistic vs. unrealistic optic flow based on head movement
	• Include naturalistic inputs to move, view in the environment
	  ○ e.g. head tracking controls vs. traditional controls
	  ○ e.g. walking in place vs. traditional controls for movement***
	• Reduce latency between input and action (50ms or less)*
	  ○ e.g. alter latency time from 30 – 90ms

Notes: *Based on Barfield and Hendrix (1995, as cited in Slater & Wilbur, 1997); **Based on Uno & Slater (1997, as cited in Slater & Wilbur, 1997); ***Based on Slater, Usoh, & Steed (1995, as cited in Slater & Wilbur, 1997).

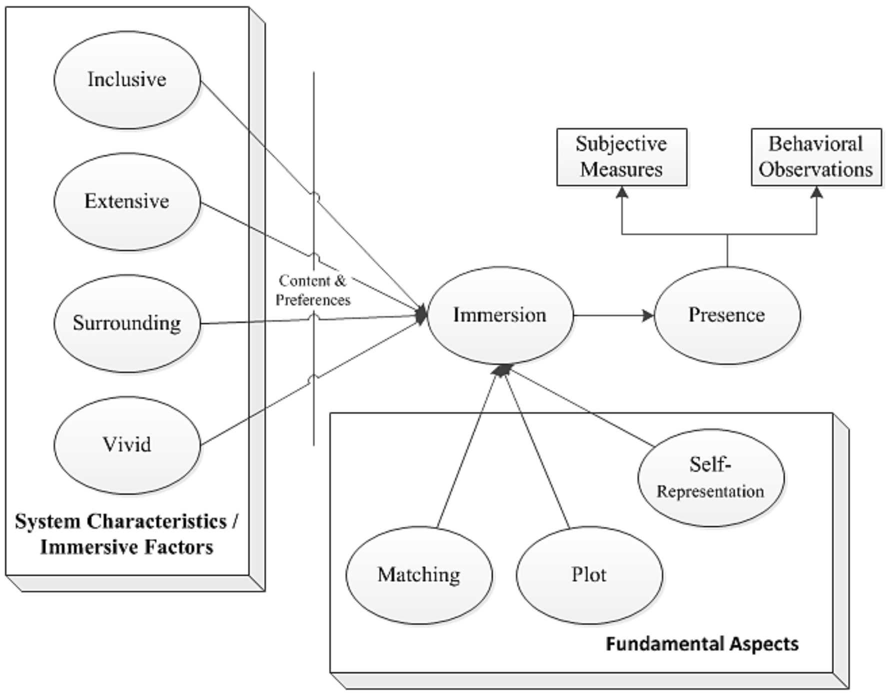


Figure 1. Slater and Wilbur's (1997) model of immersion and presence.

Technological approach: Witmer & Singer, 1998. One year after Slater and Wilbur (1997) published their definitions of immersion and presence, Bob Witmer and Michael J. Singer (1998) published an article in the journal *Presence* outlining their two new measures: the Immersive Tendencies Questionnaire and the Presence Questionnaire. Most importantly, while the Witmer and Singer (1998) definition of presence mirrored that of Slater and Wilbur (1997), their definition of immersion focused on the subjective experience, effectively adopting the psychological approach, rather than the technological approach taken by Slater and Wilbur (1997). Witmer and Singer (1998) defined immersion and presence as such:

- Immersion is “a psychological state characterized by perceiving oneself to be enveloped by, included in, and interacting with an environment that provides a continuous stream of stimuli and experiences” (Witmer & Singer, 1998, p. 227).

- Presence is “defined as the subjective experience of being in one place or environment, even when one is physically situated in another…As applied to a [VE], presence refers to experiencing the computer-generated environment rather than the actual physical locale” (Witmer & Singer, 1998, p. 225).

While the Witmer and Singer (1998) definition agreed with Slater and Wilbur (1997) that presence is a subjective sense of being in a location than the user’s actual physical environment, the authors defined immersion in a way that did not refer to characteristics of the technology, but as a subjective state in its own right. Essentially, immersion is the subjective experience of being enveloped in a stimulus flow provided by a VE in which interaction is possible, and presence is feeling as if the user has left reality behind and is actually physically present within the VE. See Figure 2 below for an illustration of their theory, followed by a description of the steps leading to the experience of presence according to Witmer and Singer (1998).

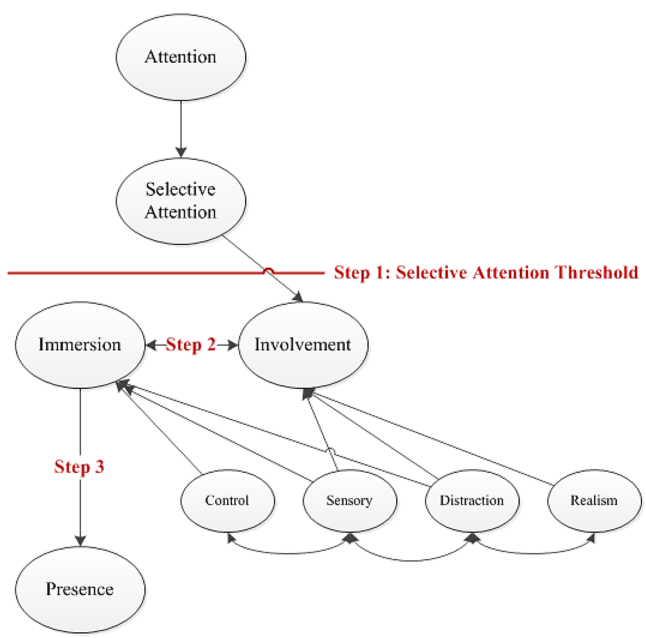


Figure 2. Witmer and Singer's (1998) model of immersion and presence.

*Step 1: Focus with selective attention***.** After allotting enough selective attention, which is focusing ones attention on meaningful stimuli (see Treisman, 1969), a threshold is reached (Witmer & Singer, 1998).

*Step 2: Experience immersion and involvement***.** An individual will begin to experience immersion and involvement once they have invested enough selective attention and breached the aforementioned attentional threshold (Witmer & Singer, 1998). As stated previously, Witmer and Singer's (1998) definition of immersion pertains to feeling as if an individual is enveloped by a VE's stimulus flow. Involvement is defined as a "psychological state experienced as a consequence of focusing one's energy and attention on a coherent set of stimuli or meaningfully related activities and events. Involvement depends on the degree of significance or meaning that the individual attaches to the stimuli, activities, or events" (Witmer & Singer, 1998, p. 227). While involvement stems directly from focus and selective attention, the researchers stated that both are needed to experience presence. Furthermore, immersion and involvement are inter-dependent (e.g., as one increases, the other increases as well), and are both influenced by a number of interacting factors. Witmer and Singer (1998) stated that these factors were based on the previous research of Sheridan (1992), Fontaine (1992), McGreevy (1992), and Held and Durlach (1992). They purported that the factors influenced either immersion or involvement, or sometimes both, and they also interacted with one-another. These factors included:

- *Control* (influences immersion) – The more control an individual has when interacting with the VE, the more immersion is experienced. Control is positively influenced by:
    - Degree: The user has increased control over the VE.

    - Immediacy: The user's inputs have an immediate effect without discernible latency.
    - Anticipation of events: The user can reliably predict the outcome of their inputs.
    - Mode: The method of control should be natural, in that it makes sense to the user and that it is intuitive.
    - Changeability: The user should be able to interact with the environment and make actual changes (e.g., open doors, interact with objects).
- *Sensory* (influences immersion and involvement) – The more senses modeled and the higher fidelity of those sensory inputs, the greater the sense of immersion and involvement will be. This is influenced by:
    - Modalities: The senses modeled will affect immersion and involvement. For example, Slater and Wilbur (1997) found that visual and kinesthetic aspects were more important than auditory stimuli; however, this hierarchy will depend on the VE's context.
    - Richness: By providing more sensory information, more immersion and involvement will be possible.
    - Multi-modal: Increasing the number of modalities will increase immersion and involvement.
    - Consistent: All of the senses should be in agreement (e.g., if it looks windy, it should also sound windy in the VE).

    - Self-movement perception realism: This is similar to the realistic perceptual requirement posed by Slater and Wilbur (1997), in that if an individual moves their head within the VE, the display should reflect the appropriate perceptual response.
    - Active search: Users should be able to actively use senses to locate objects and complete tasks in the environment.
- *Distraction* (influences immersion and involvement) – Minimizing distraction increases immersion and involvement. This will be aided by:
    - Isolation: Removing the user from the real-world, such as when using an HMD, will increase both immersion and involvement.
    - Sustained selective attention: Immersion and involvement will be enhanced if the user can focus on the VE, and is motivated to do so, in the absence of outside distractions. This will make it easier for the user to sustain attention.
    - Interface awareness: Poorly designed interfaces will detract from immersion and involvement as they are both distracting and require the user to apply effort to using the VE rather than experiencing it.
- *Realism* (influences involvement) – Increasing realism increases involvement. As an aside, it is interesting that that this is linked to Slater and Wilbur's (1997) definition of immersion, yet not Witmer and Singer's (1998) definition of immersion. Aspects of realism include:

- Scene realism: Including realistic lighting, content, and textures, as well as increasing the resolution, should all be positively associated with involvement.
- Consistency: Creating a VE that is consistent with the real-world and how the real-world operates will increase involvement.
- Meaningful: If the VE is meaningful to the user, be it because they are motivated to use it for their own enjoyment or to enhance their performance at a task, involvement will be increased.
- Separation anxiety/disorientation: When removing oneself from the VE, as involvement increases, they will be more likely to experience disorientation.

*Step 3: Increased immersion results in more presence experienced***,** in which the individual perceives that they are actually physically located in the VE, despite its inherent impossibility (Witmer & Singer, 1998).

Based on their theory, Witmer and Singer (1998) crafted the Presence Questionnaire (PQ) to quantify how present an individual felt in a VE, as well as the Immersive Tendencies Questionnaire (ITQ), which was concerned with an individual's general tendency to become involved and immersed. They validated both scales by using them in four experiments that utilized VEs by correlating single-item scores to the measures' totals. It is important to note that the purpose of the experiments was not for scale validation. While all four involved an HMD, two of the experimental VEs involved simple motor tasks, while the other two featured more complex tasks, had much higher resolutions, and were related to route learning. The participants included 152 undergraduate students from the Orlando, Florida area. As a part of the larger studies, they were given both the 32-item PQ and the 29-item ITQ. The items in the scales were measured on a 7-point semantic differential scale that also featured a mid-point anchor. The PQ generated a total score as well as scores for several proposed factors and sub-factors. Their validation study reduced the PQ to 19 items (Cronbach's $\alpha = 0.88$, $N = 152$, $M = 98.11$, $SD = 15.78$; Witmer & Singer, 1998, p. 236), and cluster analysis revealed the following three subscales:

- Involved / control (11 items; $M = 57.39$, $SD = 8.96$): This cluster of items represented the user's perception of "control of events in the VE, responsiveness of the VE to user-initiated actions, how involving were the visual aspects of the VE, and how involved in the experience the participant became" (Witmer & Singer, 1998, p. 236). Also, the authors noted that presence is not a direct result of becoming involved, however involvement should still be considered an "essential component" (Witmer & Singer, 1998, p. 239).
- Natural (3 items; $M = 12.36$, $SD = 3.44$): This cluster represented "the extent to which the interactions felt natural, the extent to which the VE was consistent with reality, and how natural was the control of locomotion through the VE" (Witmer & Singer, 1998, p. 236).
- Interface quality (3 items; $M = 14.65$, $SD = 3.4$): This cluster addressed "whether control devices or display devices interfere or distract from task performance, and the extent to which the participants felt able to concentrate on the tasks" (Witmer & Singer, 1998, p. 236).

These findings are interesting, mainly in that the many proposed factors did not emerge as significant in Witmer and Singer's (1998) analysis. For example, no subscale was exclusive to realism, the items that initially comprised the control subscale were divided between the three factors, and the natural subscale included an extra item that they did not originally hypothesize. This does alter the Witmer and Singer (1998) model, somewhat, as there is no longer a clear specification as to what affects involvement and what affects immersion, although they do make it clear that involvement is a prerequisite for immersion and that immersion is still linked directly to presence. Perhaps a simplified model is warranted, as proposed below in Figure 3.

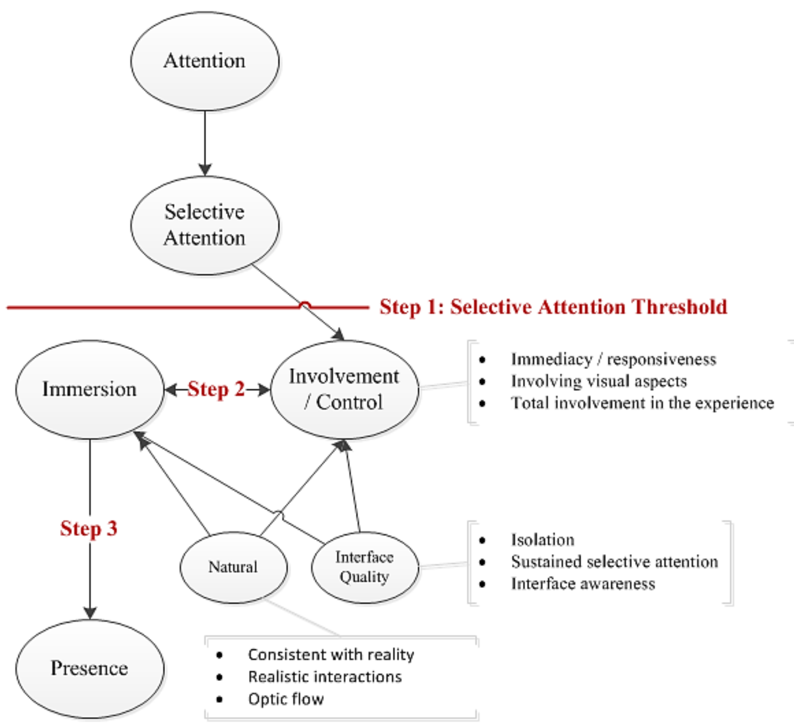


Figure 3. Witmer and Singer's (1998) model of immersion and presence after PQ validation.

Regarding the ITQ, the scale was reduced to 18 items (Cronbach's $\alpha = 0.81$, $N = 132$, $M = 76.66$, $SD = 13.61$). Cluster analysis revealed the following three subscales:

- Involvement (7 items, *M* = 26.51, *SD* = 7.24): This cluster pertained to the "subjects' propensity to get involved passively in some activity, such as reading books, watching television, or viewing movies" (Witmer & Singer, 1998, p. 236).
- Focus (7 items; *M* = 40.33, *SD* = 6.07): This cluster represented the user's "state of mental alertness, their ability to concentrate on enjoyable activities, and their ability to block out distractions" (Witmer & Singer, 1998, p. 236).
- Games (2 items; *M* = 6.21, *SD* = 3.16): This cluster focused on game playing habits, specifically "how frequently they play video games, and another asking whether they get involved to the extent that they feel like they are inside the games" (Witmer & Singer, 1998, p. 236).

Both the ITQ and the PQ were found to be reliable across the four experiments. Both were claimed by Witmer and Singer (1998) to have content validity since their items were derived directly from theory. They also examined the ways in which the PQ and ITQ scores were related to both similar and dissimilar constructs. They found that there was a significant negative correlation between PQ score and the Simulator Sickness Questionnaire (Kennedy, Lane, Berbaum, & Lilienthal, 1993) scores (r = - .426, $p < 0.001$), which is logical as becoming sick while engaging with a VE would be distracting and should lower presence (Witmer & Singer, 1998, p. 237). Witmer and Singer (1998) did not find a significant correlation between spatial ability and the PQ or ITQ scores, which was expected as they asserted that those constructs were unrelated. Witmer and Singer (1998) were able to test some elements of their model by comparing the different types of VEs from within their four studies. They were able to explore level of naturalistic inputs as they influenced presence, but did not find a difference in PQ score between head-tracking and using a joy stick for avatar locomotion. Interestingly, they found that PQ score was related to task performance in VEs. Finally, they found that ITQ and PQ scores were positively correlated across all four studies ($r = .24$, $p < .001$), however when this was examined for each study in isolation, this correlation was not found in two of the studies (Witmer & Singer, 1998, p. 238). Ultimately, the authors believed that the way they measured presence was valid, although they did warn that their measure should be subjected to advanced analytical techniques, such as factor analysis (which did eventually occur, and is discussed in a later section).

The immersion and presence debate. One year later, Slater (1999) published a response in the same journal as Witmer and Singer (1998). Slater (1999) found several flaws with the way that Witmer and Singer (1998) conceptualized immersion as well as their method of measure validation. While he agreed with their definition of presence, he disagreed with the way immersion was defined. He suggested renaming the immersion described in the original Slater and Wilbur (1997) work to "system immersion" (Slater, 1999, p. 560), defined as the objective characteristics of a VE that create a surrounding environment in which the distractions of the real-world are minimized. He also suggested that the Witmer and Singer (1998) definition of immersion should be renamed as the "immersive response" (Slater, 1999, p. 560), which represents the subjective response of the VE user regarding a feeling of being enveloped by the stimulus flow of an interactive VE.

Slater (1999) disagreed with how the PQ quantified presence, in that a total presence score is calculated as a sum of the items, none of which attempted to measure presence directly. This is deliberate on the part of Witmer and Singer (1998), who instead wanted to focus on factors leading to presence rather than including explicit questions as to whether the VE user felt present (Witmer, Jerome, & Singer, 2005). Slater's (1999) major complaint was that the PQ only measured subjective factors, in that they did not refer to quantifiable system characteristics, and that they had not been established as actual determinants of presence.

According to Slater (1999), without items that pertained specifically to presence, finding support in a correlational analysis of single item scores to a summation of all items is mathematically flawed. Essentially, Witmer and Singer (1998) defined presence in a specific way, yet did not include any items to measure that definition. Instead, they proposed items and measured those factors, and then found a significantly positive correlation between the factors and the total score. Arguably, this analysis instead revealed whether the items were all measuring the same construct, but whether that construct is presence cannot be determined in this way (Slater, 1999). It is more likely that these items were measuring immersion, instead. None of the PQ items actually pertained to presence but the experience leading to presence—which is immersion.

Slater (1999) also took issue with several other parts of Witmer and Singer's (1998) paper. He was suspicious of the subjective method of measurement and provided the following example: Imagine that there are identical systems and two identical tasks: One participant excels at the task, while the other is absolutely terrible. One participant then rates the control factors of the experience highly, and the other lowly, purely based on their performance. The one who gave the low rating would then be said that they had a lower overall level of presence, however without additional items measuring presence specifically, this cannot be confirmed. Perhaps they would have performed just as poorly in real-life, which would mean that the experience was more realistic, and actually increased their sense of presence. According to Slater (1999), this method of measuring presence is flawed in this particular circumstance, as the PQ score may have been lower despite experiencing more presence.

Slater (1999) highlighted the PQ's inability to separate out the objective system aspects and the individual subjective factors influencing presence. Slater (1999) also questioned the overall validity of the experiments, especially since they were not manipulated in ways meant to influence presence in testable ways. Slater concluded by boldly stating that he will never use the PQ to measure presence, although he does say that he will use the ITQ, despite his inherent dislike of subjective assessments, as tendency to experience immersion is valuable and worth measuring.

Singer and Witmer (1999) did respond to Slater's (1999) criticisms. They felt that the crux of Slater's argument was that immersion is only objective, and that subjective measurement is invalid. Singer and Witmer (1999) cite Sheridan (1992), who specifically stated that presence is a subjective state, which is to say that subjective measures of a subjective state are not unacceptable. Furthermore, to fully measure the construct likely requires a combination of subjective assessment *alongside* objective measures and quantifiable aspects of system characteristics. Having the system characteristics to support to experience of immersion and presence is absolutely necessary, but it is whether the individual user *perceives* that part of the system that is most important. They illustrate this concept with the following statement: "For example, frame rate could be varied between thirty and sixty frames per second and would not affect presence unless people perceived meaningful differences in the smoothness of the visual scene presentation" (Singer & Witmer, 1999, p. 569).

Singer and Witmer (1999) defended their total item score to represent presence, as the items each represent an aspect culled from the literature that should influence presence. This means, that even though they do not measure presence directly, the factors are a part of presence, and taken together represent the presence construct. This is one point where Slater's (1999) argument holds more weight: Without validating the total score against a secondary measure of presence, it is difficult to say this for sure. However, Singer and Witmer (1999) did suggest that this is a step that could be taken, and that they would expect the correlation to be positive and strongly significant. They also defended their method of validation, by re-calculating correlations between item averages and PQ scores with the item of interest removed, finding the same pattern of results.

Ultimately, the Witmer and Singer (1998) model is the one that was adopted out of the two competing operationalizations. According to Google Scholar, as of December 30, 2014, Slater and Wilbur (1997) have been cited 501 times while Witmer and Singer (1998) have been cited 2,117 times. However, having a higher citation count does not mean that Witmer and Singer (1998) were correct, as Slater (1999) had pointed out numerous valid flaws with their method and measure.

The factor structure of Witmer and Singer's (1998) PQ was examined in a follow-up validation study including 325 participants who completed the PQ after engaging with an immersive VE (Witmer et al., 2005). The researchers summarized a string of studies conducted by other researchers attempting to identify factors of presence. Many of the identified factors involved elements related to fidelity, spatial information, and involvement, but the overall number and types of factors varied widely. Witmer et al. (2005) believed that this was because the different measures were not tapping into the core construct of presence, and sought to address a need to refine the theory. Over the course of three studies, four factors were found, including "Involvement, Adaptation/Immersion, Sensory Fidelity, and Interface Quality" (Witmer et al., 2005, p. 298). This is interesting as the initial Witmer and Singer (1998) study of the PQ did not reveal a factor directly related to immersion.

The involvement factor was descriptive of focusing attention, and "is increased by performing tasks and participating in activities that stimulate, challenge, and engage the user either cognitively, physically, or emotionally" (Witmer et al., 2005, p. 299). It involves a naturalistic interface which allows the individual to immediately begin controlling the VE. This emerged as the most prevalent factor in the scale. The adaptation / immersion factor captured the "perceived proficiency of interacting with and operating in the VE and how quickly the user adjusted to the VE experience" (Witmer et al., 2005, p. 303). Labelling this factor as 'adaptation / immersion' seems to be a stretch, as it is more related to usability than immersion. The interface quality factor pertained to other aspects of interface usability, in that it was not distracting from the experience. The visual fidelity factor described the extent and fidelity of the sensory experience provided by the VE.

Witmer et al. (2005) found that involvement was enhanced by sensory fidelity, and proposed that this occurred because greater sensory fidelity captured attention more easily. They also found that low sensory fidelity had more of an impact on involvement than immersion, where low-fidelity was distracting and disruptive to the experience.

The titles of the factors presented by Witmer et al. (2005) are somewhat confusing and do not align well with the literature. For example, the PQ's involvement subscale is meant to measure aspects of a VE that involve focused attention and cognitive engagement with the task. While one item does pertain to involvement ("How involved were you in the virtual environment experience?"), the majority of PQ items align with aspects that are related to immersion, such as the usage of an avatar's senses within the VE ("How completely were you able to actively survey or search the environment using vision?") and naturalistic input mechanisms ("How natural was the mechanism which controlled movement through the environment?"; Witmer et al., 2005, p. 302). Still, the content of the PQ's items is telling. Ignoring the names of the factors, the following is apparent: Attention, sensory fidelity, and usability to minimize distractions and promote a natural experience are all determinants of presence, which is the result of a highly-immersed state. Motivation is also relevant to the experience and is captured by the involvement construct.

Consolidating Slater & Wilbur (1997) and Witmer & Singer (1998). Based on the Slater (1999) evaluation and the Singer and Witmer (1999) response, it is reasonable to suggest that both system immersion and the immersive response comprise the larger concept of 'immersion', and both have a role to play in influencing whether an individual experiences presence. This is exactly what Draper et al. (1998) suggested.

The two models of immersion and presence may be able to be combined as such: The attentional requirements aspect likely continues to be important and is well-justified by Witmer and Singer (1998). The importance of attention, while it did receive some early discussion (e.g., Draper et al., 1998), later became an important driving mechanism in contemporary presence theory (e.g., Wirth et al., 2007). The notion of system immersion is important, but it is not the only determinant to be considered. Witmer and Singer (1998) suggested that there is some sort of subjective state in-between the technology and the presence experience that they termed immersion, which Slater suggested changing to the immersive response. As Singer and Witmer (1999) suggested, it is the system immersion quality that provides the opportunity to become immersed, while the subjective experience is interpreting those system characteristics in such a way that promotes feeling enveloped in the VE's stimulus flow. In theory, as the immersive response grows stronger, it becomes more likely that presence will be achieved. Additionally, Witmer and Singer (1998) stated that the subjective experience of involvement is positively associated with immersion. Recall that involvement pertains to motivation. This notion is similar to the concept of plot that Slater and Wilbur (1997) introduced as being related to immersion, but not actually as a part of immersion itself.

There are many similarities between Slater and Wilbur's (1997) objective characteristics of immersive systems and the factors influencing the subjective experience of immersion that were listed by Witmer and Singer (1998). Witmer and Singer's (1998) distraction factor includes isolation, such as using an HMD to block out the real-world, which is identical to Slater and Wilbur's (1997) surrounding factor. Slater and Wilbur's (1997) inclusive factor, in which sustained selective attention is focused by limiting outside distractions, is also the same as the sustained selective attention aspect of Witmer and Singer's (1998) distraction factor. Also, the Witmer and Singer (1998) notion of interface awareness is easily grouped with Slater and Wilbur's (1997) inclusive factor as it suggests that there should be unobtrusive controls and peripherals. All of these elements pertain to minimizing distractions to allow for focused, sustained attention, which is also supported by the Witmer et al. (2005) factor analysis, with respect to the interface quality and immersion / adaptation factors in particular.

Witmer and Singer's (1998) control factor is similar to Slater and Wilbur's (1997) fundamental aspect of matching, in which the immediacy requirement of control aligns very well with the latency requirement of matching, and both highlighted the importance of naturalistic, easy to learn, and intuitive controls. Witmer and Singer's (1998) realism factor and Witmer et al.'s (2005) sensory fidelity factor, which includes scene realism and consistency, almost perfectly mirrors Slater and Wilbur's (1997) vivid factor, although Slater and Wilbur (1997) focused less on real-world realism as this factor instead pertained to realism within the context. It is in this sense that a cohesive, high-resolution, yet cartoonish VE could be considered vivid by Slater and Wilbur (1997), but not realistic according to Witmer and Singer (1998). Overall, this suggests that the VEs that are naturally and easily interacted with and are also realistic within context promote the experience of immersion and, eventually, presence.

Finally, Witmer and Singer's (1998) sensory factor has an enormous amount of overlap with several of Slater and Wilbur's (1997) factors, including the multiple sensory modality element of extensive, the richness aspect reflected in the vivid factor, as well as the high-resolution aspect of inclusive. Also, Witmer and Singer's (1998) self-movement perception realism overlaps with Slater and Wilbur's (1997) self-representation and realistic motion-perception requirements. Therefore, in addition to realism within context, high sensory fidelity—but not in the photorealistic sense—is a must for presence to occur.

The remaining aspects of system immersion (Slater & Wilbur, 1997), which included self-representation, plot, and the latency requirement of matching, could arguably be related to Witmer and Singer's (1998) involvement / control factor, which was stated to be separate from the immersive response, but remains an important contributor. Witmer and Singer's (1998) control factor aligns well with the aspect of plot, in that the experience in the VE is goal-driven and highly interactive. Goals are an interesting addition to the model. Goals are a strong driver of motivation, and coupled with self-evaluation and the ability to meet sub-goals, they encourage an individual to continue to strive and engage with a task (Bandura, 1982). Challenging, focused goals in particular result in higher performance than others, likely through their ability to direct attention and encourage motivation (Locke, Shaw, Saari, & Latham, 1981). Indeed, motivation—which was perhaps goal-driven—would later become an important aspect of presence (Wirth et al., 2007) and plays a role in other relevant constructs, such as flow (Csikszentmihalyi, 1990).

Ultimately, it appears that system immersion influences both the immersive response and involvement, although both have their own unique contributions. Additionally, Witmer and Singer (1998) are very clear that involvement is a separate, yet fundamental element that influences immersion (and immersion in turn influences involvement). Also relevant seems to be additional constructs, such as goals and motivation. By understanding the differences and overlap between system immersion and the immersive response, as well as considering the above additions, a combination of the two models can be proposed (see Figure 4).

Slater and Wilbur's (1997) and Witmer and Singer's (1998) model of immersion and presence can be simplified, and is supported by the findings of Witmer and Singer's (2005) factor analysis of the PQ: Presence occurs when there are high levels of immersion, and devoting one's attention is necessary for achieving immersion. Characteristics of the VE technology promote immersion by both minimizing distractions (e.g., VEs that occlude the real-world, are usable and feel natural to interact with, and are realistic within context) to allow for focused, sustained attention, as well as by providing immersive cues to the senses (high sensory resolution, multiple senses controlled by the VE). The user must then subjectively appraise these characteristics as being immersive. There is also an element of involvement, which deals with goal-driven motivation and has a reciprocal, positive relationship with immersion (Witmer & Singer, 1998).

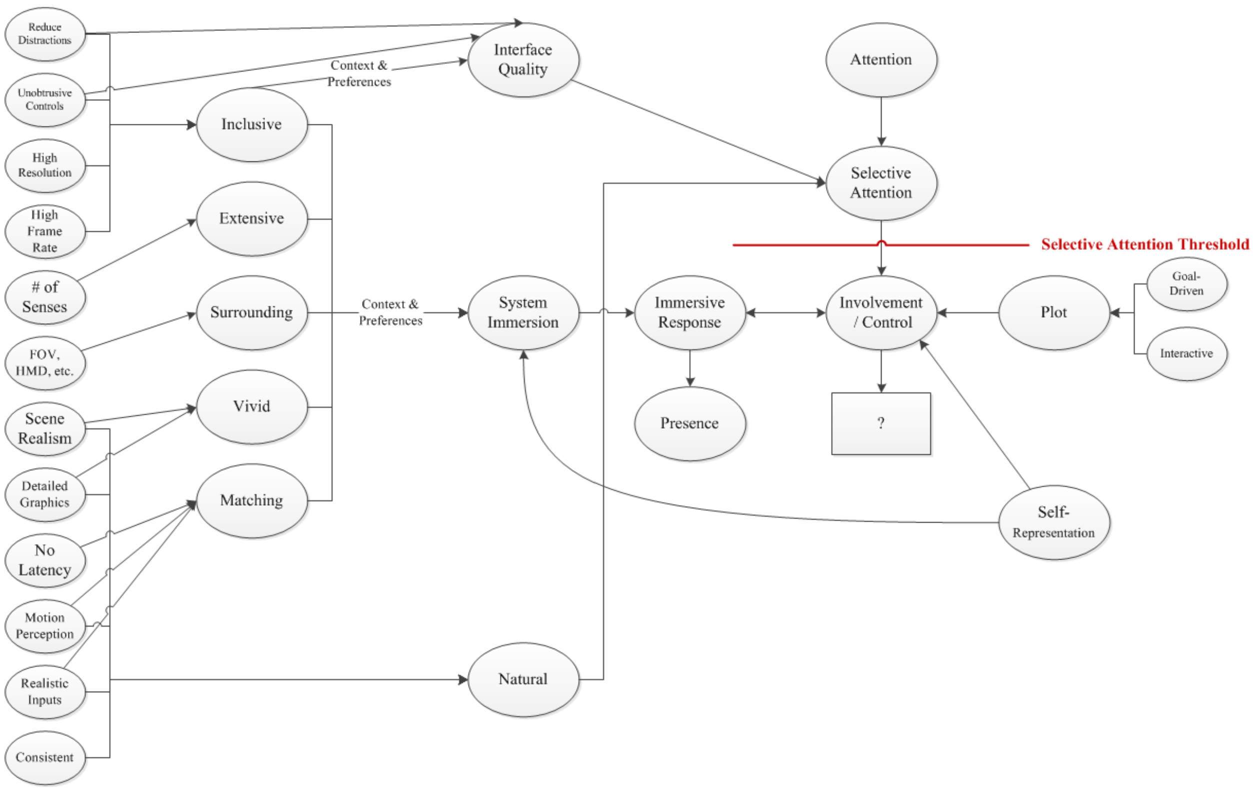


Figure 4. Proposed combination of Slater and Wilbur's (1997) and Witmer and Singer's (1998) models of immersion and presence based on theoretical overlap.

As the 1990s came to a close, an understanding of presence and immersion was just developing. IJsselsteijn et al. (2000) provided an extensive summary of the work performed in the 1990s and highlighted the potential ways to move forward. At this point, it was generally agreed upon that presence was a sense of 'being there,' while immersion was described as both an objective quality of VEs as well as the subjective experience of being enveloped in a VE's stimulus flow. Additional work during the following decade would continue to refine this definition. Furthermore, IJsselsteijn et al. (2000) stated that there are two types of presence that are of empirical interest: physical presence, in that an individual feels as if they are somewhere where they are not, and social presence, in that an individual feels as if they are communicating with another person in the same setting despite not being co-located. IJsselsteijn et al. (2000) maintained that these two types of presence are not unrelated, in that they likely have the same overlapping determinants, yet they are unique. This work focuses solely on physical presence, which will be referred to simply as presence from this point forward.

The work conducted in the 1990s revealed several different drivers of presence, all of which were described in a number of different ways. Still, they encapsulated the following core aspects, as outlined by IJsselsteijn et al. (2000):

- "Extent and fidelity of sensory information", which includes high resolution and an expansive field of view (IJsselsteijn et al., 2000, p. 521).
- "Match between sensors and the display", which is essentially mapping the user's movements in the VE to the avatar's movement, and then having the visual display respond appropriately (IJsselsteijn et al., 2000, p. 521).
- "Content factors", in that the VE has a first-person representation of the user's body, which can interact with the environment and other virtual agents, who in turn react to the user's existence and inputs (IJsselsteijn et al., 2000, p. 521).
- "User characteristics", which was an area that had been under-researched until this point, but pertains to how each individuals user's perceptual abilities, cognitive abilities, motor abilities, prior expectations, willingness to suspend disbelief, age, sex, and mental health all affect presence (IJsselsteijn et al., 2000, pp. 521-522).

Additionally, IJsselsteijn et al. (2000) provided a list of elements that disrupt presence, including poorly designed interfaces, misaligned stereoscopic equipment that results in eye strain, low-resolution textures and noise, the weight of the control and display devices, software errors, latency, low frame rates, sensory conflicts, and external distractions.

The work of IJsselsteijn et al. (2000) summarized and consolidated the theories of immersion and presence. This paved the way for further theoretical expansion and the development of complex models that applied immersion and presence directly to games.

The Early 2000s: Expanding presence and immersion to multiple domains

While much of the research in the 1990s focused on immersion, the definition of presence received attention in the 2000s and beyond. Despite explicitly examining presence, much of the previous work seemed to inadvertently focus on immersion (e.g., Witmer & Singer, 1998). In the 2000s, more of a focus was spent on simultaneously refining and expanding the presence definition by applying it more broadly to incorporate additional, relevant contexts as well as differentiating between the types of presence. Definitional unity across all of the various domains became a necessity (Wirth et al., 2007).

Lee (2004) penned a thorough and thoughtful theoretical piece examining presence through multiple contexts and through history. He stated that there were several problems with the then-current state of presence research in that many different terms were used to describe the same concept, that there was a weak theoretical foundation for the definition, and that the most popular conceptualizations ignored the three critical components of the human experience: the physical, the social, and the self.

Lee (2004) made the following key points regarding how presence should be defined: it is a psychological state; it is likely a desirable state so negative words should be omitted from the definition; and, it should encompass all types of possible presence, not just those mediated by a particular technology. Lee (2004) also raised an important point: There is a logical problem inherent to previous definitions of presence in that feeling as if you physically exist in a VE is not technologically possible (at the moment, anyway, as future technologies may be able to fool all of our senses). Instead, Lee (2004) suggested conceptualizing presence as experiencing objects that can be "perceive[d], manipulate[d], or interact[ed] with" (p. 38), which can be both sensory (e.g., on-screen objects in a VE) or non-sensory (e.g., when an individual is reading and using their imaginations). This suggested that the concept of presence applies to playing a game or watching a television show equally, despite the differing levels of interactivity between the two. Lee (2004) argued that presence is possible in these low-fidelity situations as, if perfect fidelity was a requirement for presence, no one could be considered as ever having experienced it as it is not yet possible to simulate all of the senses to the required extent. Instead, presence is subjective and our imaginations can often fill in the blanks, so to speak. Wirth et al. (2007) would later make this same argument. Finally, Lee (2004) further extended the definition of presence to refer to objects as either those that are para-authentic, in that they have real-world physical correlates, or artificial, in that they are the product of pure imagination.

Lee (2004) defined the overall construct of presence as "a psychological state in which virtual (para-authentic or artificial) objects are experienced as actual objects in either sensory or nonsensory ways" (p. 37). This allowed presence to be applied to all possible applications, from high-fidelity simulators to reading a book. He then further related this definition to the three core elements of the human experience.

Physical presence is “a psychological state in which virtual (para-authentic or artificial) physical objects are experienced as actual physical objects in either sensory or nonsensory ways” (Lee, 2004, p. 44). This is essentially the point at which an individual no longer notices the mediating technology, which could be the book or the video game controller in his or her hands. This wording of the definition removes the requirement that the user must feel like they physically exist in the mediated environment as Lee (2004) did not believe this to be possible. He also did not believe presence to be a binary state, whereas previous research purported that a VE user either was or was not present. Instead, Lee (2004) suggested that physical presence should be measured in “degrees of psychological similarities between virtual and actual objects in terms of (a) sensory perception, (b) physical manipulability, and (c) interaction quality” (p. 38).

Social presence is “a psychological state in which virtual (para-authentic or artificial) social actors are experienced as actual social actors in either sensory or nonsensory ways” (Lee, 2004, p. 45). This is a state in which the individual feels that the other agents within the environment are not artificial and that they are interacting with real people. This is different from co-presence, however, which is feeling as if one exists in a space with other people that are actually real and are concurrently accessing the environment remotely (Lee, 2004).

Finally, self-presence is “a psychological state in which virtual (para-authentic or artificial) self/selves are experienced as the actual self in either sensory or nonsensory ways” (Lee, 2004, p. 46). This is the ultimate goal of presence: For the user to feel as if they have become the avatar. They do not feel physically transported, however they no longer notice the technology and feel themselves co-located within the mediated environment.

Lee's (2004) analysis provided value in that it was among the first that attempted to unify all fields examining presence. Wirth et al. (2007) approached this problem with a similar goal, additionally stating that earlier work failed to differentiate presence from other similar concepts and that it overly focused on the technological aspects as opposed to the roles of attention and involvement. Following Lee's (2004) lead, Wirth et al. (2007) sought to consolidate and specify the definition of presence so that it was not technology-dependent and could address this multi-disciplinary research problem, was specific enough to be thoroughly tested and made distinct from other similar constructs, and so that it connected with older, stronger theoretical models from the fields of communications and psychology.

Wirth et al. (2007) defined "spatial presence", which is the same as IJsselsteijn et al.'s (2000) physical presence, as "a binary experience, during which perceived self-location and, in most cases, perceived action possibilities are connected to a mediated spatial environment, and mental capacities are bound by the mediated environment instead of reality" (p. 497). Similar to Lee (2004), this definition applies to all mediated experiences, ranging from technologically-sophisticated VEs to books. Here, it is important to note that they disagreed with Lee's (2004) notion that presence did not involve feeling as if the individual was physically present in the VE, as they maintained that "the main characteristic of Spatial Presence is the conviction of being located in a mediated environment" (Wirth et al., 2007, p. 495). Furthermore, it is this feeling that drives any sort of "media effects" stemming from the experience, such as the possibility of a surgeon's enhanced performance in telesurgery based on the assumption that if the surgeon felt co-located with the patient, their surgical performance would be increased (Wirth et al., 2007, p. 495).

The authors also stated that “presence is conceptualized as the experiential counterpart of immersion” (Wirth et al., 2007, p. 496). For example, increasing the number of senses which are controlled by the mediated environment will support the experience of presence. Similar to Lee (2004), Wirth et al. (2007) stated that lacking immersive qualities is not damaging to presence as the imagination can “compensate for that deficit in external stimulation – at least to a certain degree” (Wirth et al., 2007, p. 496).

Wirth et al.’s (2007) theory of spatial presence focused on attention and the formation of mental models. This process was hypothesized by the authors to occur in two steps: focusing attention to create a mental model of the mediated environment, and then accepting that mediated environment as where the individual is physically located.

Attention is a necessary preliminary component for this process to occur (Wirth et al., 2007). Draper et al. (1998) and Witmer and Singer (1998) acknowledged the role of attention in presence, however Wirth et al.’s (2007) theory featured a much stronger emphasis. Briefly returning to Draper et al.’s (1998) discussion of attention in presence, they cited the extensive work of Christopher D. Wickens in that humans have limited pools of attentional resources. Draper et al. (1998) proposed the following: An operator working in a remote environment must contend with two types of stimuli in two different environments: task-related information and distractions in both the remote operating environment and in the current physical environment. An operator can choose to perceive the task-specific stimuli, however individuals have sensory systems that will automatically direct attention to certain stimuli, such as distractions. According to Draper et al.’s (1998) model, immersion is “the degree to which perceptual system inputs are limited to displays of the computer-mediated environment” (Draper et al., 1998, p. 368), while telepresence occurs when an operator chooses to devote all of their attentional resources to what is occurring in the remote environment. The more attention paid has a positive correlation with the sense of telepresence achieved, and real-world distraction will break the operator out of telepresence.

Draper et al. (1998) believed that increasing immersion by manipulating characteristics of the system focuses attention by blocking out the external world. They provide the example of an HMD, which will completely focus the operator's visual senses, however they may still be distracted by auditory cues from the physical environment. Another interesting element that can focus attention is task difficulty. If a task is more difficult, and the individual is motivated to complete the task, the operator will expend more attention because the task demands it. The authors hypothesized that this should result in increased telepresence, even if it is detrimental to task performance due to the difficulty. Similarly, if a task is too easy, task performance may be high, however telepresence might be low because the user was bored and not paying attention to the remote environment.

Thus, attention is a fundamental driver of presence (and, therefore, immersion if it is indeed a downstream determinant of presence). Importantly, the development of the "spatial situation model" (Wirth et al., 2007, p. 497), or SSM, of the mediated environment is only possible if the individual is paying attention. The SSM is "a mental model (Johnson-Laird, 1983; Sanford & Garrod, 1981) of the spatial environment that the individual constructs based on (1) spatial cues she/he processes and (2) relevant personal spatial memories and cognitions (McNamara 1986)" (Wirth et al., 2007, p. 501). Wirth et al. (2007) asserted that creating an SSM is an absolutely necessary requirement for spatial presence. Therefore, attention is extremely important as a preliminary requirement for the process of experiencing presence.

Attention may be directed involuntarily (e.g., Posner, 1980, as cited in Wirth et al., 2007), or an individual can direct their attention purposefully, which is controlled attention allocation. Wirth et al. (2007) asserted that attention can also be controlled by the mediated environment, where certain characteristics can promote attention such as HMDs, which essentially follows Slater and Wilbur's (1997) conceptualization of immersion as well as the model leveraged by Draper et al. (1998). Once the individual begins to direct their attention to the mediated environment, they can begin to build the SSM. Sustaining attention will require effort, however the level of effort necessary will depend on the immersive qualities of the mediated environment (Wirth et al., 2007). For example, it will require less attention to create an SSM of a level in a game than of a location in a book because the game will provide direct sensory cues about the spatial qualities of the environment. Thus, more mental effort will need to be expended to create an SSM with equal detail from a book than from a game. Also, it is easier to pay attention to activities that an individual is interested in (Hidi, 1995, as cited in Wirth et al., 2007). Therefore, it may be that domain-specific interest also positively influences controlled attention allocation.

According to Wirth et al. (2007), the creation of the SSM depends on bottom-up feature processing as well as top-down assumptions based on the individual's own schemata. The SSM's overall correctness will vary depending on the level of information provided by the mediated environment, and whether that information is consistent with itself (e.g., all modalities should align to create a cogent experience), which will in turn allow for the formation of a more accurate SSM. SSM accuracy relies on stimuli richness and consistency, however individual differences also have an influential role. For example, those with better spatial-visual imagination are likely better at filling in the details not provided by the mediated environment and will have a more complete SSM than others in the absence of stimuli (Wirth et al., 2007).

In summary, the SSM is a mental representation of the mediated environment, while spatial presence is feeling as if one is actually located in the mediated environment. Possessing an SSM answers the question of, “Is this a space?” with a “yes”. After the SSM has been formed, the next step involves the individual testing a specific perceptual hypothesis that will answer the question of, “Am I *in* this space?”, to which an affirmative response is indicative of spatial presence.

Wirth et al. (2007) defined an individual’s egocentric reference frame (ERF) as an individual’s mental model of an environment. For example, when an individual is playing a video game, they possess an ERF for their physical environment (e.g., their living room) as well as an ERF for the game environment (e.g., the hotel lobby they are exploring in the video game). The ERF that the individual accepts as the one they are physically located within is known as the primary egocentric reference frame (PERF). Spatial presence occurs when an individual believes that the ERF provided by the mediated environment—which is the video game’s hotel lobby in this example—is the PERF. This decision is made by testing the “theory of perceptual hypotheses”, which is based on the work of Lilli and Frey (1993, as cited in Wirth et al., 2007, p. 506).

According to Wirth et al. (2007), there are three general stages of perceptual hypothesis testing, which included: *expectations*, where previous experiences guide an individual’s initial expectations; *input*, where the individual takes in information; and, *confirming*, where the individual determines if the information they have received aligns with their expectations, which in turn determines whether support for their hypothesis, which is whether a given mediated environment is the PERF, was found. A hypothesis is terminated when support is found, or else a new hypothesis is created. Also, it is possible to have multiple, simultaneous hypotheses, each of which will have different strengths relative to one-another. According to Wirth et al. (2007), the strength of the hypothesis is tied to the individual's expectations as well as the characteristics of the environment, and will determine how information is acquired in the input stage. A strong hypothesis exists if the mediated environment provides multiple cues across multiple modalities about the space, while a weak hypothesis exists if the opposite is true (Wirth et al., 2007). If an individual has a strong hypothesis, they consider information in a top-down manner to determine what matches the current hypothesis. If an individual has a weak hypothesis, without expectations to guide information search, an individual will utilize bottom-up processing to gather that information. Lilli and Frey (1993, as cited in Wirth et al., 2007) stated that strong hypotheses require less information to confirm and a greater amount of contradictory information to disprove, when compared to weak hypotheses.

Returning to the video game example, recall that a player has two ERFs: the real-world living room and the in-game hotel lobby. As a result of gameplay, the individual has formed an excellent SSM of the hotel lobby. The individual will unconsciously begin testing the hypothesis that the mediated environment—the hotel lobby—is actually the PERF. As an individual receives more information and cues that support this hypothesis, the more likely it becomes that this individual will accept the lobby as the PERF, thus experiencing presence in the video game. Every time an outside distraction occurs, the hypothesis will be questioned. If the game provides many immersive cues, such as through the use of an HMD and spatialized audio provided via headphones, the hypothesis that the hotel lobby is the PERF will be very strong due to the overall sensory consistency, richness, and persistence of stimuli stemming from the mediated environment. This will make it more difficult for outside distractions to contradict the mediated-environment-as-PERF hypothesis.

Even in the case of a weaker hypothesis where the individual does not have the luxury of playing with an HMD or with a nice pair of headphones, it is possible to mitigate the effect of contradictory evidence through motivation. Wirth et al. (2007) suggested that motivation assumes two forms: involvement and willingness to suspend disbelief. Involvement, similar to the Witmer and Singer (1998) definition, is described as “a motivation-related meta-concept that includes various forms of intense interactions with a mediated stimulus” (Wirth et al., 2007, p. 512). This involves both cognitive aspects, such as attention, as well as affective aspects, such as personal relevance. In short, when an involved individual makes connections between themselves and the mediated environment, they will be more willing to act in that mediated environment. Interestingly, Wirth et al. (2007) believed that involvement is not a determinant of presence, which is similar to what Witmer and Singer (1998) proposed. Instead, Wirth et al. (2007) stated that while likelihood to experience presence and involvement increases concurrently, they are not directly linked (e.g., an individual does not have to be involved to become present).

Suspension of disbelief occurs when an individual willingly does not pay “attention to external stimuli and internal cognitions that (might) distract from the enjoyment of the mediated story and environment” (Wirth et al., 2007, pp. 513-514). Despite feeling the weight of the HMD, an individual may choose to ignore it because they are willing to experience presence.

Even if there is an outside distraction, an individual may unconsciously disregard it due to motivation alone. What would normally be detrimental to a weak hypothesis will have less of a contradictory effect if the individual is willing to suspend disbelief (Wirth et al., 2007).

Wirth et al. (2007) purported that providing a highly-immersive mediated environment, increasing involvement, or having an ability or desire to suspend disbelief ameliorates the deficits of a weak hypothesis and allows spatial presence to occur. For example, even if an individual is not willing to suspend disbelief, the immersive qualities of the mediated environment may be influential enough to result in presence. The opposite is also true; lacking immersive qualities, pure motivation can allow an individual to experience presence.

Finally, Wirth et al. (2007) discussed how an individual's propensity toward absorption is also important. Similar to the measurement of immersive tendencies (Witmer & Singer, 1998), "trait absorption refers to an individual's motivation and skill in dealing with an object in an elaborate manner (Wild et al., 1995)" (Wirth et al., 2007, p. 515). In short, some are simply more likely than others to become engrossed without expending a large amount of attentional resources. Wirth et al. (2007) believed that trait absorption would be positively associated with attentional resource usage, involvement, and willingness to suspend disbelief.

Incorporating additional theoretical mechanisms

As the definitions of presence and immersion were refined, they also were simultaneously generalized and specified. The previous section summarized the work of Lee (2004) and Wirth et al. (2007), who strove to produce a general definition of presence as it could be applied universally. Others, which will be summarized here, also wanted to apply presence and immersion to specific contexts, such as video games. They did this by connecting immersion and presence to other theories relevant to the gaming experience. While many exist, two major models are discussed: the model of game immersion (Brown & Cairns, 2004; Jennett et al., 2008) and the model of game engagement (Brockmyer et al., 2009). Common to both are the theoretical constructs of flow and absorption, both of which are reviewed here before describing the game immersion and game engagement models in-depth.

Flow. Flow theory was first formally introduced by Csikszentmihalyi's (1975a) book *Beyond Boredom and Anxiety*. That same year, he described flow as such:

> There is a common experiential state which is present in various forms of play, and also under certain conditions in other activities which are not normally thought of as play. For lack of a better term, I will refer to this experience as "flow." Flow denotes the holistic sensation present when we act with total involvement. It is the kind of feeling after which one nostalgically says: "that was fun," or "that was enjoyable." It is the state in which action follows upon action according to an internal logic which seems to need no conscious intervention on our part. We experience it as a unified flowing from one moment to the next, in which we feel in control of our actions, and in which there is little distinction between self and environment; between stimulus and response; or between past, present, and future. (Csikszentmihalyi, 1975b, p. 43)

Flow occurs when individuals can afford attention to a stimulus and do so without any external threats or distractions. It is the result of being able to engage with an activity and achieve goals (Csikszentmihalyi, 1990). The flow state stems from experiences that mimic, or actually are, play activities that result in enjoyable, intrinsically-motivating experiences (Csikszentmihalyi, 1975b). It has been described as the "optimal experience" (Csikszentmihalyi, 1990, p. 39).

Csikszentmihalyi (1975b) surveyed individuals who willingly engaged in play-like activities, to include swimmers, chess masters, dancers, and composers, to better understand the flow experience and its root cause. The concept earned its namesake based on these interviews: "We have called this state the *flow experience*, because this is the term many of the people we have interviewed had used in their descriptions of how it felt to be in top form: 'It was like floating', 'I was carried on by the flow'" (Csikszentmihalyi, 1990, p. 40). It was also from these interviews that a fuller understanding of flow was garnered.

According to Csikszentmihalyi (1975b), the hallmark of the flow experience is the merging of action and awareness, in which an individual is wholly focused on the activity. The goals within the activity itself must be attainable, but not so easily that the experience is boring. It most often occurs in a rule-driven context, from religious rituals to games, that encourages the individual to "[center] attention on a limited stimulus field" (Csikszentmihalyi, 1975b, p. 47), which can be facilitated by reducing outside distractors. The similarities to immersion are immediately apparent; focused attention, limiting the stimulus field, and reducing outside distractions are elements that promote immersion (e.g., Slater & Wilbur, 1997; Witmer & Singer, 1998; Draper et al., 1998; Brown & Cairns, 2004).

Flow results in a loss of self-consciousness. This, however, "does not mean…that in flow a person loses touch with his or her own physical reality", which is to say that flow is not necessarily presence in the sense that the individual is mentally transported to a new environment, but instead "the [loss of] self-construct, the intermediary which one learns to interpose between stimulus and response" (Csikszentmihalyi, 1975b, p. 49). Therefore, the experience of flow is not quite presence; it is more like a highly-involved state.

Achieving a sense of control is important to the flow experience. This is also an element of immersion lauded by many researchers (e.g., Sheridan, 1992; Slater & Wilbur, 1997; Witmer & Singer, 1998). In presence and immersion, control usually pertains to the ability to control the senses in a VE (e.g., Sheridan, 1992). In flow, achieving a sense of control goes beyond the individual's ability to affect the outcome of their activity, but that they are confident in their abilities and are not worried about the possibility of failure (Csikszentmihalyi, 1975b). The activity must feature a clear progression and have goals rising from "ordered rules" (Csikszentmihalyi, 1975b, p. 53). Immediate feedback must be provided as one moves closer to achieving those goals, although this feedback does not necessarily mean the individual has a conscious evaluation of that information, but it is something that guides them automatically in the moment (Csikszentmihalyi, 1975b). Lastly, the flow experience is autotelic, in which the act itself "appears to need no goals or rewards external to itself" (Csikszentmihalyi, 1975b, p. 53); it is play purely for the sake of play. Flow is inherently intrinsic, even if the activity itself is not. If extrinsic motivators, such as competition in games or monetary gain in gambling, are present, the flow experience is still possible as the autotelic aspect of flow is a product of the experience itself, not the outcome of the activity, although these extrinsic motivators do make it more difficult to experience flow as they may become distracting (Csikszentmihalyi, 1975b).

Flow occurs when a careful balance is achieved between the difficulty of the tasks and the skill of the individual (Csikszentmihalyi, 1975b): too difficult of a challenge and too low of a skill-level results in worry, and eventually anxiety, while the opposite promotes boredom (and eventually anxiety, as well). What lies at this balance is the flow experience. It is important to note that it is not objective difficulty and skill, but the individual's *subjective* perception of the challenge and their own skill level (Csikszentmihalyi, 1975b). Interestingly, this balance between the challenge of the activity and the skill of the individual is the same mechanism proposed by Draper et al. (1998) that maintains an individual's focus in a VE.

Flow itself is not an easily-sustained experience. Momentary interruptions to flow are common, and many can only maintain flow for short bursts—although it is not impossible to sustain the flow state for a long period of time (Csikszentmihalyi, 1975b). Thus, flow is not a purely positive experience as an individual may not be in flow the entire time, however the optimal experience that is flow will occur is they individual averages around this point of challenge/skill balance.

Csikszentmihalyi continued to study flow throughout the decades. He surveyed hundreds of individuals around the world utilizing what he coined the "Experience Sampling Method" to gain a greater understanding of flow (Csikszentmihalyi, 1990, p. 4). The flow experience is seemingly universal, crossing activity domains and cultural boundaries (Csikszentmihalyi, 1990). Two important aspects of flow emerged: that attention and concentration is absolutely necessary; and, that the individual must strive to overcome challenges (Csikszentmihalyi, 1990). This is interesting since attention is related to immersion, while challenge is not; challenge is more of a cognitive aspect of the experience that makes flow unique. The motivational aspect related to goals makes it similar to what Witmer and Singer (1998) and Wirth et al. (2007) labeled as 'involvement'. Therefore, flow shares many commonalities with immersion; however the cognitive aspect involving goals, which are perhaps linked to involvement, as well as the focus on balancing challenges with skill level are the components of flow that make it distinct.

Csikszentmihalyi (1990) articulated eight specific elements of flow and enjoyment, as well as a ninth element that is an outcome, based on his extensive work. These elements are: (1) challenge/skill balance; (2) concentration; (3) clear goals; (4) immediate feedback; (5) merging of action and awareness; (6) control; (7) loss of self-consciousness; (8) time distortion; and (9) the autotelic experience, which is the outcome. He described these elements thusly:

> First, the experience usually occurs when we confront tasks that we have a chance of completing. Second, we must be able to concentrate on what we are doing. Third and fourth, the concentration is usually possible because the task undertaken has clear goals and provides immediate feedback. Fifth, one acts with deep but effortless involvement that removes from awareness the worries and frustrations of everyday life. Sixth, enjoyable experiences allow people to exercise a sense of control over their actions. Seventh, concern for the self disappears, yet paradoxically the sense of self emerges stronger after the flow experience is over. Finally, the sense of the duration of time is altered; hours pass by in minutes, and minutes can stretch out to see like hours. The combination of all these elements causes a sense of deep enjoyment that is so rewarding people feel that expending a great deal of energy is worthwhile simply to be able to feel it. (Csikszentmihalyi, 1990, p. 49, reproduced with permission of HarperCollins Publishers)

As previously suggested, for flow to occur there must be a structured, rule-based activity that poses some sort of challenge. The individual must have skills pertinent to the challenge, otherwise the activity will be meaningless and flow will not occur (Csikszentmihalyi, 1990). Thus, the optimal experience occurs when the challenge of the activity is matched by the skills of the individual. Once the balance between challenge and skill is achieved, there is a merging of action and awareness where “that person’s attention is completely absorbed by the activity” (Csikszentmihalyi, 1990, p. 53), and that level of involvement means that their actions become automatic; they do not need to think about what they are doing. The activity becomes “seemingly effortless” (Csikszentmihalyi, 1990, p. 54), although it actually requires a substantial amount of effort to attain and maintain. The goals need to be clear, and the feedback as an individual makes progress towards those goals needs to be immediate.

This combination of challenge/skill balance, clear goals, and immediate feedback, and the resulting merging of action and awareness focuses attention and a sense of concentration develops (Csikszentmihalyi, 1990). Draper et al. (1998) supported this notion. They proposed that as a task increases difficulty, attention becomes more focused. The problems and thoughts of everyday life do not creep into the mind. In flow, the activity is the only thing that matters.

Similarly, there is also a loss of self-consciousness as “in flow there is no room for self-scrutiny”, especially since, at this point, the individual’s skill is well-matched to the increasing challenge, therefore “there is little opportunity for the self to be threatened” (Csikszentmihalyi, 1990, p. 63). It is not bodily awareness that is eschewed; rather it is as if the individual loses themselves within the activity itself (Csikszentmihalyi, 1990). A sense of control also develops. It is not that the individual is in control, but that control over the activity is possible (Csikszentmihalyi, 1990). The individual is not worried about not being able to accomplish tasks, even though it is in the realm of possibility that they will fail. This also depends on their skill level—they must have the skill to match the needs of the activity, otherwise this sense of control is not possible (Csikszentmihalyi, 1990).

Time distortion was commonly reported by Csikszentmihalyi’s (1990) interviewees. Most reported that time went by much faster than they thought, but it can also seem to stretch and pass slowly. This transformation of time is not experienced in all activities as some have a strong temporal focus, such as surgery or running (Csikszentmihalyi, 1990). It is also unclear whether the distortion of time is simply a by-product of the flow experience or whether it actually contributes to the flow state by eliminating time as a distraction (Csikszentmihalyi, 1990). These eight elements culminate in the autotelic experience, which is “a self-contained activity, one that is done not with the expectation of some future benefit, but simply because the doing itself is the reward” (Csikszentmihalyi, 1990, p. 67). This is the optimal experience.

The different elements of flow are reliant on one-another (Csikszentmihalyi, 1975b). Based on the work of Csikszentmihalyi (1975b, 1990), these elements can be conceptualized as occurring in stages, to include the activity requirements (challenge/skill balance; clear goals; immediate feedback), resulting in focused attention on a limited stimulus field (merging of action and awareness) and an ability to filter out distractions, which allows for flow-related experiences (concentration; loss of self-consciousness; control; time distortion), thus facilitating the autotelic experience (see Figure 5).

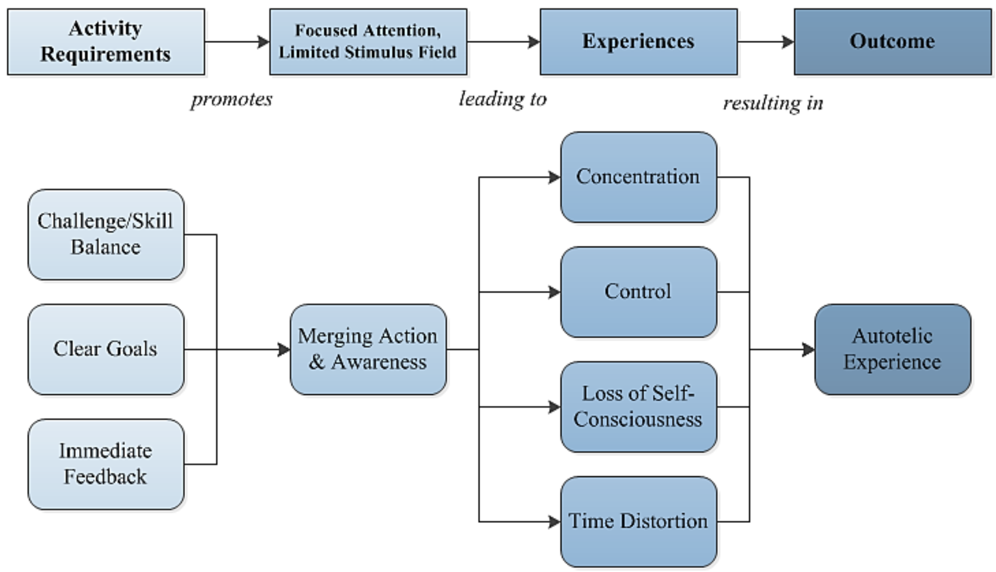


Figure 5. Csikszentmihalyi's (1975b, 1990) flow theory.

This is somewhat different from the model proposed by Chen, Wigand, and Nilan (1999), which divided flow into three stages: antecedents, which included all of the proposed activity requirements; experiences, which included merging of action and awareness, concentration, and control; and effects, which included loss of self-consciousness, time distortion, and the autotelic experience. The model proposed in Figure 5 differs in that the aspect of merging action and awareness precedes the proposed experiences, and here the only effect is the autotelic experience, which is the ultimate outcome of flow. Flow sometimes occurs at random, but it mostly only occurs when activities are structured with goals and rules, provide feedback, and allows for control—which can be achieved when a balance between challenge and skill is achieved (Csikszentmihalyi, 1990). This aligns with the proposed activity requirements. Whether an individual experiences flow in a given activity may also be due to social influences and culture, situational circumstances, and the individual's own personality (Csikszentmihalyi, 1990).

Hamari and Koivisto (2014) examined how the model of flow may be respecified in the context of gamification. In their study, the authors surveyed *Fitocracy* users. *Fitocracy* is an online social network site where users can log their exercise activities, which in turn unlocks achievements and earns them points to be used for leveling up. The *Fitocracy* users completed the Dispositional Flow Scale – 2 (Jackson & Eklund, 2002), which the authors subjected to factor analysis after respecifying the original flow model to include two factors, which differs from the model described in Figure 5: the conditions of flow (which included challenge/skill balance, control, clear goals, immediate feedback, and the autotelic experience) and the outcomes of flow (which included merging of action and awareness, concentration, time distortion, and loss of self-consciousness). They found that the model fit was adequate, but not excellent. The respecification is interesting as the autotelic experience is often described as an outcome rather than a condition of flow. The authors argued that initial interest and intrinsic motivation to use *Fitocracy* needed to be present from the outset, and this may be specific to gamification rather than entertainment games. Instead, the dispositional tendency to achieve the autotelic experience in the gamification context may instead be related to the construct of involvement, where interest in the task is necessary for flow to develop. It also may be that past flow experiences increase the likelihood of experiencing flow, which would explain why they found the outcome of dispositional flow to be a condition of flow. Given that this work focuses on the actual experience of flow, rather than dispositional tendency, in actual entertainment game context, and since involvement is addressed as a facet of the subjective gameplay experience outside of flow in a later section, the above model will not be changed to mirror Hamari and Koivisto's (2014) respecification.

Delving deeper into the notion of personality and individual differences, experiencing flow may be more difficult for those that suffer from attention and affective disorders. It also may be difficult to achieve for those who are too self-conscious. Csikszentmihalyi (1990) reported on some interesting research where those who experienced flow on a more frequent basis decreased their level of cortical activation when concentrating, meaning that paying attention to the task decreased their mental effort and these individuals were better as sustaining attention. The idea of personality influencing propensity to experience flow is similar to the idea that there are also immersive tendencies (Witmer & Singer, 1998) and trait absorption (Wirth et al., 2007). Jackson and Eklund (2002) developed the afore mentioned measure, the Dispositional Flow Scale – 2, which quantified propensity to experience flow based on the nine elements articulated by Csikszentmihalyi (1990). Ross and Keiser (2014) compared the Dispositional Flow Scale – 2 to Costa and McCrae's (1992) well-known measure of the five-factor model of personality. If propensity to measure flow is indeed an individual trait, then it should align with a specific pattern of these five factors. Previous research suggested that neuroticism and conscientiousness were strong negative and positive predictors of flow, respectively, which was also found by Ross and Keiser (2014).

As stated previously, flow incorporates some elements of immersion, yet it is distinct, perhaps in such a way that links it to the construct of involvement (Wirth et al., 2007, p. 512). Both flow and immersion require focused attention, a limited stimulus field, and reduced distractions. The descriptions of merging action and awareness and concentration are both similar to the qualities of a highly immersed state where the individual becomes absorbed in what they are doing. Indeed, loss of self-consciousness refers to becoming one with the task, which is also a characteristic of highly-immersed states, although it is not as transcendent as presence. Finally, the concept of control as it pertains to flow is complex. Slater and Wilbur's (1997) operationalization of immersion includes unobtrusive controls and naturalistic inputs, which implies mastery of the task—hence, control—but control in the sense of flow goes beyond skill mastery—which is, indeed, a necessary part of the challenge/skill balance element of flow. In flow, control is more involved as it is a sense of being in control of the situation, without fear of failure even though it is entirely possible.

Flow is set apart from immersion due to its more cognitively-focused aspects, to include its prevalence in a rule-based context with clear goals and immediate feedback to guide an individual on a clear progression through the activity. The importance of goals to motivate the individual is a part of Slater and Wilbur's (1997) plot factor of immersion as well as a part of Witmer and Singer's (1998) notion of involvement, although the aspect of goals is not well-developed in either theory. Furthermore, the involvement factor found by Witmer et al. (2005) did involve elements of being challenge and being engaged "cognitively, physically, or emotionally" (p. 299). Perhaps these elements of goals and feedback strengthen involvement, which in turn strengthens immersion, but it is the *balance* between challenge and skill that is likely the critical element for transcending involvement and achieving the flow experience.

Finally, the time distortion and autotelic experience aspects of flow are entirely unique and not described in the classic immersion literature (e.g., Slater & Wilbur, 1997; Witmer & Singer, 1998). Therefore, flow incorporates elements of immersion, yet goes beyond it by introducing cognitive-focused elements that drive the experience deeper, which likely stem from involvement. The more flow-specific, transcendental states occur as the challenge/skill balance progresses. Indeed, Draper et al. (1998) agreed with this notion, citing the earlier work of Fontaine (1992) in that there are differences between flow and presence, and likely immersion, in which flow is a much more focused, precise state with a special set of specific antecedents.

There is an incredible amount of interest in flow as it pertains to the subjective gameplay experience (Boyle et al., 2012). A simple search for "flow theory and video games" in Google Scholar returns over 171,000 results (http://www.scholar.google.com; December 30, 2014). The impetus for this considerable body of research is readily apparent. Games are rule-based activities that increase in difficulty, encouraging the player to simultaneously increase their skills. They also feature clear goals and immediate feedback that an individual is intrinsically-motivated to participate in. Given that games are also a play activity, they effectively capture all of the flow-promoting design characteristics listed by Csikszentmihalyi (1990).

It is not atypical for game players to report states of deep concentration, where they become so focused that they no longer notice the game controller in their hands and become one with the game experience. This is a highly-immersed and involved state, which becomes flow as the experience deepens, as the challenges increase and the player struggles to successfully surmount them. A sense of flow in games goes beyond being simply taken in by gameplay (i.e., immersion and involvement), but it does not necessarily involve physical transplantation into the game-world (i.e., presence). Many game players lose their sense of time, which is a commonly-reported facet of flow. Rau, Peng, and Yang (2006) found that both novice and expert gamers experienced some form of time distortion during gameplay in an online role-playing game (RPG) when it lasted 60 minutes, and expert gamers experienced time passing more quickly than the novices. The player will also become confident in their abilities and ride high on that sense of control, resulting in the autotelic experience. Therefore, games are certainly a medium in which flow is readily experienced.

Given the above connection, researchers from many different fields are interested in studying flow states in games. Clinical psychologists study flow as it relates to video game addiction (Rau et al., 2006; Jennett et al., 2008; Jo, Lee, & Park, 2008), while those in the entertainment industry seek to understand flow so that they may design more commercially-viable games (Chen, 2007). For example, Jenova Chen studied flow theory extensively (see Chen, 2007), and applied his understanding of flow theory to game design. As a founding member of thatgamecompany (http://www.thatgamecompany.com/), these principles have guided the development of highly-successful games that encapsulate the 'in the zone' playing experience.

Educators are interested in how flow can be harnessed to make game-based learning experiences more effective. Creating flow in the classroom can improve engagement with schoolwork (Shernoff, Csikszentmihalyi, Schneider, & Shernoff, 2003). According to Admiraal, Huizenga, Akkerman, and ten Dam (2011), aspects of flow align well with high-quality instruction, especially the element of challenge/skill balance. Since experiencing flow is its own reward, flow in learning should make education more fun and intrinsically rewarding. The researchers were particularly interested in this as they believed flow experienced while playing a serious game in the classroom should increase student engagement and reduce drop-out rates. Admiraal et al. (2011) conducted a study to reveal if there was a relationship between flow and game-based learning as it affected student learning outcomes. They found that flow affected in-game performance, but did not improve scores on an assessment. It is important to note that the researchers measured flow by observing in-game activities, such as "using the tools" and "navigating the urban environment" (Admiraal et al., 2011, p. 1189). Arguably, they were not accurately measuring flow at all, as this objective measure of a subjective experience had no theoretical connection to flow theory itself—simply engaging in acts of gameplay is not indicative of the optimal experience.

Pavlas (2010) was also interested in how experiencing flow while playing a serious game affected learning outcomes. He proposed a model of the subjective gameplay experience, with flow at its core. From flow emerged both a sense of play and of immersion. It is worth noting that the idea that immersion springs from flow runs contrary to other models of the subjective gameplay experience (see the upcoming discussion of Jennett et al., 2008 and Brockmyer et al., 2009). Immersion was not a focus of this work and this aspect of the model was not tested. Pavlas (2010) proposed that play, enjoyment, in-game performance, and flow positively influenced learning.

To measure flow, Pavlas (2010) utilized a validated, short form of the Flow State Scale – 2 (Jackson, Martin, & Eklund, 2008), which is the state-dependent correlate of the Dispositional Flow Scale – 2. He also measured ratings of video game self-efficacy (VGSE), which pertains to an individual's confidence in their abilities to play a video game (Pavlas, Heyne, Bedwell, Lazzara, & Salas, 2010). Also obtained was a play experience score, which was measured by the Play Experience Scale that was developed as a part of Pavlas' (2010) research. The results indicated that VGSE was a predictor of in-game performance scores and flow. Furthermore, a statistically-significant combination of play experience score, VGSE, emotional experience, and age predicted 74% of the variance in flow score. Both in-game performance scores and flow also were positively predictive of declarative knowledge scores after playing an instructional game. Flow, however, was not found to mediate the relationship between the in-game scores and learning. This finding suggested that flow is not the mechanism through which increased game performance improves learning, and that both game performance and flow are independent predictors of learning.

In summary, flow is an important part of the subjective gameplay experience and is closely related to, yet distinct from, the experience of immersion. It also appears to be directly linked to involvement, which may become flow when a balance between the challenge of the game and the skill of the player is achieved. Understanding flow—and its relationship to immersion and involvement—has implications for researchers, game designers, and educators.

Absorption. Wirth et al. (2007) suggested that an individual's propensity toward absorption may be a predictor of whether an individual becomes immersed. Absorption has its roots in the work of Tellegen and Atkinson (1974), who were interested in creating a valid and robust measure of hypnosis susceptibility. They were interested in specifying the dimensions pertinent to this particular experience, linking those dimensions to personality traits, and exploring how those dimensions correlated with susceptibility to hypnosis. They first created a set of 71 items derived from several different relevant measures. They recruited 481 female undergraduate students for their study, in which participants responded to a variety of scales, to include items from their new measure. An initial factor analysis identified 13 factors that explained 99% of the variance in the data, which was reduced to 11 factors after removing items whose factor loadings did not exceed .30. A second analysis which compared the items to Block's (1965, as cited in Tellegen & Atkinson, 1974) Ego Resiliency and Ego Control scales produced a final three-factor solution to describe susceptibility to hypnosis, one of which was absorption.

Tellegen and Atkinson (1974) found that the absorption factor was the only one to be significantly (and positively) correlated with measurements of ability to become hypnotized. Based on the items that loaded onto the factor, they defined absorption as "a 'total' attention, involving a full commitment to available perceptual, motoric, imaginative and ideational resources to a unified representation of the attentional object" (Tellegen & Atkinson, 1974, p. 274). Tellegen and Atkinson (1974) believed that the experience of absorption was likely correlated with the perception that whatever an individual was focusing on was real to them, that which is normally distracting was not, and that the individual experienced an altered self-perception of reality. They also believed that absorption is a trait of an individual that is likely related to imagination and creativity as well as Costa and McCrae's (1992) openness to experience personality factor. The items in the factor analysis became a part of the Tellegen Absorption Scale (TAS; Tellegen & Atkinson, 1974), which measures trait absorption. Glisky, Tataryn, Tobias, Kihlstrom, and McConkey (1991) found that that the TAS was highly correlated with another measure of hypnotic susceptibility and with the openness to experience personality factor.

Wild, Kuiken, and Schopflocher (1995) further examined trait absorption. They proposed that a state of absorption is the outcome of being involved in a task when either of the following four experiences occurs: the aesthetic experience (Bullough, 1912 and Dewey, 1934, as cited in Wild et al., 1995); flow (Csikszentmihalyi, 1975a, 1975b, 1990); intrinsic motivation (Deci & Ryan, 1985 and Lepper & Greene, 1978, as cited in Wild et al., 1995); and the peak experience (Maslow, 1962, as cited in Wild et al., 1995). They noted that these are similar, yet distinct constructs. For example, the aesthetic experience and peak experience will happen without engaging in a specific behavior, while flow and intrinsic motivation do require deliberate action on the part of the individual (Wild et al., 1995). Also, flow is described as a positive emotion, however the authors maintained that it is possible to be involved with a task when it is a negative experience.

Wild et al. (1995) proposed that trait absorption has both a motivational component, where an individual will self-select into potentially absorbing tasks, and a cognitive component, where those who are more prone to absorption will have faster reaction times to tasks such as figure-ground comparisons and be more able to experience cross-modal elaboration, in which an individual experiences color in response to music or sounds. Trait absorption, as measured by the TAS (Tellegen & Atkinson, 1974) was found to be significantly and positively correlated with a measure of the openness to experience personality factor, particularly for the aesthetic subscale. It was also linked to faster reaction times in difficult figure-ground differentiation tasks and was significantly and positively correlated with cross-modal elaboration. Finally, trait absorption was also found to be significantly and positively correlates with likelihood to seek out absorptive experiences, such as listening to music or watching plays. These results suggested that there is both a motivational and cognitive component to trait absorption, which is reflected in real-world behaviors. It is likely important, then, to measure trait absorption as an important individual difference that may affect how a player experiences a game.

Wirth et al. (2007) briefly discussed absorption as it related to spatial presence, however Sas and O'Hare (2004) examined the connection between presence and absorption much more in-depth, along with the related concepts of empathy and imagination. They defined presence as a "psychological phenomenon through which one's cognitive processes are oriented toward another world, either technologically mediated or imaginary, to such an extent that he or she experiences mentally the state of being (there), similar to one in the physical reality, together with an imperceptible shifting of focus of consciousness to the proximal stimulus located in that other world" (Sas & O'Hare, 2004, p. 524).

Sas and O'Hare (2004) conducted a study in which they hypothesized that those who are more empathetic and able to become absorbed would be more likely to experience presence. Of interest were scores on the Interpersonal Reactivity Index (Davis, 1980, as cited in Sas & O'Hare, 2004), the TAS (Tellegen & Atkinson, 1974), and the Creative Imagination Scale (Barber & Wilson, 1978). They also utilized a single question to measure willingness to experience presence, which stated: "To what extent were you willing to be transported to the virtual world?" (Sas & O'Hare, 2004, p. 529). In their study, participants were asked to explore a desktop-based VE for 20 minutes in search of a thief. Sas and O'Hare (2004) found several positive correlations between presence and empathy (r = .49), absorption (r = .39), creative imagination (r = .35) and willingness to suspend disbelief (Sas & O'Hare, 2004, p. 532). They also conducted a multiple regression to determine which cognitive factors were more influential in predicting scores of presence. They found that creative imagination and absorption were highly correlated (r = .56), and suggested that they were likely the same construct (Sas & O'Hare, 2004, p. 533). Although there are different guidelines as to the cut-off value for when two constructs are actually the same based on correlation, this value is still low compared to the r = .70 threshold suggested by Tabachnick and Fidell (2007). Collapsing the two scores into one general factor of absorption, the combination of absorption along with willingness to suspend disbelief predicted 45% of the variance in presence score. Although creative imagination was not found to be a significant factor, the researchers believed it to be important, just not as important as absorption and willingness to suspend disbelief. Sas and O'Hare's (2004) results suggested that presence and absorption tendency were somewhat positively associated, which merits further examination, while creative imagination was strongly, positively correlated with absorption tendency.

Absorption is not only a trait; it is also a state, according to Agarwal and Karahanna (2000). The authors were interested in connecting this state of absorption, which they labeled "cognitive absorption" (Agarwal & Karahanna, 2000, p. 666), to software usage. Their operationalization of cognitive absorption was based on Tellegen and Atkinson's (1974) definition of absorption, but it also incorporated elements of flow (Csikszentmihalyi, 1975b, 1990) and cognitive engagement (Webster & Ho, 1997, as cited in Agarwal & Karahanna, 2000). They believed that trait absorption was an antecedent of the state of cognitive absorption. They also stated that certain elements of flow, specifically control, attentional focus, curiosity, and intrinsic motivation were important parts of the cognitive absorption experience. Cognitive engagement, as defined by Webster and Ho (1997, as cited in Agarwal & Karahanna, 2000) referred to a “state of playfulness” that is essentially flow, without the control requirement (p. 669). Their definition of cognitive absorption was essentially the result of blending the theory of absorption (Tellegen & Atkinson, 1974) with the theory of flow (Csikszentmihalyi, 1975a). Agarwal and Karahanna (2000) defined cognitive absorption as “a state of deep involvement with software” (Agarwal & Karahanna, 2000, p. 673). The authors listed five dimensions of cognitive absorption, which included:

1. Temporal dissociation, which was the “inability to register the passage of time while engaged in interaction” (Agarwal & Karahanna, 2000, p. 673).
2. Focused immersion, which was the “experience of total engagement where other attentional demands are, in essence, ignored” (Agarwal & Karahanna, 2000, p. 673).
3. Heightened enjoyment, which was “capturing the pleasurable aspects of the interaction” (Agarwal & Karahanna, 2000, p. 673).
4. Control, which represented “the user's perception of being in charge of the interaction” (Agarwal & Karahanna, 2000, p. 673).
5. Curiosity, which involved “tapping into the extent the experience arouses an individual’s sensory and cognitive curiosity” (Agarwal & Karahanna, 2000, p. 673).

Cognitive absorption clearly incorporates elements of both immersion and flow. The dimensions of temporal dissociation and control are both very similar to their counterparts in flow, and the concept of heightened enjoyment is essentially the autotelic experience. What appears to stand out is the concept of curiosity, which the authors attributed to the flow experience on the suggestion of Trevino and Webster (1992, as cited in Agarwal & Karahanna, 2000).

Agarwal and Karahanna (2000) also considered several different technology usage theories, to include Innovation Diffusion Theory (Brancheau & Wetherbe, 1990, as cited in Agarwal & Karahanna, 2000), the Technology Acceptance Model (TAM; Davis, 1989), the Theory of Reasoned Action (Ajzen & Fishbein, 1980, as cited in Agarwal & Karahanna, 2000), and the Theory of Planned Behavior (Ajzen, 1985, as cited in Agarwal & Karahanna, 2000). All of these theories centered on that notion that what an individual believes about a piece of software will affect how they use it.

Agarwal and Karahanna (2000) proposed that cognitive absorption was a downstream predictor of both perceived ease of use and perceived usefulness, which are factors derived from Davis' (1989) TAM, which ultimately will predict an individual's intention to use a specific piece of software. Perceived ease of use of the software and perceived usefulness is also influenced by the individual's self-efficacy for that software. Agarwal and Karahanna (2000) also believed that cognitive absorption would have an effect on workload, where mental workload for a given task should be lower as being in a state of cognitive absorption would mean that all attentional resources are devoted to the task at hand and are available for use. They proposed a model of cognitive absorption that was couched within the TAM (Davis, 1989). It is interesting to note that trait absorption was absent from this model, despite the authors' acknowledgement that they believed it is an antecedent of cognitive absorption.

Agarwal and Karahanna (2000) created the Cognitive Absorption Scale with subscales representing each of the five proposed dimensions. Two-hundred and eighty-eight students were asked to reflect on their use of the internet and to respond to the items in the measure. Their proposed model was evaluated through structural equation modeling (SEM) using partial least squares because of that method of analysis' ability to be used in low sample sizes and because it has less-strict requirements dictating acceptable residual distributions. They found their model to be an acceptable fit, with adequate factor loadings on the proposed constructs, as well as having high internal reliability.

Cognitive absorption did have a positive influence on the five proposed factors, as well as on perceived usefulness and perceived ease of use of the internet. Self-efficacy did not have a major role in the model compared to the other constructs. This may be because Agarwal and Karahanna (2000) measured generalized self-efficacy rather than application-specific self-efficacy. Given that Pavlas (2010) found VGSE to be significantly correlated with flow, perhaps inclusion of application-specific self-efficacy as a downstream predictor of cognitive absorption could be included in future models of the construct. Playfulness and personal innovativeness were significant, positive predictors of cognitive absorption. They concluded that the experience of cognitive absorption was found to change behavioral beliefs, such as perceived usefulness and perceived ease-of-use, which was also influenced by personal traits, such as playfulness and personal innovativeness (Agarwal & Karahanna, 2000).

In summary, trait absorption (Tellegen & Atkinson, 1974) was found to be significantly and positively correlated with presence, as were two other individual differences: creative imagination and willingness to suspend disbelief (Sas & O'Hare, 2004). Therefore, these three individual differences are important to consider when attempting to understand the subjective experience of gameplay. Furthermore, the state of cognitive absorption (Agarwal & Karahanna, 2000) is also worth considering, however the theory behind cognitive absorption is extremely similar to (and is based directly on!) that of flow (Csikszentmihalyi, 1975a), to the point that they are likely the same construct.

Game immersion. Brown and Cairns (2004) strove to define immersion with respect to gaming. They generated a model of the subjective gameplay experience based on interviews of seven experienced game players. Their model is entirely atheoretical; however, it does have excellent face validity as closer examination reveals that their model resembles prior work, although their labels are problematic and confusing given the two decades of prior research. Brown and Cairns (2004) stated that immersion is "used to describe the degree of involvement with a game" (p. 1298). This proposition disagreed with what was proposed by Witmer and Singer (1998), in which immersion and involvement were separate but related constructs, yet this disagreement may be in name only. They proposed three degrees of involvement which are progressive: it begins with engagement, which leads to engrossment, which then leads to total immersion. They believed immersion to be the final state, which was described as a state of presence (Brown & Cairns, 2004). For each stage, Brown and Cairns (2004) proposed specific "barriers" that must be overcome before progressing to the next level of involvement (p. 1298). These barriers were a combination of individual differences, such as game preferences, game-specific elements, and contextual factors. See Figure 6 for an illustration of their model.

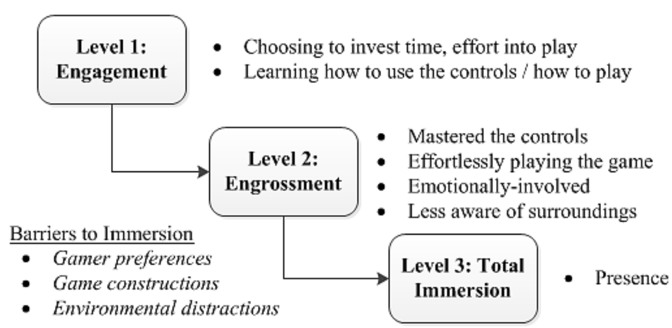


Figure 6. Brown and Cairns' (2004) model of game immersion.

According to Brown and Cairns (2004), for engagement to occur, the player must consciously invest time, effort, and attention into playing. They also must have access to the game, meaning they are able to choose something they enjoy—which is similar to the domain-specific motivational element of presence described by Wirth et al. (2007)—and are able to master the controls so that they are not a hindrance to the experience and will not consume any mental resources. Therefore, the game needs to be usable, which aligns with the usability-driven interface requirements that minimize distraction and allow for focused attention as found in Witmer and Singer (1998). As the player becomes more involved, they become engrossed and less aware of their surroundings (Brown & Cairns, 2004). Barriers to engrossment included elements of game design, such as overall quality of the game where better stories will likely lead to more emotional involvement in the storyline, while the barriers to total immersion are empathy and atmosphere (Brown & Cairns, 2004). This experience is rare, with the other two levels being far more likely to occur. To reach total immersion, the player must become empathetic with the main character in a cohesive VE. Interestingly, Brown and Cairns (2004) also believed that immersion will be driven by "the number of attentional sources needed as well as the amount of each attentional type" (Brown & Cairns, 2004, p. 1299). This is similar to the idea that the number of senses modeled will increase immersion (e.g., Slater & Wilbur, 1997), and highlights the role of attention (Witmer & Singer, 1998). The spirit of Brown and Cairns' (2004) work aligns with previous research in the area of immersion and presence, although their labels and operationalizations do not.

Jennett et al. (2008) expanded greatly on the work of Brown and Cairns (2004), tying their new definition of immersion more strongly to theory. They believed that games are a viable way to study immersion because games themselves are designed in such a way to deeply immerse players. They maintained that immersion is a subjective experience in which players provide complete attention to and “lose” themselves in a game, to the point where outside distractions and time passing goes unnoticed, and that the player feels as if they are actually “in the game” (Jennett et al., 2008, p. 641).

There are some obvious points of overlap between Jennett et al.’s (2008) conceptualization of immersion and that of Csikszentmihalyi’s (1975b, 1990) flow theory. They stated that immersion involves: “Lack of awareness of time;” and “Loss of awareness of the realworld” (Jennett et al., 2008, p. 642). These elements very clearly align with previous research on flow in games (e.g., Rau et al., 2006). Jennett et al. (2008) acknowledged that this is one point of conceptual overlap, as is the ability to involve a person in the game by providing an appropriate amount of challenge, which is another critical factor driving the flow experience (Csikszentmihalyi, 1990). Not only this, they believed that immersion is “evidently a precursor for flow because that sense of being so involved that nothing else matters is practically a colloquial definition of immersion” (p. 642), yet flow is still unique in that is an “optimal experience” in that it is difficult to attain, while immersion is “not always so extreme” (Jennett et al., 2008, p. 642). The notion that immersion is a lower cognitive state than flow is later supported by Brockmyer et al. (2009) and Procci, James, and Bowers (2013).

Jennett et al. (2008) deemed the following to be the requirements for experiencing immersion while playing a game: concentration, challenge, control, emotional involvement, and real-world dissociation. Again, these elements have an enormous amount of overlap with the flow experience, to include concentration, challenge, control, and real-world dissociation, to a certain extent. They also made the point that not all games will make the player experience presence despite being fun (e.g., a player may be immersed in the speed-puzzle game *Tetris*, but does not feel as if he or she were actually there), and not all games will be immersive (e.g., a VE that is boring). This is a fairly confusing statement as Jennett et al. (2008) did not keep their use of terminology consistent with previous work. It may be more clear to say that it is possible to be completely involved—Witmer and Singer's (1998) definition, not Jennett et al.'s (2008)—by gameplay to the point of flow, but not to feel present, just as it may also be possible to feel a sense of immersion and presence in a VE, but not be caught up in the thrilling experience of playing a game.

While the terminology's disconnect between the game immersion model (Brown & Cairns, 2004; Jennett et al., 2008) and previous research (e.g., Slater & Wilbur, 1997; Witmer & Singer, 1998) is disorienting, the following is apparent: The subjective experience of gameplay is progressive. It begins with focused attention, and greater levels of immersion allow a player to experience presence. Involvement and flow likely behave in a similarly progressive way.

Game engagement. Jennett et al. (2008) took an earnest first step forward in describing the subjective gameplay experience, although their construct operationalizations were largely inconsistent with past research in the area. They also failed to distinguish unique constructs (e.g., immersion, flow) in their model of game immersion. Brockmyer et al. (2009) addressed this need. Rather than using Jennett et al.'s (2008) term *game immersion*, they described the subjective gameplay experience as *game engagement*. They stated that game engagement is "a generic indicator of game involvement…that can be conceptualized as representing a progression of ever-deeper engagement in game-playing" (Brockmyer et al., 2009, p. 624). Similar to Brown and Cairns (2004) and Jennett et al. (2008), Brockmyer et al. (2009) believed that game engagement is a progressive experience, where immersion, presence, flow, psychological absorption, and dissociation were all a part of overall game engagement, and that they were all unique. The model of game engagement leveraged the following definitions:

- *Immersion* is "becoming engaged in the game-playing experience while retaining some awareness of one's surroundings (Banos et al., 2004; Singer & Witmer, 1999)" (Brockmyer et al., 2009, p. 624).
- *Presence* is "(1) being in a normal state of consciousness and (2) having the experience of being inside a virtual environment (Mania & Chalmers, 2001; Mikropoulos & Strouboulis, 2004; Ryan, Rigby, & Przybylski, 2006; Tamborini & Skalski, 2006)" (Brockmyer et al., 2009, pp. 624-625).
- *Flow* is "the term used to describe the feelings of enjoyment that occur when a balance between skill and challenge is achieved in the process of performing an intrinsically rewarding activity (Csikszentmihalyi & Csikszentmihalyi, 1988; Moneta & Csikszentmihalyi, 1996, 1999). Having a specific goal and an immediate performance feedback structure increase the likelihood of flow…Flow states also include a feeling of being in control, being one with the activity, and experiencing time distortions" (Brockmyer et al., 2009, p. 625).
- *Psychological absorption* is an altered state in which "there is a separation of thoughts, feelings, and experiences and affect is less accessible to consciousness (Glicksohn & Avnon, 1997)" (Brockmyer et al., 2009, p. 625), and is essentially the same as Agarwal and Karahanna's (2000) definition of cognitive absorption.

- *Dissociation* is “a clinical term…defined as ‘the lack of normal integration of thoughts, feelings, and experiences into the stream of consciousness and memory’ (Bernstein & Putnam, 1986, p. 727)” (Brockmyer et al., 2009, p. 625), where absorption is likely dissociation’s non-pathological counterpart.

Brockmyer et al. (2009) believed that most game players experience some low level of immersion. They also believed that most of those same players may also be able to experience presence if the conditions support it. Beyond that, a player might achieve the mind-altering experience of flow or cognitive absorption. Please note that Brockmyer et al. (2009) referred to the state of cognitive absorption as ‘absorption’. The authors did not specifically articulate that their notion of absorption was the same as Agarwal and Karahanna’s (2000) concept of cognitive absorption, however they will be considered equivalent in this discussion as they are both explicitly purported in both papers to be the state manifestation of Tellegen and Atkinson’s (1974) trait absorption.

While previous research highlighted the overlap between flow and cognitive absorption (Agarwal & Karahanna, 2000), Brockmyer et al. (2009) stated that the two are theoretically at odds based on two points. The first difference is affect; games can be frustrating and induce anxiety, however both frustration and anxiety are antithetical to the flow experience (Csikszentmihalyi, 1990). They believed that frustration and anxiety were instead a part of cognitive absorption. The other difference is motivational; flow is the result of intrinsic motivation, while cognitive absorption, they stated, is not. Therefore, Brockmyer et al. (2009) proposed that achieving flow occurs when the gameplay experience is positive and intrinsic, while cognitive absorption occurs when motivations are extrinsic and gameplay is anxietyinducing or frustrating. Lastly, the authors cited previous research in which absorption and presence were not correlated (Murray, Fox & Pettifer, 2007, as cited in Brockmyer et al., 2009), which suggested that an individual does not need to experience presence in an order to achieve these higher states (i.e., flow and cognitive absorption). This is similar to the findings of Jennett et al. (2008), which supported the notion that there is a common low-level game engagement state that branches into two possible concurrent streams—one for presence and one for the altered cognitive states. According to Brockmyer et al. (2009), the altered cognitive state that is experienced depends on the affect and motivation associated with the game. See Figure 7 for an illustration of this model.

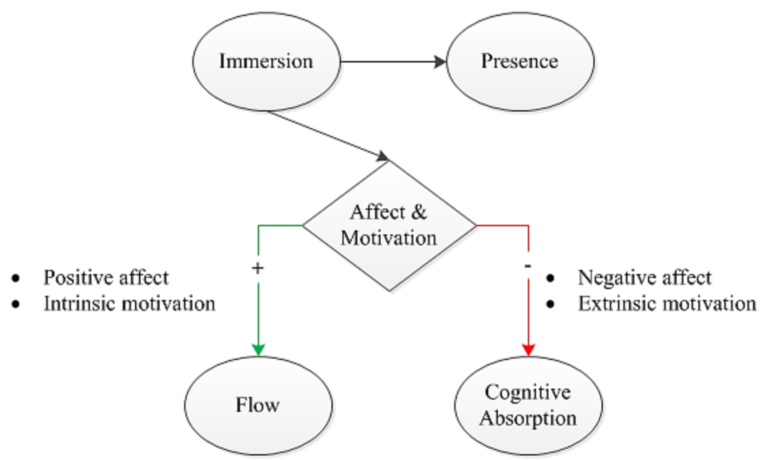


Figure 7. Brockmyer et al.'s (2009) model of game engagement.

After a thorough literature review, as well as feedback from focus groups, Brockmyer et al. (2009) crafted a measure that quantified aspects of immersion, presence, flow, absorption, and dissociation. They used both Rasch analysis and classical tests to validate their game engagement measure. After a series of pilot studies, they produced the 19-item Game Engagement Questionnaire (GEQ). One-hundred and fifty three junior high school students completed the

GEQ as part of a larger study on media usage. Half of the participants were recruited from an urban school district, while the other half were recruited from a rural school district. No differences were found between these two groups with respect to total GEQ score, therefore their responses were collapsed into one sample.

Rasch analysis revealed Brockmyer et al.'s (2009) measure to be highly reliable. Rasch analysis was also used to determine item hierarchy based on how difficult participants found each item to agree with. The data suggested that the immersion items were the easiest to agree with, followed by the presence, flow, and then absorption items. The items within each grouping also were found to be interchangeable, but the groupings themselves were not (e.g., immersion items were interchangeable with immersion items in this ranked hierarchy, but not with presence items). The authors believed that the GEQ represented a hierarchical progression of game engagement states.

After establishing a reliable measure, Brockmyer et al. (2009) conducted a second study to determine if higher scores on the GEQ actually measured greater levels of game engagement, while controlling for both age and tendency to experience dissociation as measured by the Dissociative Experiences Scale (DES; Bernstein & Putnam, 1986, as cited in Brockmyer et al., 2009). They recruited 107 male undergraduate students who played at least three hours of games a week. They believed that an all-male sample that played that particular amount of games on a weekly basis would most likely be able to experience flow or absorption in the laboratory setting in which this study was conducted. Participants completed a background questionnaire as well as the GEQ regarding their typical gameplay experiences, a measure of trait aggression, and the DES. Participants then played the first-person shooter game *S.T.A.L.K.E.R.: Shadow of*

Chernobyl for 30 minutes.

Brockmyer et al. (2009) hypothesized that those who were more engaged with gameplay would be less susceptible to outside distractions, which was supported by Wirth et al. (2007) who proposed that individuals who are immersed and motivated will be similarly less affected by external stimuli. After 25 minutes of play, the researchers played an audio tape with three statements: “Excuse me, did you drop your keys?”, followed by a somewhat louder “Excuse me, did you drop your keys?”, and then an even louder “Did you drop your keys?” (Brockmyer et al., 2009, p. 631). The researchers recorded at what point the participants responded. After gameplay, they completed the GEQ a second time.

Brockmyer et al. (2009) conducted a series of multiple regressions to see if GEQ score predicted participants’ responses to each of the statements, controlling for age at step one, tendency to dissociate at step two, and typical GEQ score at step three. Their use of regression for these data is somewhat problematic as their outcome variable was not actually quantitative; instead the predicted variable was to which statement they responded (i.e., statement 1, 2, or 3), which can be considered an ordinal variable as best. They did not find any significant results predicting a response for statements two or three, but they did find a significant result for whether participants responded to the first statement: Participants with higher DES and lower GEQ scores were more likely to respond to the first statement. Brockmyer et al. (2009) also computed a number of correlations. They found that the typical GEQ score was highly correlated with the post-game GEQ score (r = .72, p < .01), and that the typical GEQ score and the DES were significantly, positively, but not strongly correlated, which suggested that they are similar, yet distinct constructs (r = .39, p < .01; Brockmyer et al., 2009, p. 631).

Brockmyer et al. (2009) stated that the GEQ is reliable and valid for measuring game engagement in violent video games. Furthermore, their results suggested that game engagement exists along a continuum. They also suggested that more research is needed with a more ethnically-diverse sample and in a wider age range. More research is also needed to improve the items with respect to the higher game engagement states of flow and absorption, perhaps with the help of experts in the area, and they even explicitly suggest that the GEQ should be compared to the TAS (Tellegen & Atkinson, 1974).

Procci et al. (2013) more recently utilized the GEQ in a study of the gameplay experience. Their review of the literature highlighted three key individual differences that have been of interest to gaming scientists: gender, age, and game experience. They also cited previous research that suggested individual differences, such as gender, affected the flow experience. The goal of their study was to determine if the individual differences that affected gameplay preferences similarly affected the subjective gameplay experience. Undergraduate students were recruited to play a browser-based platform puzzle game. This particular game was chosen as previous research had found that this type of game was equally enjoyable to males and females (Procci, Bohnsack, & Bowers, 2011). Participants played the game for at least 15 minutes and then they responded to the items on the GEQ. First, the researchers were interested in broadly testing Brockmyer et al.'s (2009) model, specifically if low levels of game engagement (i.e., an average of the scores on the immersion and presence subscales of the GEQ) predicted higher levels of game engagement (i.e., an average of the scores on the flow and cognitive absorption subscales of the GEQ).

Procci et al. (2013) found that low-level engagement was strongly predictive of high-level engagement, where for every one-point increase of low game engagement score, the high game engagement score doubled. The researchers also examined whether individual differences, which included gender, age, and experience as quantified by VGSE, predicted both low- and high-level game engagement, and whether those individual differences moderated the relationship between them. Low-level engagement was not predicted by individual differences, however high-level engagement was. Age was the only individual difference that was significantly influential, where high-level game engagement decreased by nearly a full point for every increase in year of age. Age also weakened the relationship between low- and high-level game engagement: if gender, VGSE, and low-level game engagement were held constant, the older individual would experience lower levels of high-level game engagement than the younger individual. These results provided support for the Brockmyer et al. (2009) model in that game engagement exists on a progressive gradient, and found that age influenced higher levels of game engagement, just as Pavlas (2010) found age to be a significant contributor to predicting flow scores.

The subjective gameplay experience. Immersion, presence, flow, and cognitive absorption have all been discussed as potential contributors to the subjective gameplay experience. The work of Brown and Cairns (2004) and Jennett et al. (2008) attempted to consolidate the theories of immersion, presence, and flow into the concept of game immersion. While their work is noteworthy because it began with interviews of actual game players, many of the construct's labels were misaligned in light of decades of past research. They did propose that the subjective gameplay experience was hierarchical, which was supported by Brockmyer et al. (2009) and Procci et al. (2013).

Brockmyer et al.’s (2009) work is notable as it made a clear distinction between the four different constructs as well as laid the foundation for an examination of their interrelated nature, where low levels of game engagement included immersion and presence, while high levels of game engagement included flow and cognitive absorption. It is worth discussing, however, whether flow and cognitive absorption are the same construct. Brockmyer et al. (2009) argued that flow and cognitive absorption differed based on affect and motivation, which may be an over-simplification. Regarding affect, frustration and anxiety are destructive to the flow experience, but an individual can be somewhat frustrated and somewhat anxious while still engaged in flow as long as they are able to adjust accordingly (Csikszentmihalyi, 1990). Flow is not a static experience, but one that wavers along a central channel, briefly touching both positive and negative emotions, where the optimal experience occurs when most of the time is spent at the balance point between challenge and skill (Csikszentmihalyi, 1990). One can persist, despite being somewhat frustrated, because they are driven by their motivation to overcome the challenge and achieve the autotelic experience (Csikszentmihalyi, 1990). Regarding motivation, games often impose both intrinsic and extrinsic motivators (Recchia, 2013). While a flow state is inherently intrinsic and even though games may not be, the drive to continue to play for the sake of playing is (Csikszentmihalyi, 1975b). Therefore, the dichotomy between flow and cognitive absorption posed by Brockmyer et al. (2009) is artificial.

Furthermore, Tellegen and Atkinson (1974) defined the state of absorption in a way that was very similar to the experience of flow, including focused attention, minimized distractions, and an altered perception of time. Also, as previously noted, both Wild et al. (1995) and Agarwal and Karahanna (2000) asserted that flow is a part of absorption. Furthermore, Agarwal and

Karahanna's (2000) five dimensions of cognitive absorption dovetail with the experience of flow. Flow and the state of cognitive absorption are indeed the same construct. It is then pertinent to only focus on flow rather than also examining cognitive absorption, given the lack of dissimilarity found elsewhere in the literature and the contrived distinction made between the two by Brockmyer et al. (2009). Therefore, the constructs with which to examine the subjective gameplay experience should focus on immersion, involvement (which was proposed by Witmer and Singer in 1998, but seems to have been since ignored), presence, and flow.

The subjective gameplay experience: Putting it all together

A review of the literature on the subjective gameplay experience revealed a complex, evolving theory centered on presence and immersion that was influenced by several different fields, from those who designed information systems with an engineering approach to positive psychologists. The result was chaos, where constructs began to be used interchangeably in the literature despite being distinct. This review attempted to step through that evolution to better understand how one might describe the subjective gameplay experience.

One approach would be to consolidate all of the possible theories into one mega-model where constructs come together based on overlapping mechanisms. While this review has revealed enough information and generated numerous illustrations of constructs, the end result would be a massive, cumbersome model. A similar problem was found in the information systems literature regarding the TAM (Davis, 1989). Viswanath Venkatesh spent roughly a decade expounding upon the TAM (Venkatesh, 2000; Venkatesh & Davis, 2000; Venkatesh & Bala, 2008) and also compared the TAM to seven other similar models to systematically create a consolidated model, which he named the Unified Theory of Acceptance and Use of Technology

(UTAUT; Venkatesh, Morris, Davis, & Davis, 2003). The UTAUT is an extremely complex model, which while realistic, is unwieldy. In his critique of the TAM, Bagozzi (2007) stated that "the exposition of UTAUT is a well-meaning and thoughtful presentation. But in the end we are left with a model with 41 independent variables for predicting intentions and at least eight independent variables for predicting behavior …The [information systems] field risks being overwhelmed, confused, and misled by the growing piecemeal evidence" (p. 245). Therefore, while a similar approach to Venkatesh's could be applied to the study of the subjective gameplay experience, the resulting model may ultimately add to the growing confusion. Rather than creating a mega-model, this work's contribution is to create a unified front regarding the construct operationalizations of the subjective gameplay experience, which will serve as a well-reasoned step forward out of the chaos. The first step in this process is to identify the relevant constructs, which was completed in the previous section. The next steps are to clearly define these constructs based on the past scientific literature, and to provide a general, but testable model of how they may interact.

Definitions. The constructs this review has identified as relevant to the subjective gameplay experience, as suggested by Brockmyer et al. (2009), are immersion, presence, and flow (Jennett et al., 2008; Brockmyer et al., 2009; Boyle et al., 2012). Given that cognitive absorption is essentially flow, it will not be included in this discussion. Involvement has been mentioned several times in the literature, but has not received as much attention as the other constructs. Still, involvement as a construct is also worthy of exploration (Witmer & Singer, 1998; Wirth et al., 2007). The definitions of these four constructs, based on decades of research and as applied to the subjective gameplay experience, are as follows.

Immersion. The definition of immersion should include: (1) Slater and Wilbur's (1997) "system immersion" (Slater, 1999, p. 560), which are the objective characteristics of a VE that create an all-inclusive experience in which the distractions of the real-world are minimized and immersive cues are provided to the user's senses; and (2) Witmer and Singer's (1998) "immersive response" (Slater, 1999, p. 560), which represents the subjective response of the user within an interactive VE regarding a feeling of being enveloped by the stimulus flow, which is facilitated by those immersive cues. Immersion involves attention, sensory resolution, and usability (Witmer et al., 2005). Furthermore, when the technology is unable to provide adequate immersive cues, individual differences and cognitive factors can mitigate this deficit and increase the subjective experience of immersion to the point that presence is possible (Sas & O'Hare, 2004; Wirth et al., 2007). Immersion is defined with respect to games accordingly:

> *Immersion is "a [subjective] state characterized by perceiving oneself to be enveloped by, included in, and interacting with [a video game] that provides a continuous stream of stimuli and experiences" (Witmer & Singer, 1998, p. 227). Immersion requires focused attention on a limited stimulus field and minimized distractions (Csikszentmihalyi, 1990; Sheridan, 1992, 1994; Witmer & Singer, 1998), which can be promoted by the video game system itself. Immersion may be enhanced by the capability of the video game's technology to provide the player immersive cues. This includes the ability to interact with the video game through a virtual representation of the player (Slater & Wilbur, 1997). Interaction must seem natural with regard to the input mechanisms and the game's response to the player (Witmer & Singer, 1998; IJsselsteijn et al., 2000). Immersive cues are also strengthened by increasing the "extent and fidelity [and resolution] of sensory information" (Slater & Wilbur, 1997; IJsselsteijn et al., 2000; Wirth et al., 2007). Lacking immersive cues, involvement and individual differences may mitigate the deficit, thus helping the player to experience immersion (Witmer & Singer, 1998; Sas & O'Hare, 2004; Wirth et al., 2007).*

The subjective experience of immersion in video games is a function of both the technological capabilities of the game system and the individual's subjective perception of being 'enveloped' by the gameplay experience. This definition does include interactive elements, even though immersion itself is possible in non-interactive contexts such as books (Wirth et al., 2007). Since this definition applies to the subjective experience of video games, interactivity is assumed.

Immersion begins with focusing attention (Sheridan, 1992, 1994; Witmer & Singer, 1998), which can be a conscious effort on the part of the individual or facilitated by the game's technological capabilities. This is achieved when the stimulus field is limited and distractions are minimized, such as when a surrounding HMD is used so that the individual may only focus on the game and so that the distraction of physical reality is occluded from their senses. Distractions may also be related to the game itself. Witmer and Singer (1998) highlighted the importance of interface quality. A game that has glitches, or has game mechanics, controls, and interfaces that are difficult to master and understand will be unusable, distracting, and detract from the experience of immersion (Brown & Cairns, 2004; Jennett et al., 2008).

Similarly, the method of interacting with the environment through a representation of the player must feel natural and be free of latency (Witmer & Singer, 1998, p. 236). For example, the player should be in control of their avatar's senses. Therefore, the game's display should update to accurately reflect the magnitude of the player's input (e.g., realistic perception of movement; Slater et al., 1994). The method of physically playing the game should also be as natural as possible. A highly immersive game goes beyond usable controls, in which the player's physical inputs are accurately reflected by the avatar (Slater & Wilbur, 1997). For example, a game will be more natural and, thus, immersive if the player moving their physical head results in the avatar moving its head in the exact same way at the exact same speed. Including every one of these aspects is not necessary, nor do they realistically represent capabilities of all games.

However, including more of these naturalistic features increases the strength of the immersive cues provided by the game. The bare minimum seems to include the ability to control and interact with the environment, in a non-distracting, logical way.

Furthermore, in line with the idea of providing a naturalistic experience, immersion increases when more of the player's senses are controlled by the game, thus creating a cohesive experience. Therefore, also increasing the number and resolution (e.g., level of detail) of the senses provided to the player will increase the immersive quality of the game (Draper et al., 1998; Slater & Wilbur, 1997; IJsselsteijn et al., 2000; Wirth et al., 2007). Photorealism is not a requirement of immersion, as the world can be unnatural stylistically as long as all of the elements make sense and are coherent within their own context (Slater & Wilbur, 1997).

Deficits in the technological capabilities of the video game to provide immersive cues may be damaging to the individual's subjective appraisal as to whether a player feels immersed, but this can be mitigated by both involvement and individual differences (Wirth et al., 2007). For example, as involvement increases, so should immersion (Witmer & Singer, 1998). Lacking immersive qualities, pure motivation can allow an individual to experience immersion in the game, simply because they are striving to become immersed. Also, individual differences can also mitigate the immersive cue deficit, such as willingness to suspend disbelief (Sas & O'Hare, 2004; Wirth et al., 2007) and creative imagination (Sas & O'Hare, 2004). An individual's own propensity to experience immersion may also be a factor to consider (Witmer & Singer, 1998).

Involvement. Witmer and Singer (1998) proposed that involvement is a concept that is distinct from immersion, yet the two act reciprocally where one increases the strength of the other. Wirth et al. (2007) also highlighted the importance of involvement. Involvement has otherwise received little attention in the gaming literature, although it seems to encapsulate the cognitive engagement and motivational aspects of the subjective gameplay experience and should be included as an additional fourth factor to accompany immersion, presence, and flow.

Research in the persuasion and advertising domain has focused on involvement. A study by Petty and Cacioppo (1984) found that when personal relevance for a given topic was low, an essay was found to be more persuasive if it had more arguments, even if the strength of the argument overall was identical to an essay with fewer arguments listed. When participants were given essays that were personally relevant, they were more persuaded by the quality of the arguments rather than the quantity. Therefore, personal relevance was found to be a motivational factor that influenced persuasion, where quantity trumped quality when personal relevance was low.

The terms personal relevance and involvement were used interchangeably by Petty and Cacioppo (1984). Indeed, when Zaichkowsky (1985) strove to strictly define involvement, her approach centered on personal relevance. Involvement, as applied to advertisement and purchasing decisions, was explicitly defined as: "A person's perceived relevance of the object based on inherent needs, values, and interests" (Zaichkowsky, 1985, p. 342), which is also shaped by individual differences, characteristics of the target object, and situational factors (Zaichkowsky, 1985). Echoed here are elements of the TAM (Davis, 1989), in which perceived usefulness, assuming that usefulness is interchangeable with relevance, was found to be a downstream predictor of intention to use software, just as it is a factor in persuasion (Petty & Cacioppo, 1984) and purchasing behavior (Zaichkowsky, 1985). This definition also addressed both cognitive (e.g., objective information about utility) and affective (e.g., subjective emotional appraisal) components that were proposed to be a part of the involvement construct (Zaichkowsky, 1994). Wirth et al.'s (2007) operationalization of involvement also included cognitive and affective components.

Therefore, involvement is a motivational aspect that influences whether an individual will interact with an object driven by individual differences, such as personal relevance and interest, and is shaped by the characteristics of the object and situational factors. Involvement with respect to games may be defined as:

> *Involvement is a motivational factor (Wirth et al., 2007) regarding gameplay that is "experienced as a consequence of focusing one's energy and attention on a coherent set of stimuli or meaningfully related activities and events. Involvement depends on the degree of [perceived relevance] that the individual attaches to the stimuli, activities, or events" (Witmer & Singer, 1998, p. 227). Involvement is "increased by [playing video games] that stimulate, challenge, and engage the user either cognitively, physically, or emotionally" (Witmer et al., 2005, p. 299). Involvement has a reciprocal relationship with immersion, where increasing a sense of immersion similarly increases a sense of involvement, and vice-versa (Witmer & Singer, 1998).*

Involvement shares the same attentional requirements as immersion, in that involvement requires focused attention on a limited stimulus field and minimized distractions (Witmer & Singer, 1998; Wirth et al., 2007). Indeed, an individual's own goal-driven motivation to play the game may also make them more willing to focus on gameplay and ignore distractions (Wirth et al., 2007). Witmer and Singer (1998) originally stated that involvement depended on "the degree of significance or meaning" and individual placed on the VE (p. 227). Essentially, involvement is determined by the level of perceived relevance a player experiences. Ultimately, involvement is unique from immersion in that its focus is more cognitive in that it relies on goal-driven motivation (Witmer & Singer, 1998) as well as has an affective component related to perceived relevance and desire to play (Wirth et al., 2007).

Presence. The definition of presence in video games is adopted from Wirth et al.'s (2007) definition of spatial presence:

> *Presence is a state of "conviction of being located in [the game] environment" (Wirth et al., 2007, p. 495). It is "a binary experience, during which perceived self-location and…perceived action possibilities are connected to [the game environment], and mental capacities are bound by the [the game environment] instead of reality" (Wirth et al., 2007, p. 497).*

Presence occurs when the individual experiences an overwhelming amount of immersion—be it directly from the prevalence of immersive cues provided by the game, an increased level of involvement that reciprocally increased immersion, or due to the influences of individual differences such as creative imagination, willingness to suspend disbelief, and immersive tendencies—in which the individual accepts the game environment as their PERF and that this state is either present or absent (Wirth et al., 2007). The greater the sense of immersion, the easier it will be for presence to occur and the less likely it will be that distractions break the individual out of this fully-immersed state.

Flow. The definition of flow is adopted directly from Csikszentmihalyi (1975a, 1975b, 1990) and incorporates the nine elements of flow and enjoyment: (1) challenge/skill balance; (2) concentration; (3) clear goals; (4) immediate feedback; (5) merging of action and awareness; (6) control; (7) loss of self-consciousness; (8) time distortion; and (9) the autotelic experience.

> *Flow is the "optimal experience" (Csikszentmihalyi, 1990, p. 39), in which attention is focused on a limited stimulus field provided by the video game and where outside distractions have been minimized, resulting in a merging of action and awareness where "that person's attention is completely absorbed" (Csikszentmihalyi, 1990, p. 53) and playing the game becomes "seemingly effortless" (Csikszentmihalyi, 1990, p. 54). This occurs in an activity with clear goals, immediate feedback as one works toward those goals, and a progressive level of challenge which encourages an individual to increase their skills (challenge/skill balance) so that the activity is not boring or frustrating. This results in a state of concentration, in which only the activity matters, a loss of self-consciousness where an individual loses themselves in the activity "with no room for self-*

scrutiny" (Csikszentmihalyi, 1990, p. 63), and a sense of control in which an individual feels confident in their abilities and does not worry about failure. The individual experiences a sense of time distortion, in which time either slows down or speeds up. The end result is the euphoric, intrinsically-motivating autotelic experience which individuals seek to recreate "simply because the doing itself is the reward" (Csikszentmihalyi, 1990, p. 67).

Flow is affected by the design of the activity itself, the culture and circumstances in which the individual experiences it, and the propensity for the individual to experience flow. Flow shares many core aspects of immersion, yet its additional focus on cognitively-driving features, such as goals, creates an intrinsically-motivating experience that is distinct. These cognitively-driving features may tap into the elusive element of involvement (Witmer & Singer, 1998; Wirth et al., 2007). Flow is a state beyond involvement in which a balance between the challenge of the game and skill of the player is found and moves progressively forward. A state of flow is not presence as it does not explicitly involve or require the "conviction of being located in a mediated environment" (Wirth et al., 2007, p. 495). Flow may be essentially the same as a state of cognitive absorption, and both dispositional flow tendencies and propensity to absorption may influence whether flow is experienced.

A revised model of game engagement. The subjective gameplay experience is characterized by the experiential states of immersion, involvement, presence, and flow. Brockmyer et al. (2009) proposed that these states are progressive, as outlined in their game engagement model. Involvement is the experience of motivation, while immersion is the experience of envelopment in stimuli provided by the game. Presence, which is a sense of being physically located in the game, and flow, the optimal experience, are both states that grow out of immersion and involvement, respectively.

Facing the risk of adding fuel to the mounting fire of confusion surrounding the subjective gameplay experience, one is hesitant to apply a new label to this emerging construct. It may not be such a bad idea, however, to take a page out of Brockmyer et al.'s (2009) book and to consider the usage of the term *game engagement* as a "generic" description of the subjective gameplay experience (p. 624). This term lends itself well to this particular situation. Recall that Brockmyer et al. (2009) delineated states of low-level and high-level game engagement, where low-levels of game engagement pertained to immersion and presence, while high-levels of game engagement pertained to the mind-altering states of flow and cognitive absorption. Perhaps it would be pertinent to revise their model based on the definitions proposed above: low-level game engagement may instead refer to the reciprocal states of immersion and involvement, while the concept of high-level game engagement applies to those states that are not mind-altering, as originally suggested, but instead are uniquely specific and require additional effort to obtain, which would include both presence and flow.

Low levels of game engagement, which is the result of the reciprocal relationship between immersion and involvement, is necessary to experience high levels of game engagement as characterized by flow and presence, but presence is not necessary to experience flow just as flow is not necessary to experience presence. Refer back to Figure 4, which shows the many complex ways in which Slater and Wilbur's (1997) and Witmer and Singer's (1998) notions of immersion overlap. This review suggests that the question mark stemming from involvement is likely flow. The revised game engagement model (R-GEM) based on the original game engagement model as proposed by Brockmyer et al. (2009) and the conceptual overlap presented in Figure 4, is illustrated below in Figure 8. See Appendix A for a bulleted outline of the constructs and their influences.

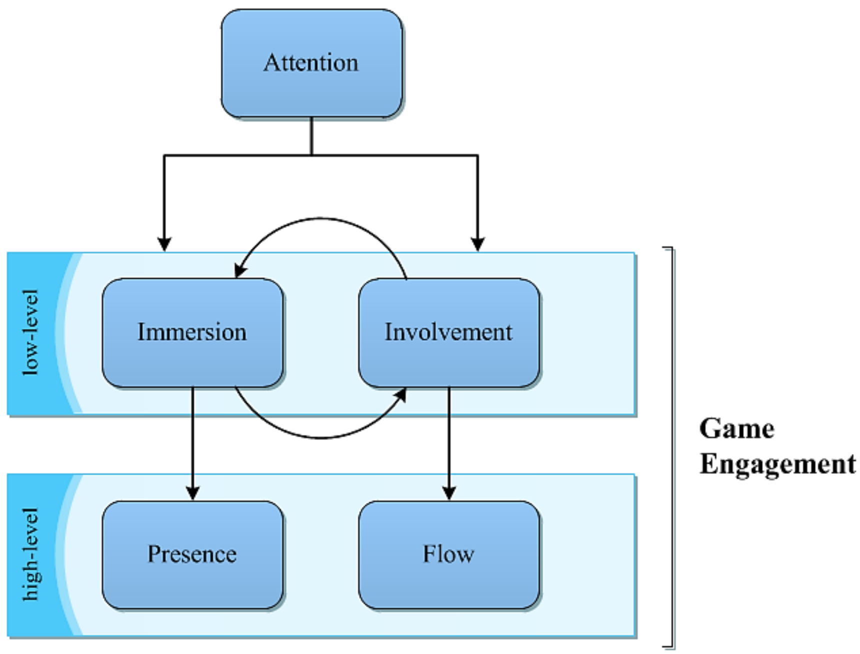


Figure 8. The revised game engagement model (R-GEM).

Low levels of game engagement are characterized by immersion and involvement. Immersion is the subjective feeling of being enveloped by the games' stimuli and experiences, while involvement pertains to motivation to play. Both immersion and involvement require focused attention on a limited stimulus field. Focused attention can be fostered by minimizing internal (e.g., ability to master controls and basics of gameplay, understandable interfaces, lack of glitches) and external (e.g., reduced background noises) distractions through high-quality, usable designs and unobtrusive game peripherals that limit external sensory interference (e.g., HMDs, headphones). Attention can also be focused due to sheer determination on the part of the player: if the player wants to play, they will be motivated to focus attention on their own.

An overall sense of immersion is supported by the technological capabilities of the system to provide immersive cues, and is enhanced by both the motivational aspects of involvement and select individual differences. Technological capabilities include realistic interactivity and input mechanisms, as well as higher sensory resolution and increased number of senses provided by the game. Refer to the work of Slater and Wilbur (1997) outlined in Table 1 for ways these aspects can be manipulated to increase the strength of the immersive cues. Involvement has a positively reciprocal relationship with immersion. Motivation to play the video game, shaped by a general desire to enjoy games as well as game-specific interest, which is essentially perceived relevance, drives the experience of involvement, which then also increases immersion. Relevant individual differences that increase the likelihood of low-level game engagement include willingness to suspend disbelief (Sas & O'Hare, 2004; Wirth et al., 2007), creative imagination (Sas & O'Hare, 2004), and immersive tendencies (Witmer & Singer, 1998).

Presence is a state that occurs when the player accepts the game environment as their PERF, meaning that they believe they are physically located within and interacting with the game. Presence occurs when there is a very high level of immersion, which, stated again, is influenced by the combination of immersive cues provided by the video game (e.g., high-fidelity, increased field of view, HMD, naturalistic inputs, etc.) and individual differences. The stronger the immersive cues, the easier it is to reach the state of presence, although deficits can be mitigated by involvement, which will increase an overall sense of immersion, and individual differences. This is one of the potential experiential outcomes of high-level game engagement.

Flow is the other potential experiential outcome of high-level game engagement. Flow is the optimal experience of intrinsically-motivated enjoyment while playing a game. It is a goal-driven high in which the player's perception of time becomes distorted. While it has nine very specific elements that have been articulated by Csikszentmihalyi (1990), there is a large amount of apparent overlap with immersion. The experience of immersion mirrors that of flow's elements of merging of action and awareness, concentration, loss of self-consciousness, and control (to a certain extent). There are also unique requirements for achieving the flow state, which includes goals, immediate feedback, and challenge/skill balance. It may be possible that the goals and immediate feedback elements are part of the involvement aspect of low-level game engagement; however, flow transcends this state as the progressive challenge/skill balance pushes the experience to become the high-level game engagement state of flow. Not surprisingly, the flow experience has the exact same requirements as immersion and involvement: a need for focused attention on a limited stimulus field with minimized distractions. Flow also features unique experiences, in that it is intrinsically-rewarding (the autotelic experience) and players in flow will experience time distortion.

Individual differences also affect the flow experience, such as the dispositional tendency to experience flow (Jackson & Eklund, 2004). Trait absorption may also be related to flow given that it was hypothesized to influence cognitive absorption (Agarwal & Karahanna, 2000), and the near-perfect overlap of cognitive absorption and flow. Finally, factors such as age and VGSE have been shown to have an impact on the flow experience (Pavlas, 2010; Procci et al., 2013).

As described above, the requirements of presence and immersion are essentially the same. Indeed, presence is an extremely immersed state in which the immersion is so powerful that the individual feels as if they are physically in the game environment, as if they are breaking into a new level of immersion that is its pinnacle. Flow shares many common aspects with immersion, to the point that it is also a potential precursor of flow, but is unique given its cognitive-focused features, which stem from involvement. The high-resolution, naturalistic elements that result in presence are not necessary for flow to occur. Therefore, eight assertions about the revised game engagement model can be made, and are listed below in Table 2.

Table 2. Eight assertions of the revised game engagement model.

1	Immersion pertains to being enveloped in the stimulus flow of the video game, supported by the presence of immersive cues.
2	Involvement pertains to goal-driven motivation to play the video game.
3	Immersion and involvement share several common factors: focused attention and limited internal (usability) and external (limited stimulus field) distractions.
4	A sense of immersion can increase the experience of involvement, and vice-versa. Low-level game engagement is defined by the state of this reciprocal relationship.
5	Presence and flow are two different products of high-level game engagement, both of which require low-level game engagement, however immersion is more influential in presence while involvement is more influential in flow.
6	Presence occurs when there is an overwhelming amount of immersion, which can continue to be high when immersive cues are lacking due to the mitigating influence of involvement and/or individual differences.
7	Flow occurs when a balance is achieved between the mounting challenge of the game and the growing skill of the player.
8	It is possible to be both present and in flow, as well as only experiencing presence or only experiencing flow.

Furthermore, while this revised model of game engagement has been simplified, the literature review has made it apparent that there are many, many potential predictors to be considered. Heeding Bagozzi's (2007) warning about the utility and testability of such a model, the approach taken here should be to crudely validate this model by manipulating the different mechanisms of the game with guidance from the literature and drawing hypotheses as to how these manipulations are expected to affect the experiential states of immersion, involvement, presence, and flow based on the R-GEM illustrated in Figure 8. Testing the above listed assertions (Table 2) will determine whether the R-GEM is valid, thus reducing some of the construct confusion currently present in this field of research and paving the way for future empirical efforts as more granular aspects of the model are refined.

Construct-specific measurement issues

Given the complexity and confusion surrounding these different constructs, it is important to briefly review measurement issues connected to each. In the early 2000s, as definitions of immersion and presence became somewhat more concrete, the next issue to address was that of valid, reliable measurement. IJsselsteijn et al. (2000) summarized the ways in which presence had been measured, which included both subjective and objective assessments. While this originally applied only to presence, the notion can be extended to all of the subjective states discussed in this section. The subjective assessments included post-test self-reported rating scales, continuous presence assessment, and psychophysical ratings. Regarding post-test self-reported rating scales, IJsselsteijn et al. (2000) provided the PQ (Witmer & Singer, 1998) and questions pertaining to presence from Slater et al. (1994) as key examples. Additionally, as the experience of presence is likely to change throughout the task, the authors believed that a single post-task assessment may not be accurate. IJsselsteijn et al. (2000) suggested overcoming this problem through the use of continuous presence assessment methods, in which an individual is asked to report on their current level of presence by using a sliding bar. While Slater and Steed (2000) argued that a method similar to this is disruptive to the experience of presence itself, IJsselsteijn et al. (2000) defended it, stating "observers are unlikely to report on a belief of being in the displayed scene, since they are usually well aware of actually being in a media laboratory. Rather, they report on a sensation of being there that approximates what they would experience if they were really there. This does not necessarily conflict with providing a continuous rating - especially given the fact that the measurement device requires very little attention or effort to operate" (p. 523). Still, it does seem that this method would be somewhat intrusive to the on-going experience of play. Another type of subjective measure of presence is psychophysical measurement, such as when participants are provided pictures and asked to report if it is a VE or the real world (e.g., Schloerb,1995, as cited in IJsselsteijn et al., 2000). IJsselsteijn et al. (2000) criticized this method as the results are likely to reflect the participants' ability to judge pictures rather than measure their experience of presence in VEs.

IJsselsteijn et al. (2000) stated that subjective measures would fail if users did not understand the definitions of the constructs of interest or if they did not fully comprehend what they were being asked. As such, they suggested that objective corroborative measures also were needed. This aligns with Slater and Steed's (2000) desire to incorporate objective behavioral reactions into their measurement of presence, however this remains difficult as the designer of the VE will often need to include an element that causes a startle reaction. This is problematic as this additional element may not serve any purpose other than to assess presence. It will also limit the use of off-the-shelf VEs, such as games, in this type of research as they may not have a feature that supports this method of evaluation. This is precisely why Slater and Steed (2000) opted to investigate hand movements rather than reflexive behaviors when quantifying behavioral responses to presence, as nothing needed to be added to the VE. Others suggested that task performance is a viable way to measure presence, however, some contend that simply experiencing presence does not automatically translate into performance gains (see Slater, 1999). Others (Freeman et al., 1999, as cited in IJsselsteijn et al., 2000) found that postural response was associated with presence, which is not an adequate standalone assessment, but could be an interesting corroborative measure worth investigating.

IJsselsteijn et al. (2000) also highlighted that the use of physiological measures to quantify presence was rising in popularity, such as heart rate and galvanic skin response (GSR). IJsselsteijn et al. (2000) stated that GSR is associated with arousal, while heart rate is linked to "hedonic valance" (p. 525), in which pleasant stimuli causes increases in heart rate while unpleasant stimuli results in the heart rate slowing. GSR does not seem to be a valid measurement option due to mixed findings in the literature. For example, IJsselsteijn et al. (2000) reported on a study by Wiederhold, Davis, and Wiederhold (1998) in which participants experienced a VE with or without an HMD. GSR increased across both conditions, however there was no statistical difference between the two, which is not what one would expect given the presence of the HMD. Sallnas (1999, as cited in IJsselsteijn et al., 2000) was interested in testing the effect of haptic feedback on presence and co-presence utilizing both subjective measures and GSR. The data suggested that the addition of haptics increased task performance and subjective measures of presence, but had no effect on GSR. Later work would find that heart rate may be more useful than GSR with respect to measuring immersion and presence (e.g., Drachen, Yannakakis, Nacke, & Pedersen, 2010).

IJsselsteijn et al. (2000) suggested using dual-task measures as a final method of providing corroborate objective assessments of presence, although they did not report on any research that had been conducted to this effect. Given attention's role as an important downstream component of low-level game engagement, a dual-task paradigm could be an effective way to measure presence. In this paradigm, as more attention is paid to the primary task, which would be whatever the user is supposed to be doing in the VE, performance on a secondary task should decrease as presence increases. Unless the dual-task measure can be built into the game—which is possible, see Sharek and Wiebe (2012)—it may be too distracting and detract from the overall experience. Therefore, dual-task paradigms are of limited value in this type of research if embedding the task to be seamlessly integrated with gameplay is infeasible.

While IJsselsteijn et al.'s (2000) review was specific to presence, the overall message applies to all of the constructs of interest: Subjective assessment of game engagement may be fallible, therefore secondary objective measures are warranted, such as whether the participant exhibited any startle responses or if their body movements changed. The use of a single post-task measure may be questionable as the experience likely changes throughout the task, therefore physiological measurements of these constructs should also be established.

Given the exploratory nature of this work, only subjective measures will be utilized. If these methods hold promise, future work should then examine ways in which these ratings and theoretically-driven stimuli affect physiological measurement. The ways the major constructs of game enjoyment (i.e., immersion, involvement, presence, and flow) have been measured will now be discussed.

Immersion. The measurement of immersion has been covered extensively in a previous section. Immersion should be measured as the subjective response to system characteristics that envelop a user within the stimulus flow. It is possible to measure immersion with the PQ (Witmer & Singer, 1998). Although the PQ is purported to measure presence, it purposefully only measures the subjective response to factors that promote presence (Witmer et al., 2005). Therefore, the PQ essentially measures immersion, as well as involvement and the attentional requirements of low-level game engagement.

Jennett et al. (2008) also developed a questionnaire to measure game immersion, and while their definition of immersion does not align well with the one established above, it is worth discussing briefly. Their measure of game immersion utilized Agarwal and Karahanna's (2000) five factors of cognitive absorption, as well as Brown and Cairns' (2004) proposed elements of game immersion. The immersion questionnaire was composed of sixteen pairs of positively- and negatively-worded questions, for a total of 32 items, each of which was rated on a 5-point Likert scale. They also wanted to address a comment from Slater (1999), in which he critiqued Witmer and Singer (1998) for not actually having any items tied to the actual construct of presence itself. They added an additional question in which respondents were asked to rate how immersed they felt on a scale of 1 to 10. Jennett et al. (2008) also were interested in objective measures of immersion, such as task performance and eye tracking. They conducted three experiments to this effect.

The purpose of the first experiment was to determine if the immersion questionnaire could discriminate between non-immersive and immersive computer-based tasks. The non-immersive task involved clicking boxes as soon as they appeared on a computer monitor, while the immersive task was playing the game *Half-Life*. Participants also completed tangrams before and after the experimental intervention to determine if there was an effect on task performance. Jennett et al.'s (2008) results indicated that the immersive task was rated more highly on the immersion questionnaire than the non-immersive task. The analysis of the task performance data did not yield any meaningful results.

Deciding to move away from task performance, the purpose of the second experiment was to determine if eye tracking would be a more useful—and less invasive—complimentary metric for quantifying immersion in comparison to the tangram task. Jennett et al. (2008) hypothesized that immersion would be associated with eye movement changes, specifically the number of fixations per second, although they did not hypothesize whether that number would increase or decrease. Once again, participants either completed the non-immersive clicking task or played *Half-Life* for 10 minutes. Eye tracking was used to count the number of fixations per second as well as fixation duration. Afterward, the participants completed the immersion questionnaire. Jennett et al. (2008) found that fixations decreased over time in the immersive task, while they increased in the non-immersive task. The analysis regarding fixations and score on the immersion questionnaire revealed extremely conflicting findings, where many in the non-immersive task rated it highly immersive and vice-versa.

Based on the results of these two studies, Jennett et al. (2008) found it necessary to refine their questionnaire. They developed an entirely new set of questions with similar wording, but removed the positive-negative pairs as this confused participants. The new questionnaire featured six sections with 31 items, where each item was once again rated on a 5-point Likert scale. Most of the items were positively worded, but some were reverse-coded. A total score was found by summing all 31 items. The first three sections were based on flow, cognitive absorption, and elements of Brown and Cairns' (2004) theory, which included the subscales focused on attention, temporal dissociation, and presence. Believing that task characteristics would also likely influence the experience, the fourth section of the questionnaire was based on challenge. The last two sections pertained to emotional involvement and enjoyment.

The newly-revised immersion questionnaire was administered to a sample of 244 participants recruited from online gaming forums and a factor analysis was conducted. Respondents were asked to think about the last time that they played a game, and then completed the immersion questionnaire. Principal component analysis revealed a single overall factor with five smaller factors that they determined were player characteristics , such as “cognitive involvement, real world dissociation, emotional involvement” and game characteristics, such as “challenge [and] control” (Jennett et al., 2008, p. 654). The authors did not expand on the precise meanings of these factors.

A third experiment was then conducted. The purpose of this experiment was to help clarify the confusing findings from the second experiment, in which some participants rated the box-clicking task as highly immersive. Jennett et al. (2008) believed that this finding may have had to do with the pace of the interaction, as some participants reported liking the task because they wanted to see how quickly they could click the boxes, thus making a game out of a supposedly boring, non-immersive task. They believed that if the new boxes did not immediately appear after the click, this would slow down the pace of the task, and result in less immersion. They had four conditions in their study, which only involved the box clicking task, where participants were instructed to click boxes as quickly as possible. The conditions varied how quickly the boxes appeared on the screen. After the task, participants completed the immersion questionnaire, as well a measure of anxiety and another of affect. Jennett et al. (2008) found that the pace of the task did not affect immersion, likely because the task just was not game-like enough. Ratings of anxiety and negative affect did increase as the pace increased.

Overall, the results of Jennett et al. (2008) were not illuminating regarding their concept of game immersion, nor validating for their measure. Ultimately, the problem with Jennett et al.’s (2008) work is that it combined many different theoretical approaches and does not distinguish well between them, having subsumed everything under the description of ‘immersion’ when it truly was not.

Approaching immersion in games from a different angle, Qin, Rau, and Salvendy (2009) were interested in measuring immersion in game narrative. Game narratives are distinct from other narrative forms because they may be non-linear, are interactive, and are usually advanced through the mastery of game mechanics (Qin et al., 2009). While not all well-liked, time-absorbing games feature a narrative (e.g., *Tetris*), many do. Narrative had been largely ignored in the immersion and presence literature, aside from brief consideration by Slater and Wilbur (1997), who believed that a VE's plot may increase immersion (which may actually have been a reference to involvement).

Qin et al. (2009) sought to define, measure, and validate factors of immersion in game narratives. They defined immersion in a game narrative as a descriptive term in which players felt "totally submerged in their fictional surroundings" (Qin et al., 2009, p. 113). As the result of two studies, they stated that narrative is driven by goals, which are sometimes provided by the player themselves rather than the game, as well as curiosity to see where the game takes them next. The player must also comprehend the storyline. This results in concentration, a sense of control over gameplay, and an understanding of the key parts of the storyline. This culminates in a sense of empathy with the game's characters, which aligns very well with the emotional involvement aspect of immersion proposed by Jennett et al. (2008). The concept of interest here is yet another amalgam of theories, and while it features an approach very similar to the game engagement model proposed in this work, it is not an examination of the construct of immersion, as defined in the previous section. That is not to say that Qin et al.'s (2009) work is not informative, but it perhaps should not be labeled as 'immersion'.

Involvement. Involvement is a motivational construct. It is influenced by individual differences, characteristics of the video game, and situational factors (Zaichkowsky, 1985). This may include; the perceived relevance of the game as it pertains to enjoyment and need fulfillment (Davis, 1989); motivation through the presence of goals, either imposed by the game itself or the individual onto the game (Csikszentmihalyi, 1990, also similar to Qin et al., 2009); game-generic and game-specific preferences (Procci et al., 2011); and whether the individual feels as if playing a game in a given setting would be an enjoyable and worthwhile experience. Involvement should be measured by addressing all of the above listed aspects of these three factors. Two potential validated measures for measuring involvement exist: The Revised Personal Involvement Inventory (Zaichkowsky, 1994) and items from the expanded perceived usefulness subscale of the TAM (Davis, 1989).

Personal Involvement Inventory**.** Judith Lynne Zaichkowsky developed (1985) and later refined (1994) the Personal Involvement Inventory (PII). The PII was developed for use in advertising research and was composed of 20 semantic differential scale items (Zaichkowsky, 1985). The initial validation effort involved eight different datasets, in which content validity, reliability, criterion-related validity, and construct validity for the PII was established. Zaichkowsky (1994) addressed concerns of scale length and validity in a follow-up validation study.

Zaichkowsky (1994) determined via calculation that the scale could be reduced to 10 items from 20 while still achieving a Cronbach's α of .90, which is an acceptable value in applied domains (Nunnally, 1978, as cited in Zaichkowsky, 1994). She first established content validity by having expert judges rate the set of word pairs identified in the initial study (Zaichkowsky, 1985). The judges identified 35 word pairs related to involvement. These 35 word pairs were rated by a sample of 54 undergraduate students with respect to different consumer goods. Items with low average correlations across all of the goods were dropped, as were items that were redundant based on high inter-item correlations. Twenty-two items with a Cronbach's α over .90 were retained.

Zaichkowsky (1994) then examined the test-retest reliability of those 22 items in a new sample of 52 undergraduate students, who rated a set of advertisements twice, three weeks apart. Nine items with test-retest correlations below .60 were dropped, as were three items with high inter-item correlations. This resulted in a final set of 10 items, which would become known as the Revised Personal Involvement Inventory (R PII), where Cronbach's α ranged from .91 to .95 and the test-retest correlations ranged between .73 and .84. Participants also were asked to describe their reasoning behind the ratings they gave the advertisements. To establish content validity, these qualitative responses were scored by expert judges—who were not exposed to either the advertisements or the participants' ratings—as having either low, medium, or high levels of involvement. There was a significant relationship between the experts' judgments and the ratings provided by the participants, whose scores were divided into the low (10-29), medium (30-50) and high (51-70) categories.

The R PII was also subjected to factor analysis (Zaichkowsky, 1994). Initially, a single factor solution was found. To address previous concerns that involvement has both affective and cognitive components, the scale was divided into those two subscales. Whether these items align with Jennett et al.'s (2008) factors of emotional involvement and cognitive involvement is unknown. The subscales were found to be internally consistent (Cronbach's α = .86 to .95), and significantly and positively correlated (r = .58 to .70). A follow-up factor analysis found that the model of the two correlated subscales had a better fit than the initial single factor solution (Zaichkowsky, 1994). Utilizing the R PII will likely be very useful for measuring involvement in gameplay, although it requires some edits to the instructions to make it more applicable to this domain.

Technology Acceptance Model (TAM). The perceived relevance of an object is a major determinant of involvement (Zaichkowsky, 1985). Perceived relevance in the video game context is the amount to which the individual's needs (e.g., for enjoyment, a method of relaxation, as an escape) are met by the game. This is essentially how Witmer and Singer (1998) described the concept of involvement. The construct of perceived relevance, specifically termed "perceived usefulness" is also central to the TAM (Davis, 1989, p. 320). In the TAM, perceived usefulness is defined as "the degree to which a person believes that using a particular system would enhance his or her job performance" (Davis, 1989, p. 320). Given that this is identical to the concept of perceived relevance presented here, albeit in a different domain, the term perceived relevance will be used.

The original TAM includes two subscales: one for perceived ease of use and one for perceived usefulness/relevance. First, Davis (1989) created 14 items for each aspect of the model. This number was based on the Spearman-Brown Prophecy formula, which suggested that 10 items would be adequate to reach a reliability of .80. Four items were added to that number so that the best items could be identified. Early validation work involved participants who rank-ordered items to ensure that the candidate items accurately represented the constructs of interest. This reduced the number of items to 10 per construct. The initial set of 20 items were then used in a field study of 112 computer users. The results of that study lead to the elimination of four additional items per construct.

The final set of 12 items (6 per construct) was evaluated in a second study that compared two graphics programs. The items for perceived usefulness/relevance across the two studies were found to be internally consistent (Cronbach's α = .97 and .98). Several validation studies have used these items with similar results (e.g., Davis, Bagozzi, & Warshaw, 1989; Venkatesh, 2000; Venkatesh & Davis, 2000; Yi & Hwang, 2003; Venkatesh et al., 2003; Venkatesh & Bala, 2008; Evans, 2010).

Examination of the perceived usefulness items from Davis (1989) reveal that it may be difficult to apply them to gaming without significant rewording. Instead, these items serve as an inspiration for additional items that will complement the R PII by ensuring all relevant areas of the construct are measured.

Presence. Dissatisfied with the use of subjective measures to quantify presence (e.g., the PQ), Slater and Steed (2000) sought to validate an equation to predict presence as well as develop a new method of quantifying presence through behavioral observation, as was suggested by Slater and Wilbur (1997). Slater and Steed (2000) did not address immersion in this work other than to reiterate that their approach defines immersion as a description of the technology and not a subjective state.

Slater and Steed (2000) summarized the many different ways that presence can be measured. For example, including some sort of startle event, as suggested by Held and Durlach (1992, as cited in Slater & Steed, 2000), and measuring the response (did the participant jump in their seat?) is one method, while others measured the after-effects of VE use (Welch, 1997, as cited in Slater & Steed, 2000). They also mentioned that several different subjective scales have attempted to quantify presence, but that they continue to ultimately disagree with their implementation. In addition to definitional disagreement, they also stated that these scales rely on summating Likert scale scores, and that they staunchly believed that ordinal data cannot be summated or averaged to find a total overall score. Citing their own previous work (Slater, Steed, & Usoh, 1993, as cited in Slater & Steed, 2000), they suggested instead setting a rating threshold for each item and summating for how many items that threshold was breached.

Citing Bystrom, Barfield, and Hendrix (1999), Slater and Steed (2000) stated that attention is a key component in that where and what the individual is attending to is an important determinant of presence. These transitions of feeling as if one is present in the VE (V) or in the real environment (R) can occur when the user is reminded that they are using a VE, such as a glitch causing the frame rate to drop or an outside distraction such as a phone ringing. They believed that quantifying this transition is the key to measuring presence. They labeled this transition as a break in presence (BIP), and whether a BIP occurred (and how often) was obtained through means of self-report. Slater and Steed (2000) noted that while it is impossible to have a user declare that they have transitioned from R to V, as this would immediately bring them back to R, it should not be disruptive for the user to signal when they transition from V to R, as they have already been distracted and removed from V for whatever reason.

Slater and Steed (2000) created an equation to predict the likelihood of experiencing presence in a VE (*p*) based on the predicted state (high vs. low presence) and the number of BIPs. They attempted to validate this equation with an experiment, as well as sought to determine if the *p* value was positively correlated with both subjective measures of presence as well as behavioral observations, which would address some concerns outlined in Slater (1999).

In their experiment, participants played a three-dimensional chess game in a VE that utilized an HMD with head tracking and a three-dimensional mouse. The VE was constrained to have a frame rate that did not dip below 20, a latency of 120ms, and an expansive field of view of 67°, as well as a first-person perspective and motion tracking to promote realistic matching between the user and the avatar, all of which should have been suitable to promote a sense of presence as prescribed by Slater and Wilbur (1997).

Given the nature of the task, the individual would have to stretch their virtual body to reach all of the chess pieces on the board. Slater and Steed (2000) used this as an experimental manipulation, and participants were divided into conditions where they would either complete task using motion tracking (high-activity group, n = 10), so that they reached out their physical hand to move the virtual one, or use the three-dimensional mouse (low-activity group, n = 10).

After completing a brief demographic survey, the participants were trained to state whenever they experienced a BIP by viewing gestalt illusions (e.g., is it two faces or a vase?) and saying the word "now" whenever they perceived the figure to switch with the ground. While participants completed the chess-playing task, reported BIPs were counted. Afterward, participants responded to 5 items by rating them on a 7-point Likert scale, where higher scores indicated more presence experienced. This was based on previous experiments in which they found success with this method (see Usoh et al., 1999). The first question was used as a discriminator item, which determined whether the high presence (scores 6 and above) or low presence (scores 5 and below) equation was used to calculate p. Depending on the equation selected, the counted BIPs were used to generate p. For the remaining four items, rather than summing the Likert-scale data, they counted the number of times participants reported a score of

6 or above. This resulted in a total number that is the individual's "presence count" (Slater & Steed, 2000, p. 427).

Slater and Steed (2000) did not find that *p* differed between conditions; however, they did find a significant, positive correlation between *p* and the presence count. The researchers also were interested in whether physical activity was associated with presence in an attempt to validate a method of quantifying presence using objective behavioral responses. They did not find that there was a significant relationship between hand movement and *p* in the low-activity group, which makes sense given that these individuals did not have to move their hands to complete the task. They did find, however, a significantly positive relationship in the high-activity group between hand movement and *p*, $r^2 = 0.73$, $t(15) = 5.69$, $p < .05$, where $p = -1.90 +$ 19.11[hand movement] (Slater & Steed, 2000, p. 428). The researchers suggested that this may be a feedback loop, in that those who experienced more presence moved their hands more, and seeing their avatar also move their hands in union with their own movements would increase presence, thus resulting in even more hand movement.

These results relied on the strength of the equation to calculate *p*. While the researchers found that the use of the discriminatory question was valid, they did encounter a problem with BIP reporting. Not all of the participants were trained using the gestalt illusions, and half of those participants who were not trained with the illusions did not report a BIP, while all of the participants that were trained did. Slater and Steed (2000) took this to mean that the number of reported BIPs may have been artificially high. This is troublesome, so while these results are interesting, this indicates that special attention needs to be made to refining the BIP training procedure so that more accurate results are found. Given that the items used to generate the presence count were strongly correlated with *p*, this indicated that these subjective items are quantifying some aspects of presence. Slater and Steed (2000) also suggested that future work should have participants estimate the percentage of time they felt present in the VE and using the 50% as a critical point.

Witmer and Singer's (1998) PQ did not measure presence directly, and it is arguable that the PQ actually measures immersion. As a result of a follow-up factor analysis, Witmer et al. (2005) did provide four additional items that were more directly related to presence, stating that the efficacy of these new items could be explored in future research efforts.

Flow. Csikszentmihalyi (1990) measured the flow experience using the Experience Sampling Method, in which participants were asked to report on their experiences and feelings at several points during the day over the course of about a week. This allowed him to study in-depth the flow construct and the role it played in everyday lives. Ever since, many others have devised new ways to measure flow.

Delle Fave, Massimini, and Bassi (2011) provided an excellent summary of the different ways flow had been measured effectively. Some methods involved interview and direct observation, which is exactly how Csikszentmihalyi (1975a, 1975b, 1990) chose to conduct his research. Delle Fave et al. (2011) stated that this method is excellent for exploratory research. There are also many self-reported measures of flow, either to assess flow in general, or in response to a specific event (Delle Fave et al., 2011). For example, the Flow Questionnaire was developed by Csikszentmihalyi (1975a) and later expanded by Dell Fave and Massimini (1988, as cited in Delle Fave et al., 2011). It featured both scale-ratings and open-ended questions divided into three major sections. Respondents were asked to read passages about flow and then rate and report on scenarios in which they may have experienced flow themselves. It broadly examined flow as a component of everyday life.

One of the more prominent of the situation-specific flow questionnaires is the Flow State Scale (FSS) developed by Jackson and Marsh (1996). The FSS measured the flow experience in physical activities and featured 36 items rated on a five-point Likert scale. Initially, researchers developed a large item pool that covered aspects of each of the nine specific elements of flow proposed by Csikszentmihalyi (1990). Flow experts identified 54 potential items from this pool. These items were administered to 252 individuals who regularly engaged in physical activities. The results of this analysis resulted in item revision and clarification. The revised items were then administered to 394 more participants, and a factor analysis examined model fit for both the 54-item model as well as a 36-item model, which featured four items for each of the nine elements. The analysis revealed that the 36-item model was the stronger of the two, with factor loadings that ranged from .56 to .88 (Jackson et al., 2010a, p. 21). The analysis also revealed that the nine first-order factors were all significantly and positively correlated (r = .18 to .72; Jackson et al., 2010a, p. 21). These nine first-order factors also all loaded onto one global flow factor, although the variability of these factor loadings was very high and ranged from .39 to .91 (Jackson et al., 2010a, p. 21). The researchers also developed the Dispositional Flow Scale (DFS) that measured propensity to experience flow in a given activity (Jackson, Kimiecik, Ford, & Marsh, 1998). Similar to the FSS, it also featured 36-items divided into 9 subscales, each of which was rated on a five-point Likert scale. These nine subscales were referred to as 'dimensions' in both the FSS and DFS.

Jackson and Eklund (2002), in an effort to reduce the variability found in the previous validation effort (Jackson & Marsh, 1996), produced the Flow State Scale – 2 (FSS-2) and the aforementioned Dispositional Flow Scale – 2 (DFS-2). With the help of Csikszentmihalyi, additional items were developed to produce an item pool that consisted of the original 36 items from the FSS/DFS, as well as 13 new items. This item set was administered to 597 individuals who participated in physical activities, some of which either completed the DFS-2 or the FFS-2 items, although some completed both. Those who completed the FSS-2 items were asked to engage in an activity and then complete the survey. The reported average activity length was roughly 25 minutes. Five of the original items were replaced with new ones, and the scale remained 36 items long.

Confirmatory factor analysis (CFA) was conducted by Jackson and Eklund (2002) and the new versions of the scales were found to have acceptable model fit. Factor loadings for the nine first-order factors averaged .78 and .77 for the FSS-2 and DFS-2, respectively (Jackson et al., 2010a, p. 24). The correlations between these nine first-order factors ranged between $r = .13$ to .76 for the FSS-2 and between $r = .24$ to .78 for the DFS-2 (Jackson et al., 2010a, p. 24). The nine first-order factors also loaded onto the global flow factor for an average of $r = .66$ for the FSS-2 and $r = .71$ for the DFS-2 (Jackson et al., 2010a, p. 24). Cronbach's α ranged from .80 to .90 for the FSS-2 and .81 to .90 for the DFS-2 (Jackson et al., 2010a, p. 31).

A follow-up validation study reported in Jackson and Eklund (2002) included 987 participants who engaged in physical activities and then completed either the DFS-2, the FSS-2, or both. For those that were asked to complete the FSS-2 after their activity of choice, the average length of time was once again about 25 minutes. Acceptable model fit was found. Cronbach's α ranged from .80 to .92 for the FSS-2 and .78 to .86 for the DFS-2 (Jackson et al., 2010a, p. 31). Therefore, the FSS-2 and DFS-2 were found to be reliable and valid measures of the flow experience in physical activity.

Jackson et al. (2008) produced short versions of both the FSS-2 and the DFS-2. The short flow scales contained 9 items—one for each of the major flow dimensions. While the long versions of the FSS-2 and DFS-2 are capable of producing a single global flow score, it is preferable to use the scale to examine the different dimensions. The short versions, however, are more suited to quantify the global flow score, especially in situations where flow is perhaps not the focal construct of interest or when there are time constraints. To create these short scales, existing items from each factor were selected based on the strength of their factor loadings. If the factor loadings were similarly high, the item with greater face validity was chosen for inclusion (Jackson et al., 2008).

Jackson et al. (2008) conducted a large-scale validation study in a sample of 1,552 participants that included both the short and long versions of the FSS-2 and DFS-2. Participants either reflected upon their physical activity of choice, or engaged in it, and then responded to the surveys. Participants were divided into four groups: 652 completed the long DFS-2; 499 completed the long FSS-2; 692 completed the shortened DFS-2 (S DFS); and 865 completed the shortened FSS-2 (S FSS). The activities ranged from yoga to hockey, and over half of the participants engaged in their given activity at least twice a week.

Jackson et al.'s (2008) method of construct validation incorporated both a within-network (internal structure) and a between-network (relationship analysis) approach. The within-network analysis involved a CFA, while the between-network analysis compared flow scores to theoretically-related constructs, specifically intrinsic motivation, self-concept, and psychological well-being. The measures were found to be acceptably reliable: Cronbach's α = .80 to .89 for the DFS-2; Cronbach's α = .76 to .90 for the FSS-2; average Cronbach's α = .77 for the S DFS; average Cronbach's α = .76 for the S FSS (Jackson et al., 2008, p. 568).

The short versions of the measures were modeled as both nine independent factors as well as nine factors embedded in the larger 36-item structure. A series of CFAs indicated satisfactory fit for the DFS-2, FSS-2, and S DFS. The CFA conducted on the S FSS data indicated poor fit. Choosing to conduct the CFA in only a those participants who engaged in one type of activity—volleyball, as it had the most individuals in its subgroup—Jackson et al. (2008) found acceptable fit. They believed that the S DFS is more robust to measuring global aspects of flow than the S FSS, and that situation-specific elements affect its efficacy.

Jackson et al.'s (2008) correlation analysis revealed that the short scales were representative of their longer counterparts (r = .73 and .97 for the DFS; r = .76 and .89 for the FSS; p. 537). The predicted between-network correlates also were found to be significant in the hypothesized ways. Finally, factor loadings were acceptably consistent across samples. The results of this large scale validation effort suggested that the S FSS and S DFS were representative of the original measures, although they are less diagnostic, with the S FSS being less valid than the S DFS.

While extensive validation testing has found that the DFS-2 and FSS-2 (Jackson & Eklund, 2002) are exceptionally strong measures of the flow experience (Jackson et al., 2008), these measures were developed with physical activity in mind. Still, Jackson, Eklund, and Martin (2010a) reported that the FSS-2 and DFS-2 have been used successfully in a variety of domains. For example, a CFA conducted by Wang, Liu, and Khoo (2009) found that the DFS-2 was a valid measurement of propensity to experience flow in online gaming. Still, in a study of 762 undergraduate gamers that responded to the DFS-2, their responses did not align with the factors proposed by the scale developers (Procci et al., 2012). This indicated that future psychometric work may be needed to refine the DFS-2, and likely the FSS-2, for gaming. As reported in a previous section, Hamari and Koivisto (2014) examined the factor structure of the DFS-2 in a gamification context. Again, gamification refers to the application of game features to traditionally non-gaming tasks, such as exercise. The authors and found the fit to the original model to be adequate, but not excellent.

Other measures specific to flow in games have been developed. Fu, Su, and Yu (2009) developed EGameFlow, which is for measuring the flow experience in serious games. The scale's dimensions strayed somewhat from the nine elements of flow and included: concentration, goals, feedback, challenge, autonomy (which was essentially control), immersion (which included elements of this work's definitions of both immersion and involvement), capacity for social interaction, and knowledge improvement. Delle Fave et al. (2011) believed that the most accurate scales are those that measure Csikszentmihalyi's (1990) nine elements of flow, such as the FSS-2 and the DFS-2, but this was not the case with EGameFlow.

EGameFlow was validated in a sample of 166 students taking an online software course (Fu et al., 2009). Students played Flash-based minigames which introduced players to operating system concepts. Item analysis was conducted, which reduced the scale from 56 to 48 items. The scale was found to be highly reliable (Cronbach's α = .94; Fu et al., 2009, p. 111) and a CFA revealed acceptable fit, where the model explained 74.92% of the variance in the data. Furthermore, the scale and subscales scores were significantly, positively correlated with self-reported enjoyment scores. Therefore, Fu et al. (2009) stated that their scale demonstrated both reliability and validity, and that future studies should focus on validating EGameFlow in more complex serious games.

Summary

The subjective gameplay experience is a combination of experiential states, to include immersion, involvement, presence, and flow. A revised model of game engagement (R-GEM) was proposed. With focused attention, the reciprocal states of immersion and involvement are possible, which are the low levels of game engagement. Low levels of game engagement can transition to higher levels of game engagement. Increased levels of immersion may result in presence. Deficits in the immersive cues provided by the game can be mitigated by involvement or individual differences. High levels of involvement may result in flow when a balance between challenge and skill is achieved.

The purpose of this work is to examine whether the R-GEM is representative of the subjective gameplay experience. Regarding this effort, there are many different ways to measure these subjective states. As Jackson et al. (2008) were careful to warn, as is "the case with all experiential phenomena, flow cannot easily be quantified by psychometrics or fully illuminated through investigative interviewing…A multimethod approach to researching flow will result in improved understanding of flow, and how to achieve it" (p. 582). This advice, of course, applies to the revised model of game engagement proposed in this work. Still, one must begin somewhere. The experimental approach reported in this work relies on subjective measurement to determine whether the revised game engagement model has merit. This will serve to organize and clarify future research efforts that examine the subjective gameplay experience. Then, additional measures of these constructs, such as physiological approaches, can be examined in an effort to further develop and validate the revised game engagement model.

CHAPTER THREE: METHODS

This study was conducted in two phases: an online prescreening survey and an in-person experimental session. To test the R-GEM, undergraduate students from the University of Central Florida were recruited based on their prescreening responses to play the game *Minecraft* during the in-person experimental session (Persson, 2009), after which they immediately responded to a number of surveys about their experience. The gameplay experience was manipulated with respect to the strength of the immersive cues provided as well as game difficulty so that players were afforded the opportunity to experience a variety of different game engagement experiences (e.g., experiencing presence but not flow, or experiencing both presence and flow).

Participants were randomized to either the high immersion or low immersion condition. Those in the high immersion condition played *Minecraft* with high-resolution, detailed textures and advanced lighting effects which was achieved through the use of programming modifications (‘Modded *Minecraft*’). Per Slater and Wilbur (1997), this enhances the vividness aspect of the game to make it more immersive. Those in the high immersion condition also played in full-screen mode at a resolution of 1920x1080 pixels with 128x128 pixel textures. Doing so does not increase the amount of game that the player sees, but does fill more of their field of view with the game. This is another aspect proposed to strengthen the immersive cues provided by the game (Slater & Wilbur, 1997). Participants in this condition also played the game while wearing noise-cancelling headphones. This minimized distractions and provided spatialized audio cues, both of which should increase immersion (Slater & Wilbur, 1997). Finally, while this experiment took place in a laboratory setting, participants played in an area blocked off by curtains from the rest of the room and with the lights dimmed to minimize the cues that suggested they are indeed in the University of Central Florida’s Psychology building. This helps to increase the strength of the cues provided by the game and make presence more likely to occur (e.g., Wirth et al., 2007).

Those in the low immersion condition played with the standard Minecraft textures and without the use of lighting shaders (‘Vanilla *Minecraft*’) to decrease the vividness of the experience. They played in windowed mode, which is set to a native resolution of 854x480 pixels to decrease the field of view and also featured 16x16 pixel textures. They also played with the Windows Start bar hidden and against a black background. External speakers were used to play game sound, the lights in the laboratory remained on, and curtains were not used to block off the experimental area. This decreased the strength of the cues provided by the game and made presence less likely to occur, in theory. Ultimately, these manipulations were used to ensure that those in the high immersion condition had a much richer sensory experience with stronger cues to support accepting the *Minecraft* world as the PERF than those in the low immersion condition.

Despite these differences, the R-GEM does not suggest that these sensory changes will affect actual gameplay enjoyment. Essentially, a player may experience highly enjoyable, involving gameplay regardless of the immersion manipulation. Those who are most likely to become involved—to the point of experiencing flow—are those individuals that are regular players of a game and who are playing the game in the way they most enjoy and at the level of difficulty that matches their skill level. Brockmyer et al. (2009) recruited experienced gamers for their study, as they were far more likely to experience flow in a laboratory setting because of their familiarity with gaming. Therefore, for this study, only experienced *Minecraft* players were recruited with the exception of a handful of non-players used in a pilot study. Whether the participants experienced flow was manipulated by providing gameplay experiences that aligned with their preferences and skill levels. For those in the congruent preferences condition, they played at the Normal difficulty, while those in the incongruent condition played on the Peaceful difficulty. Normal difficulty was chosen as, early on in testing, initial survey data indicated that most of the *Minecraft* players played on this particular difficulty setting. Playing on the Peaceful difficulty made the gameplay experience very boring for these experienced players due to the misaligned challenge/skill balance (Csikszentmihalyi, 1990), meaning that it was unlikely that these players would experience flow.

It is important to note that, as presence and flow are driven by subjective states, it may be entirely possible for an individual to feel so immersed in the non-immersive condition that they experience presence, or for an experienced player to still really enjoy *Minecraft* on the easiest setting and experience flow. Therefore, subjective measures of all states of the R-GEM (attentional requirements, immersion, involvement, presence, and flow) were measured, and scores of the different game engagement states were examined in lieu of the experimental manipulations.

Participants

Sample size was determined by utilizing the power calculation software *G*Power* (Faul, Erdfelder, Lang, & Buchner, 2007; Faul, Erdfelder, Buchner, & Lang, 2009). Assumptions included a power level ($1 - \beta$) of .80, a degree of freedom of 1, and a large effect size (f = .40). A two-way ANOVA with a total of four groups—two levels for each of the two independent variables—would require a total sample size of N = 52, with 13 in each group.

Pilot study. For an initial small pilot study, 16 participants from the University of Central

Florida were recruited, half of which were experienced *Minecraft* players. Participants were randomized to one of the four conditions. They played *Minecraft* and then completed the post-game surveys. This is simply to ensure that the manipulations did not cause any issues and provided some information how as how to improve the study (e.g., to use a difficulty of Normal instead of Hard for the congruent preferences condition).

Participants. For the experiment, 1,210 students were initially recruited from the University of Central Florida. They were recruited through the UCF Psychology Sona System (for Sona credit) as well as by word of mouth and posted advertisements (for no compensation). They completed a number of baseline questionnaires online, that were used to screen for 60 *Minecraft* players suitable to complete the experiment.

Measures

In light of the careful attention given to definitional issues, accurate measurement is absolutely necessary in order to effectively examine relationships in the R-GEM. The literature is wrought with differing definitions for each of the constructs of interest, so while the measures used should be reliable and validated, they should *also* have face validity with respect to the definitions put forth by this work. The most influential measures for each construct have been identified and were then evaluated for suitability. Adjustments were made to items as was deemed necessary to ensure face validity. See Appendix B for a complete list of the measures.

Demographic surveys. Basic demographic information about participants, as well as their gameplay habits and information about their level of VGSE was collected. It is suggested that demographic information, especially that which pertains to experience with video games, should be collected at the end of a game-based experiment as the questions may prime participants to behave in a certain way which would invalidate the results (Boot, Blakely, & Simons, 2011). For this study, however, it may be possible that having participants report on their VGSE and preferences for gameplay *after* playing the game will result in these scores being shaped by gameplay itself. Since these aspects will be considered as predictors rather than outcomes, it is important to measure them before exposure to the game to eliminate this potential confound. This also applies to the collection of individual differences data. Therefore, this information was collected during an online pre-screening survey before participants reported to the experimental session.

Basic demographic information**.** A basic demographic survey was administered to participants. Participants were asked to report their age, gender, ethnic background, year in school, and primary language.

Gaming preferences and behaviors**.** Participants completed a detailed gaming preferences survey, which has been used in previous studies conducted by the UCF RETRO Lab. They reported their comfort with video games as well as with computers by rating each on a 7-point Likert scale from 1 ("not at all") to 7 ("very"). Data about gameplay habits were also collected. Participants were asked to report how many hours per week that they play video games, in general, to select from a list all of the ways that they play video games (i.e., computer, console, cell phone, tablet, or handheld gaming device), and to select the medium on which they play the majority of games. Participants also were asked to list their top three favorite video games. Participants rated on the 7-point Likert scale described above as to how much they enjoy playing video games in general. They also were asked if they had ever played the game *Minecraft*. If they had played it, they were asked to estimate how many hours they had spent playing the game as well as to rate their comfort with and enjoyment of *Minecraft* on the same 7-point Likert scale described above. Finally, they were asked if they had ever played any video game using an HMD, and if they had ever played *Minecraft* with an HMD.

Video game self-efficacy (VGSE). Participants completed the 10-item Video Game Self-Efficacy Scale (VGSES). The VGSES was created by Pavlas et al. (2010), who adapted Schwarzer and Jerusalem's (1995) Generalized Self-Efficacy Scale to fit the gaming context. Previous research found that scores on the VGSES were significantly, positively correlated with the previously described comfort with games and hours spent gaming measures (Procci et al., 2013). Participants rated each of the 10 items with respect to themselves on a 4-point Likert scale from 1 ("not at all true") to 4 ("exactly true"). A total VGSE score was computed.

Minecraft screening. If participants reported that they had played Minecraft, they completed an additional survey where they responded to questions about their typical *Minecraft* gameplay behaviors. Participants reported if they played *Minecraft* on a computer, console, cell phone, or tablet, as well as identified their preference. They also reported what versions of *Minecraft* they had played, again identifying their preference. They also were asked about their typical *Minecraft* gameplay methods, to include screen size and resolution, and their use of texture packs, mods, and shaders. They rated their enjoyment of the three different *Minecraft* gameplay modes, as well as listed their preferred difficulty in Survival Mode. They were asked whether they played *Minecraft* in windowed mode or full-screen mode, with or without sound, and whether they preferred to play alone. These questions were used to determine whether a potential participant qualified for inclusion in the in-person experimental session.

Individual differences influencing game engagement. The literature revealed several different individual differences that may influence different aspects of the revised game engagement model. Immersive and involvement tendencies, dispositional tendency to experience flow, willingness to suspend disbelief, creative imagination, and trait absorption were measured after completing the demographic surveys during the online prescreening portion of the study.

***Immersive/involvement tendencies*.** The ITQ measures an individual's propensity to become immersed in activities in general (Witmer & Singer, 1998). The ITQ is an 18-item scale, where each item is rated on a 7-point Likert scale. Witmer and Singer (1998) found that the ITQ was reliable (Cronbach's $\alpha = 0.81$; p. 235). The item scores are summed, and Witmer and Singer (1998) found that the average ITQ score was 76.66 out of 126 possible points ($SD = 13.61$; p. 236). In addition to the total score, the ITQ produces three subscale totals: The involvement subscale had seven items ($M = 26.51$ /49, $SD = 7.24$; Witmer & Singer, 1998, p. 236) and pertained to the tendency to become involved in passive activities; the focus subscale also had seven items and measured respondents' ability to concentrate and block out distractions ($M = 40.33$ / 49, $SD = 6.07$; Witmer & Singer, 1998, p. 236); and, the games subscale was comprised of two items ($M = 6.21$ /14, $SD = 3.16$; Witmer & Singer, 1998, p. 236) that quantified gameplay behaviors, such as frequency of play and extent of involvement in gameplay. Despite its name, the ITQ measures elements of immersion *and* involvement, as well as the requirement of focused attention. The ITQ may instead be more useful to measuring general low-level game engagement tendencies rather than only immersion as it covers all of these related aspects. Therefore, a total ITQ score, rather than subscale scores, was used to measure general immersive/involvement tendencies.

***Dispositional flow*.** To measure the tendency to experience flow, the SHORT

Dispositional Flow Scale (S DFS) was used (Jackson et al., 2008; Jackson et al., 2010a, 2010b). Participants were asked to reflect on their experiences in general when playing games. The S DFS features 9 items, one for each of the flow dimensions, each of which is rated on a 5-point Likert scale from 1 ("never") to 5 ("always"). Jackson et al. (2008) reported that the S DFS was reliable (Cronbach's α = .77; p. 568). The S DFS mean score will be used in this study rather than its 36-item DFS-2 counterpart as it has been found that the S DFS is a simple, yet valid measure of general flow tendency. The short version of the DFS was also found to be more robust than the short version of the FSS-2, and is perfectly acceptable for use in situations in which granular details regarding the nine elements of flow proposed by Csikszentmihalyi (1990) are not necessary (Jackson et al., 2008). Therefore, a mean dispositional flow score based on the S DFS was computed for each participant.

Willingness to suspend disbelief. Lombard and Ditton (1997) were the first to suggest that willingness to suspend disbelief was an important individual difference to consider regarding likelihood to experience presence in a VE. Willingness to suspend disbelief encompasses overlooking obvious signs that a VE, or any other mediated environment, is not real simply because the individual wants to fully engage with the experience (Lombard & Ditton, 1997). Unfortunately, there is no standardized measure for willingness to suspend disbelief. Sas and O'Hare (2004) did create their own single item used to measure willingness to suspend disbelief, where participants were asked to rate a statement ("To what extent were you willing to be transported to the virtual world") on a 7-point Likert scale from 1 ("not at all") to 7 ("completely"; p. 529). Sas and O'Hare (2004) found this single item to be significantly and positively correlated with scores from other validated presence questionnaires. They also found that, in a multiple regression analysis predicting presence score, this item explained roughly 45% of the variance. Therefore, the score from Sas and O'Hare's (2004) single item was used to measure willingness to suspend disbelief in this study.

Scales with more than one item are typically preferred as the use of multiple items tends to average out error and are robust against carry-over effects from previous items (Diamantopoulos, Sarstedt, Fuchs, Wilczynski, & Kaiser, 2012). Even still, Diamantopoulos et al. (2012) found that single item measures can be used in studies if only several requirements are met, most of which are impossible in empirical research. As such, they suggested that a minimum of four items should be used. To address this issue, three additional items about willingness to suspend disbelief have been created to form the new Willingness to Suspend Disbelief Scale:

1. How unwilling are you to suspend disbelief while playing the game?
2. How unwilling are you to become a part of a game?
3. How willing are you to ignore reality and really get into a game?

These items were subjected to correlational analysis to ensure that they all represent the same construct. Based on this analysis, either the single-item score from Sas and O'Hare (2004) or the mean score across all four items will be used, where higher scores indicate more willingness to suspend disbelief.

***Creative imagination*.** Creative imagination also has been found to be an important individual difference that explains whether an individual may feel present (Sas & O'Hare, 2004). Wilson and Barber (1978) developed the Creative Imagination Scale (CIS) to quantify an individual's level of creative imagination. First, an individual is provided a set of instructions in which they are encouraged to think creatively (Barber & Wilson, 1977). The usage of these instructions is suggested by Barber and Wilson (1978), however the original text was shortened for use in this study. The original instructions are nearly two pages in length, however the one paragraph introduction suffices. The CIS takes an individual through 10 exercises where they are asked to imagine different things. For consistency, an audio recording of each of the 10 exercises was used. Afterward, participants completed a brief survey about their experience during the exercises. Each item was rated as a number between 0 and 4, where each level of scoring corresponded to how real the participant felt the experience to be (e.g., 0 = "0% Not at all the same" to 4 = "90+% Almost exactly the same" as reality; Wilson & Barber, 1978, p. 237). The CIS takes roughly 18 minutes to complete.

Wilson and Barber (1978) administered the CIS to 217 college students taking psychology courses. These data were used to generate a set of norms, and the authors provided *t*-score conversion tables. Reliability was examined through both test-retest ($r = .82$) and split-half ($r = .89$) correlation, both of which were found to be significantly positive (Wilson & Barber, 1978, p. 239). Validity was examined through both correlation and principal components analysis. All items in the scale were found to be highly correlated with the total score, indicating that all items measured the same construct. The CIS score was found to be moderately correlated with a similar construct, further suggesting validity. Finally, the principal components analysis found that all items significantly loaded onto one factor that accounted for about 46% of the variance, which again suggested that the scale was valid as well as determined that it captured one global aspect that was purported to be creative imagination (Wilson & Barber, 1978). Sas and O'Hare (2004) also reported that the measure was highly reliable, listing a Cronbach's α of

.83 that was reported by Barber and Wilson (1979, as cited in Sas & O'Hare, 2004, p. 531). In this study, the total CIS *t*-score was calculated for each participant.

***Trait absorption*.** To measure trait absorption, which is the tendency to become fully engrossed in whatever an individual is focusing their attention on, the TAS was used (Tellegen & Atkinson, 1974) with instructions provided by Kihlstrom (2011). Please note that the TAS has become a part of Tellegen's Multidimensional Personality Questionnaire™ (MPQ™; Copyright © 1995, 2003 by Auke Tellegen. Unpublished test. Used by permission of the University of Minnesota Press. All Rights Reserved; "MPQ Overview," 2011). The TAS has 34 true-false items, where higher rates of endorsement indicate more absorption tendency, such as becoming easily captivated by stimuli, having cross-modal experiences, or experiencing an altered awareness ("MPQ Standard," 2011). As reported in an earlier section, Tellegen and Atkinson (1974), Glisky et al. (1991), and Wild et al. (1995) found the TAS to be a valid measure. The TAS is also reliable. For example, in a longitudinal study involving 88 participants, 45 of which were female, Kremen and Block (2001) found the Cronbach's α to be .87 at age 18 and .92 at age 23, and that the correlation between the administrations was also high (r = .66 for the males and r = .53 for the females; p. 254). The TAS included several different subscales (Glisky et al., 1991), although the focus of this study will be on only the overall total score, which was calculated for each participant.

Game engagement states. One of the more difficult challenges is the accurate assessment of the many subjective states of game engagement: immersion, involvement, presence, and flow. Furthermore, it is also important to measure the pre-requisite for low-level game engagement: focused attention. Each of these measures were collected during the in-person experimental session after playing the game.

***Focused attention*.** Recall that the PQ deliberately does not measure presence directly as its purpose was to measure its influences, which were immersion, involvement, and the attentional requirements for both (i.e., low-level game engagement; Singer & Witmer, 1999). Witmer et al. (2005) conducted a factor analysis, which divided the PQ's items into the involvement, adaptation/immersion, sensory fidelity, and interface quality clusters, yet, as discussed previously, the labeling of these subscales can be considered somewhat misleading.

Because of this, each of the 32 items of the third version of the PQ as provided by Witmer et al. (2005) have been analyzed according to the definitions set forth in this work. Three items were eliminated from the item pool because they too closely resembled items that should belong in a flow measure (e.g., "Were you involved in the experimental task to the extent that you lost track of time?"; Witmer et al., 2005, p. 302). The remaining 28 items were then sorted into three new subscales: immersion (18 items), involvement (1 item), and attentional requirements (10 items). These items have also been modified to relate directly to video games. The proposed attentional requirements subscale contains items such as "How much did the control devices interfere with playing the game?" and "To what extent did events occurring outside the virtual environment distract from the game?" (adapted from Witmer et al., 2005, p. 302). Participants rated each of these items on a 7-point Likert scale from 1 ("not at all") to 7 ("very much"), and some were are reverse-coded. To quantify aspects of attention and focus, a total score from the revisited PQ's 10-item attentional requirements subscale was calculated, where higher scores indicate greater levels of focused attention. A total score was used rather than a mean as the original PQ used totals as part of its score calculations.

Immersion. Immersion can be measured with items from the PQ (Witmer & Singer, 1998; Witmer et al., 2005). As mentioned above, an analysis of the third version of the PQ identified 18 items that applied to immersion, such as the individual's subjective response to the sensory (e.g., "How much did the visual aspects of the environment involve you?") and naturalistic (e.g., "How much did your experiences in the virtual environment seem consistent with your real world experiences?") aspects of the VE (Witmer et al., 2005, p. 302). Two items were removed from this item set because they pertained to haptics, which are not a part of the game played in this study. Each of the remaining 16 items were rated on a 7-point Likert scale from 1 ("not at all") to 7 ("very much"), and some items were reverse-coded. A total score for the sixteen immersion items from the PQ was found, where higher scores indicate greater levels of immersion.

Involvement. Involvement was measured after gameplay by the Revised PII (Zaichkowsky, 1994), as well as by four additional items inspired by the perceived usefulness subscale of the TAM (Davis, 1989), and the lone item from the PQ that measures involvement (Witmer et al., 2005).

The Revised PII (Zaichkowsky, 1994) is a 10-item semantic differential scale for measuring involvement with an advertised product. For each item, the participant was asked to mark the location on a 7-point scale between two words in a word pair (e.g., interesting and uninteresting) that represents their appraisal of a product. These ten items are divided into two five-item subscales: affective and cognitive. The affective subscale refers to the individual's emotional appraisal of the product, while the cognitive subscale refers to the individual's perception of the product's utility. The Revised PII is internally consistent (Cronbach's α = .91 to

.95), reliable (*r* = .73 to .84), and features content validity (Zaichkowsky, 1994, pp. 61-62). The Revised PII returns a score between 10 and 70, which is then classified into three levels of involvement: low (10-29); medium (30-50); and high (51-70; Zaichkowsky, 1994, p. 62).

Since the Revised PII is geared toward advertisements, the utilitarian nature of the items in the cognitive subscale may not apply to game engagement. This was addressed by editing the instructions slightly ("Think about why you might want to play a video game (for fun, to relax, etc.). You will judge the video game you just played against a series of descriptive scales according to how YOU perceived playing the game fit your gameplay needs."), as well as by including a sentence to anchor the word pair ratings ("Playing the game was ____ to me and my gameplay needs.").

An additional four items, inspired by the content of the format of the items from the perceived usefulness subscale of the TAM (Davis, 1989), were created. These four items were rated on a 7-point Likert scale from 1 ("strongly disagree") to 7 ("strongly agree"), and half of the items were reverse-coded. These items addressed aspects of involvement in gameplay not specifically addressed by the Revised PII:

1. I did not feel motivated to play this game.
2. Playing this game fulfilled my gameplay needs.
3. I developed a desire to play and enjoy this game.
4. I disliked playing this game.

Finally, participants also were asked to respond to the single involvement item from the PQ: "How involved were you in the virtual environment experience?" (Witmer et al., 2005, p. 302). They were asked to rate this item on the same 7-point Likert scale used for the TAM-inspired items, therefore the wording was changed slightly from its original form to be consistent with the other items.

The overall measure of involvement was aggregated based on the results of the ten-item Revised PII (Zaichkowsky, 1994), the four items inspired by the perceived usefulness subscale of the TAM (Davis et al., 1989), and the single involvement item from the third version of the PQ (Witmer et al., 2005). The items were summed to generate a total score from 15 to 105, where higher scores indicate greater involvement. Like Zaichkowsky (1994), the scores were divided into three sections to classify participants as having experienced low (15-44), medium (45-75), or high (76-105) levels of involvement during gameplay, although only the total score was used for the purposes of analysis in this study.

Presence. To measure presence, a combination of items from Slater and Steed (2000) and Witmer et al. (2005) was used. Recall that Slater and Steed (2000) created an equation for presence based on subjective ratings and the number of breaks in presence (BIPs) experienced. In Slater and Steed's (2000) study, participants responded to five items, each of which was rated on a 7-point Likert scale, where higher scores were more indicative of having experienced presence. For each item, if the participant rated it a 6 or 7, their presence count was increased by one point as Slater and Steed (2000) believed this to mean that they had likely experienced presence. Presence count was found to be significantly, positively, and strongly correlated with Slater and Steed's (2000) measure of presence that was derived from BIPs. Given that reporting a BIP may be disruptive to the holistic gameplay experience, only the five subjective items were used to measure presence and the participant will not be trained to report BIPs. Additionally, Witmer et al. (2005) also proposed four new items to add to the PQ, each of which concerned presence directly. These items augmented Slater and Steed’s (2000) five existing items.

Therefore, presence was measured by having participants respond to all nine items—five items from Slater and Steed (2000) and four items from Witmer et al. (2005)—each of which was rated on a 7-point Likert scale, where higher scores were indicative of experiencing presence, and then calculating a presence count (Slater & Steed, 2000). Like Slater and Steed (2000), the number of times the participant endorsed each presence item with a 6 or 7, a point was added to the participant’s presence count. Since there are nine total questions, the presence count ranged between 0 and 9.

Finally, Slater and Steed (2000) also suggested that participants should estimate the percentage of time that they felt present in the VE (presence estimate), where those who reported values above 50% may be considered as having experienced presence (presence classification). Participants were asked to estimate this. This is a largely novel measure, but it is interesting to see if it is associated with the presence count, but this portion of the analysis was entirely exploratory.

Flow. To measure the mechanics of the flow experience, the FSS-2 (Jackson & Eklund, 2002; Jackson et al., 2010a, 2010b) was used. Based on the results of their validation studies, Jackson et al. (2008) suggested that the long versions of the scales should be used when an in-depth examination of flow was the focal point of the study. The short scales are more practical, as they are simplified and shorter, but only provide broad information about the flow experience (Jackson et al., 2008).

The FSS-2 is a 36-item scale. Participants were asked to reflect on their gameplay experience, and then to rate each item on a 5-point Likert scale from 1 (“strongly disagree”) to 5 ("strongly agree").There are four items for each of the nine flow dimensions, which were derived from Csikszentmihalyi's (1990) nine elements of enjoyment and flow: Challenge/Skill Balance; Merging of Action and Awareness; Clear Goals; Unambiguous Feedback (which was previously labeled 'immediate feedback'—they are essentially the same); Concentration; Control; Loss of Self-Consciousness; Transformation of Time (which was previously labeled 'time distortion'); and, the Autotelic Experience. Jackson et al. (2008) advised that means for each dimension should be used, rather than an overall flow score. Averages across a variety of domains for each dimension can be found in Jackson et al. (2010a).

As noted previously, the FSS-2 has been validated in the sports domain, but it also has been found applicable to several non-physical activities (Jackson et al., 2010a). Previous studies have established validity through extensive testing and found support for nine dimensions that, while they do all measure one global construct, each represent unique aspects of the experience (e.g., Jackson & Eklund, 2002; Jackson et al., 2008). Previous studies have also found the FSS-2 to be highly reliable, where Cronbach's α for each of the subscales consistently ranged from .80 to .90, with a mean that hovered around .85 (Jackson & Eklund, 2002; Jackson et al., 2008; Jackson et al., 2010a, p. 31).

Still, Procci et al. (2012) did not find the FSS-2's dispositional counterpart, the DFS-2, to align well with gamers' tendency to experience flow in gameplay. One of the study's stated shortcomings was that it did not survey participants after they played a game, therefore it is not clear whether the FSS-2 shares the DFS-2's limitation. The FSS-2, however, is still one of the most well-validated measures of flow available, and does measure each of the nine elements of flow and enjoyment—which is a critical component of effective flow measures (Delle Fave et al., 2011). Therefore, the FSS-2 it is worth reconsidering for measuring flow in gamers.

Since the FSS-2 taps into each of the nine elements of flow and enjoyment, the dimension subscales are likely useful when applied to the different aspects of the R-GEM. For example, the subscales pertaining to Goals and Unambiguous Feedback overlap with the involvement construct, but the Challenge/Skill Balance subscale should be unique to flow. Mean scores for each of the nine mean dimension subscales were found, as well as an overall mean flow score. Additionally, standardized scoring tables have been provided by Jackson et al. (2010a), which allows for conversion to standardized *t*-scores. However the authors also stated that the use of the standardized scores need to be interpreted with caution if the distributions across factors are not similar. Therefore, only the subscale means will be used in this study.

In light of Procci et al.'s (2012) findings, it may be useful to take a secondary approach to flow measurement in games, and collect data related to the phenomenological experience of flow separate from the nine elements. The creators of the FSS-2 also developed the CORE Flow State Scale (C FSS; Martin & Jackson, 2008) that addressed this very need. The C FFS (Martin & Jackson, 2008) is unique from the state and dispositional flow scales because its items are targeted at measuring aspects of what an individual in flow may experience. Rather than asking them to quantify elements of the flow experience (e.g., how clear they thought goals were), they are instead asked to rate items that holistically describe flow experiences (e.g., if they felt like they were 'in the zone'; Jackson et al., 2010b). The C FSS is a 10-item scale that features experiential descriptions, for each of which participants rated their agreement on a 5-point Likert scale from 1 ("strongly disagree") to 5 ("strongly agree") after playing the game. These items were based on previous qualitative research conducted by Jackson (1992, 1995, 1996, as cited in

Martin & Jackson, 2008), and reflect actual statements from individuals describing their flow experiences. All items are summed and then divided by 10 to find a mean CORE flow score (Jackson et al., 2010b). Jackson et al. (2010a) warned that if more than two responses are missing, the score may no longer be valid.

Martin and Jackson (2008) examined the validity of the C FSS in a sample of 220 volleyball players. Participants were asked to complete the C FSS immediately after playing. They found a mean CORE flow score of 3.37 out of 5 (*SD* = 0.75; Martin & Jackson, 2008, p. 148). The C FSS was found to be highly reliable (Cronbach's α = .92; Martin & Jackson, 2008, p. 148), and the scores were nearly normally distributed with only minor skew and kurtosis. Factor analysis found that each of the items loaded strongly onto a central flow factor, ranging between .59 and .85, for an average of .83 (Martin & Jackson, 2008, p. 147). A CFA revealed good fit across several indices, and the scores of the C FSS were found to be significantly correlated in hypothesized ways with a number of theoretically-related constructs. Invariance was also found across diverse samples. Based on these results, Martin and Jackson (2008) stated that the C FSS is both reliable and valid.

In summary, to measure flow, participants completed both the FSS-2 (Jackson & Eklund, 2002) and C FSS (Martin & Jackson, 2008) immediately after gameplay. The mean score of the items from the C FSS was used as an overall, global measure of the flow experience, while the nine FSS-2 dimension means was used to assess the R-GEM on a more granular level.

Model-validating variables. There are some assumptions present in the R-GEM that should be empirically evaluated. Given that the attentional requirements do focus on whether poor usability results in distraction, overall usability was measured. Additionally, the evolving balance between challenge and skill is an important aspect of what may cause involvement to progress to flow. This relationship should be thoroughly examined, as well. Witmer and Singer (1998) found that experiencing simulator sickness was negatively correlated with the PQ. Therefore, ratings of simulator sickness will also be collected. The literature review (e.g., Brockmyer et al., 2009) revealed that there may be some question as to whether the state of cognitive absorption, as defined by Agarwal and Karahanna (2000), differs from that of flow (Csikszentmihalyi, 1975a, 1975b, 1990). Therefore, cognitive absorption will also be measured. Measurement of these subjective states occurred after gameplay.

Usability**.** To measure usability, the System Usability Scale (SUS; Brooke, 1986, as reprinted in Brooke, 1996) was administered after gameplay. The SUS is a brief, 10-item measure of summative usability. Many view the SUS as "an industry standard", and it has been found to be highly reliable, valid, and, most importantly, useful (Brooke, 2013, p. 29). Regarding reliability, Bangor, Kortum, and Miller (2008) reviewed nearly a decade's worth of research on the SUS and found it to be highly reliable (Cronbach's $\alpha = 0.91$; p. 581). The researchers also noted that the SUS has excellent face validity. A factor analysis from Lewis and Sauro (2009) found that the SUS actually had two factors, usability and learnability, however the single SUS score is still perfectly valid to use alone. Regarding applicability, SUS scores have been found to be significantly, positively correlated with actual market success in smartphones (Bangor, Joseph, Sweeny-Dillon, Stettler, & Pratt, 2013) and with an objective measure of usability related to number of errors made, where higher scores indicated fewer errors (Peres, Pham, & Phillips, 2013). This further supports the notion that the SUS is valid and useful.

Each item of the SUS is scored on a scale from 1 to 5, which is then converted to scores from 0 to 4, where some items are reverse-coded. The total is then multiplied by 2.5 for a score between 0 and 100 (Brooke, 1996). This score is not to be interpreted as a percentage (e.g., the game was 88% usable), but can be converted into usability percentiles. Sauro (2011) reviewed 500 different studies that used the SUS and found the average score to be 68, which is the 50th percentile. He also found that SUS existed on a curve, where a score of 80.3 and above was the top 10 percentile of scores. In this study, the raw SUS score is used.

***Simulator sickness*.** Simulator sickness is a group of symptoms, such as nausea, blurred vision, and dizziness, which may occur during simulator exposure. Measuring simulator sickness is especially relevant as it should be antithetical to the experience of game engagement (e.g., Witmer & Singer, 1998). It is important to measure simulator sickness as it may be detrimental to the game engagement experience, thus confounding results unless accounted for.

Kennedy et al. (1993) determined that simulator sickness is similar to, yet unique from motion sickness in that it has different and less severe clusters of symptoms. Kennedy et al. (1993) analyzed a dataset of 1,119 pre- and post-test scores of the Pensacola Motion Sickness Questionnaire (MSQ, Kellogg, Kennedy, & Graybiel, as cited in Kennedy et al., 1993) used across 10 simulators in order to determine what factors may be unique to simulator sickness. Of the 28 symptoms in the MSQ, 16 were retained for analysis based on their frequency of occurrence in non-motion simulators. Factor analysis revealed that there were three main symptom clusters in simulator sickness: oculomotor disturbances, such as blurred vision and headache; disorientation, such as dizziness and vertigo; and nausea, such as stomach awareness and burping (Kennedy et al., 1993). These 16 items were grouped together as part of the Simulator Sickness Questionnaire (SSQ; Kennedy et al., 1993).

Stanney and Kennedy (1997) reported that the SSQ produced reliable, valid symptom profiles across four different VEs using HMDs. Others have reported similar findings. For example, as evidence of validity, the disorientation subscale score was found to be significantly, positively correlated with postural instability after simulator usage (Kennedy, Berbaum, & Lilienthal, 1997). Balk, Bertola, and Inman (2013) conducted a factor analysis in 530 participants and found that their results mirrored those of Kennedy et al. (1993), further suggesting validity. As for reliability, Kennedy et al. (2001) reported a split-half correlation for the SSQ to be $r = .80$ (p. 8), which is acceptably reliable.

The SSQ is administered in two parts: There is a pre-task questionnaire which asks the participant to report their general state of health and to rate their current experience of 16 symptoms on a 4-point scale, where higher scores indicate greater severity, and a post-task questionnaire in which they once again rate the symptoms on the same scale (Kennedy et al., 1993). Only those who are currently healthy should have their SSQ scores calculated (Kennedy et al., 1993). The SSQ is scored by multiplying each symptom's post-task score by the weight provided by Kennedy et al. (1993, see p. 212), and then summed for the weighted total. From this, weighted total scores for the Nausea, Oculomotor, and Disorientation subscales are found. A Total Severity score is found by summing the three weighted subscale scores and applying a conversion formula provided by Kennedy et al. (1993, see p. 212). This method results in a score of 0, which actually means that zero symptoms were experienced, and produces a standard deviation of 15 in Kennedy et al.'s (1993) sample. Higher scores indicate more simulator sickness, and normative data based on the analysis of 1,119 scores can be found in Kennedy et al. (1993, p. 214).

The participants completed the SSQ's baseline questionnaire and pre-task symptom ratings as a part of the introductory surveys the day of the in-person experimental session, and then they completed the post-task symptom ratings after gameplay. For this study, the Total Severity score was calculated.

Cognitive absorption**.** Cognitive absorption was measured using the items provided by Agarwal and Karahanna (2000). Recall that the theory of the state of cognitive absorption, as described by the authors, is essentially a combination of flow theory along with that of cognitive engagement, which was also grounded in flow theory. The purpose of including these items is to determine whether cognitive absorption and flow are independent constructs.

Agarwal and Karahanna's (2000) Cognitive Absorption Scale (CAS) includes 20 items divided amongst five subscales: Temporal Distortion, which included five items about the subjective distortion of time, which is also described similarly in flow theory; Focused Immersion, which included 5 items pertaining to both the ability to concentrate and ignore distractions as well as the subjective appraisal of becoming immersed; Heightened Enjoyment, which were four items adapted from the perceived enjoyment subscale of the TAM (Davis, Bagozzi, & Warshaw, 1992, as cited in Agarwal & Karahanna, 2000); Control, which included three items pertaining to a sense of control, which is also similar to flow theory; and, Curiosity, which included three items that pertained to being curious about the task. The Control and Curiosity items were adapted from a previous flow measure (Webster et al., 1993, as cited in Agarwal & Karahanna, 2000). Each of these items is rated on a 7-point Likert scale from 1 ("strongly disagree") to 7 ("strongly agree"), with 4 as a midpoint anchor ("neutral"). Some items were reverse-coded.

Agarwal and Karahanna (2000) conducted a validation study in 288 students who completed the measure of cognitive absorption as part of a survey on their experiences with internet usage. The items were found to be reliable, with Cronbach's α ranging from .83 to .93 (average Cronbach's α = .90; Agarwal & Karahanna, 2000, p. 679). Factor analysis produced factor loadings ranging from .64 to .87 (average of .75; Agarwal & Karahanna, 2000, p. 684), which the authors took to imply validity. For this study, an overall mean score was calculated and then compared to the CORE flow score, while subscale means were compared to the FSS-2 dimension subscales.

Equipment and software

Computer hardware. The computer used in this experiment was a custom-built machine running the Windows 8.1 64-bit operating system. The computer had an Intel Core i7-4770K processor (model BX80646I74770K, 3.5 GHz), an ASUS motherboard with an Intel 787 chipset (Maximus Hero IV model, 6Gb/s), an EVGA GeForce GTX 760 graphics card (model 04G-P4-2768-KR, 4GB VRAM), and 16 GB of DDR3 1600 RAM (G.SKILL Trident X series, model F3-1600C7D-16GTX). The computer also featured a 3 TB Seagate Barracuda hard drive (model STBD3000100, 7200 rpm), which was set to never sleep by changing settings under the Power Options menu in the operating system. Since *Minecraft* writes a large amount of data to disk, the hard drive spinning up and down may cause lag unless this setting is changed. A Logitech 2.4 GHz wireless mouse (model MK320) and wired Logitech keyboard (model K120) also were used as the input devices.

Display. The game was displayed on an HP LCD Monitor (model 2709m, 0.311 mm pixel pitch), which measured 27" inches on the screen's diagonal. The monitor was connected to the computer via HDMI and was routed through a ViewHD HDMI 1x2 splitter (model VHD-1X2MN3D). The monitor also featured a maximum resolution of 1920x1080 pixels. As mentioned previously, those in the low immersion condition played *Minecraft* in windowed mode, which meant that their game filled only 854x480 pixels in the center of the screen. The remaining screen space was a black background and the Windows task bar was minimized. These participants were not allowed to change the window size of the game. Those in the high immersion condition played Minecraft at the full screen resolution of 1920x1080 pixels.

Sound. Those in the high immersion condition wore a pair of Panasonic noise canceling headphones (model RPHC200K) set at 50% of the maximum computer volume. Those in the low immersion condition heard sound through two standing 5 watt Logitech speakers (model Z130) connected to the computer via USB. The speakers were set at 1/3 of the computer volume, which was also set at 50%.

Game. *Minecraft* version 1.7.2 (Persson, 2009) was used in this study. *Minecraft* is a popular sandbox adventure game that can be played on the PC, console, tablet, or cellphone. There are two main gameplay modes: Creative Mode, in which players are free to build whatever they want in the virtual environment using any of the game's materials; and Survival Mode, in which the player attempts to survive in a world where they can be injured by aggressive creatures at night and may starve to death if they do not secure a food source and eat on a regular basis.

The defining characteristic of *Minecraft* is that it is a low-resolution, voxel-based world in which the virtual environment (e.g., dirt, tree trunks, gravel, water) is made up of large, perfectly square blocks. The virtual environment features a surface world, rich with different environmental biomes (e.g., tree-filled forests, deserts with dunes made up of sand blocks, snowy tundra), and expansive underground cave systems and abandoned mines. In Creative Mode, these blocks can be used like virtual Legos. In Survival Mode, players harvest materials from blocks and craft them into items that help them survive. Players must work to gain these materials, such as by mining for ores and cutting down trees for wood using tools they have crafted.

These blocks are procedurally-generated: Every new world in *Minecraft* is nearly limitless, with the exception of known depth and height restrictions and the memory capabilities of the computer the game is played on. Despite the vastness and complexity of the virtual environment, the game itself is of very low graphical fidelity. The only non-cube features of the virtual environment are the models for the player character and in-game creatures, which are still blocky to fit the style of the game. Furthermore, the textures applied to the blocks are 16-bit, which is a very low, visibly blocky resolution. This is in stark contrast to other modern games, which strive to appear as realistic as possible by using much higher texture resolutions. These aspects make *Minecraft* unique.

All participants began in a newly-generated world in Survival Mode, with the default world type and without cheats enabled. Each world used the seed “minitx” for map generation. This ensured that the game will offer roughly the same experience to all players. In Survival Mode, players can die if they are attacked by monsters or if they do not find food fast enough and starve. When players die, they lose everything in their inventory and are returned to the original spawn point or to the last bed that they slept in. Those in the congruent preferences condition had the game set to the Normal difficulty level, while those in the incongruent preferences condition played on the Peaceful difficulty level. When playing on the Peaceful difficulty level, monsters do not spawn and the player’s hunger bar does not deplete. The only way the player can die on this difficulty is if they fall from a high place or if they stay too long under water without coming up for air.

Those in the low immersion condition played *Minecraft* in windowed mode (854x480 pixels) while those in the high immersion condition played at the full screen resolution (1920x1080 pixels). Also, those in the low immersion condition played Vanilla *Minecraft*, in which no textures or mods were used. Those in the high immersion condition played *Minecraft* with several modifications to make the experience more vivid.

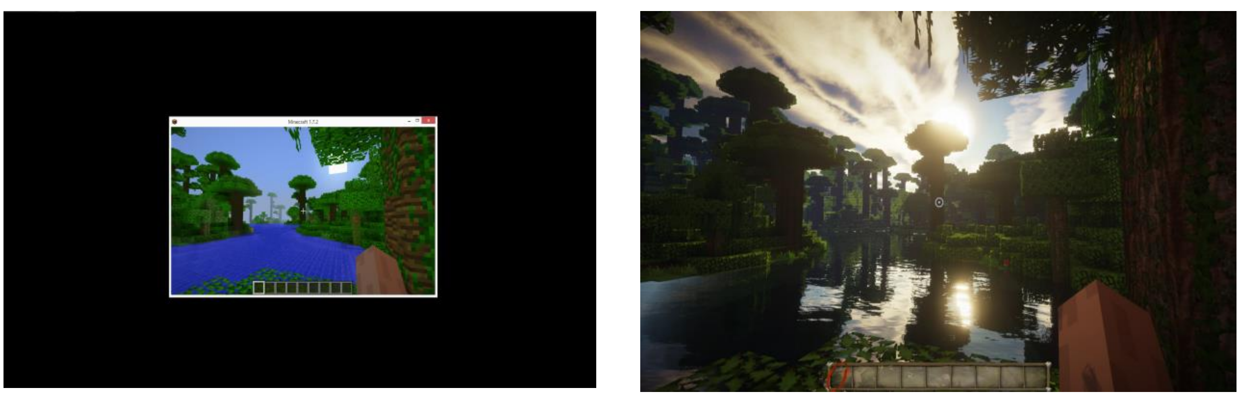
In *Minecraft*, resource packs are used to change in-game textures and feature various resolutions. Those in the high immersion condition played with the Chroma Hills resource pack (version 1.0.8, 128x128 texture resolution; "Chroma Hills", n.d.) which contains normal maps. Applying a normal map to a texture provides complex shape information for an otherwise simple, or even flat, model ("Normal Map", n.d.). When a normal map is applied, light will reflect and shade the textured model more realistically and as if it was actually 3D rather than flat.

Participants in the high immersion condition also played with the use of shaders. In *Minecraft,* shaders are used to add several features to the game environment that provide additional detail, which supports the experience of immersion. To achieve this, several modifications were made. Specifically, *Minecraft* Forge 1.7.2 (version 10.12.0.1033; "Minecraft Forge Downloads", n.d.) was installed to support modifications. The OptiFine mod (version 1.7.2-HD-U-D1; "OptiFine Downloads," n.d.) was also used to add video setting options and to improve overall performance. The use of shaders was supported by installing the GLSL Shaders 2.3.12 mod (version 2.3.12; Karyonix, 2014).

The shader used was Sonic Ether's Unbelievable Shaders (SEUS, version 10.1 preview 2; Sonic Ether, 2014), which added several immersive elements to *Minecraft*, to include dynamic lighting and shadows, plants that waved in the wind, realistic clouds, water physics, special sky effects based on the weather, and even puddles that formed on the ground when it rained. Many of these effects were only made possible due to the normal maps found in the Chroma Hills resource pack. Additional normal maps were created using Adobe Photoshop CS6 and NVIDIA's texture tool for the acacia leaves, spruce leaves, packed ice, and hay block textures.

Several settings were changed to further optimize performance. Specifically, the native *Minecraft* clouds were turned off and the render distance was set to 16 in the Video Settings. Cloud shadows also were turned off in the Shaders menu. In the *Minecraft* launcher, the game was set to use up to 10 GB of RAM to prevent crashes. The game also occasionally suffered from a server tick exception loop crash, which was fixed by making changes to a Forge config file (Caveman, 2012). Auto-saving was set to every 30 minutes to prevent lag, which may occur when saving the game. The ability to set custom colors was added by installing the Optifine mod, however this was toggled to "off" because it occasionally turns some in-game textures black (SycloneSJS, 2014).

The shader itself was also tweaked for the purposes of this experiment. Given that it was not a finished build, there were a handful of bugs and incomplete functionality was present. For example, the night time was too dark to for the player to see and the trees no longer waved in the wind. The night was made less dark by following a tutorial posted by YouTube User GOzzi (GOzzi, 2014). Occasionally, broken block particles and dropped blocks would turn black if they were in a shadow. The particles bug was fixed by commenting out line 37 in the gbuffers_textured.fsh file. The dropped blocks bug was fixed by adjusting values in line 61 of the gbuffers_textured_lit.fsh file. Minor edits to the code were made to have trees and tall grass wave in the wind (“SEUS Modifications”, 2014). There was a bug where blocks being destroyed turned blue, however this was fixed by toggling “tweakBlockDamage” to “On” in the shader’s settings menu. An error related to the gbuffers_weather.fsh file was fixed by deleting an extra ‘f’ from line 21, and commenting out line 40 fixed a bug where it would rain indoors. The shader also added puddles, which form when it rained in *Minecraft*. In this version of the shader, the effect was distracting, in that the ground became too shiny and reflective despite the use of normal maps. Changing some values in the shader’s code made the effect much more subtle, which was essential for the success of this experiment. Finally, the shader featured its own skybox, which clashed with the skybox found in the Chroma Hills resource pack. The file “sky1.png” was removed from the resource pack so that only the shader’s skybox was displayed during play. See Figure 9 below for a comparison of the *Minecraft* immersion conditions.



Low Immersion **High Immersion**

Figure 9. Minecraft texture pack comparison by immersion condition.

Compared to the set-up for the low immersion condition, playing *Minecraft* in the high immersion condition is very demanding of the computer’s resources. During play in the high immersion condition, the frames per second (FPS) rate hovered around 60, while it could easily exceed 100 in the low immersion condition. Serendipitously, turning the VSync option on in *Minecraft*'s Video Settings will limit the maximum FPS to match monitor's refresh rate, which happened to also be 60 FPS. In this study, VSync was turned on, therefore participants in both conditions played Minecraft at 60 FPS, which only momentarily dipped in the high immersion condition to about 50 when a lot of new blocks were loaded at once. This frame rate is far above the typical rates found in television and motion pictures, which range between 24 to 30 FPS ("Frame Rate", n.d.). Also, given that Huynh-Thu and Ghanbari (2008) found that decreasing the frame rate of video from 25 to 12.5 did not seem to affect perceptions of video quality, it is highly unlikely that the drop from 60 to 50 FPS would be noticeable. Furthermore, this frame rate also is far above the value Barfield and Hendrix (1995, as cited in Slater & Wilbur, 1997) suggested was needed for presence. The difference in frame rates was negligible between conditions.

Video of gameplay was captured so that it could be reviewed in the event of otherwise unexplainable outliers (e.g., were settings changed, were there glitches?). The gameplay video was captured by feeding the video output from the computer through a 1x2 HDMI splitter. One of the outputs went directly to the monitor, while the other went to a Hauppage HD PVR Rocket (model 01530). The PVR Rocket recorded video of whatever was streaming through the HDMI cable directly to a 32 GB Corsair Flash Voyager GO USB stick (model CMFVG-32GB-NA). The use of the PVR Rocket did not have any effect on the game FPS because of the use of the splitter and the nature of the device itself.

Room set-up. Only one computer was used in this experiment. The monitor and speakers were positioned on the computer station desk and marked with tape so that the research assistants could ensure that they were in the same position for every participant. The computer and computer station were configured based on the participant's condition. Two curtains were installed on either side of the computer station. For those in the low immersion condition, the curtains were pushed to the back wall and the computer station was open, while those in the high immersion condition had the curtains closed around them. The curtains were used to limit the distracting external cues from the laboratory environment, which should increase immersion. See Figure 10 for what the computer station looked like with the curtains partially closed.

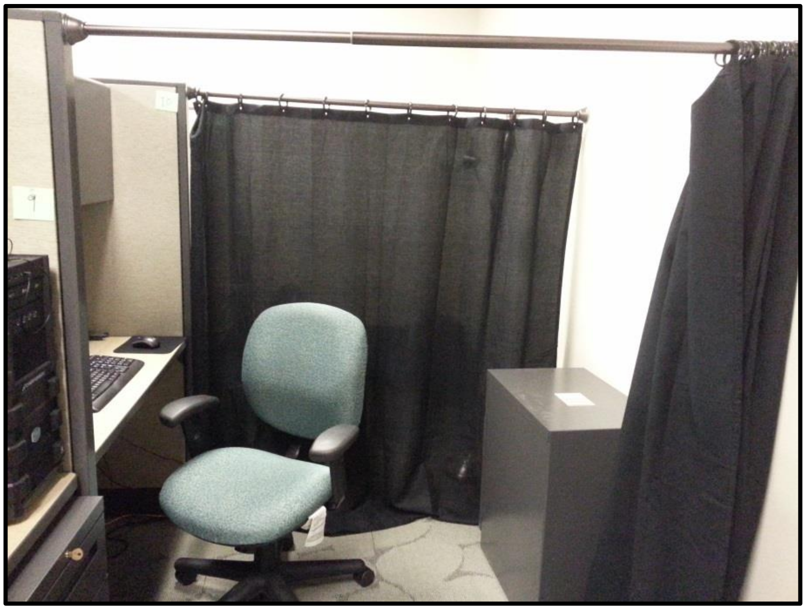


Figure 10. Computer station set-up with curtains partially closed.

Procedure

As stated previously, this study was completed in two stages: A baseline pre-screening that occurred online and an in-person experimental session. Based on the online pre-screening, 101 individuals were recruited for the in-person study. Only participants who were over the age of 18 were allowed to participate.

Online pre-screening. The baseline pre-screening was completed using an online survey system, Survey Monkey (http://www.surveymonkey.com). The participant received a copy of the informed consent document, which they were asked to read. They were not, however, asked to sign the informed consent document since this study qualified for a waiver of documentation of informed consent. Participants were required to click a button that indicated they have read and agreed to the terms listed in the informed consent document. Then, the participant completed the introductory surveys. This included the demographic surveys (i.e., basic demographic information, gaming preferences and behaviors, and VGSES) as well as the various measures for the individual differences hypothesized to influence game engagement (i.e., immersive/involvement tendencies, dispositional flow, willingness to suspend disbelief, creative imagination, and trait absorption). They also completed the *Minecraft* screening survey.

In-person experimental session. Participants that qualified for inclusion were invited to the laboratory to complete the experiment. Participants completed this study individually with a research assistant. This study was conducted in the University of Central Florida's Psychology building. The computer and computer station were configured for the randomly pre-assigned condition for the incoming participant before his or her arrival. When the participant reported to the time slot, they were asked to re-review the informed consent document. Given the nature of this study, both the informed consent document and the research assistant thoroughly explained the risk of simulator sickness to the participant. The participant was instructed to tell the research assistant if they started to feel dizzy or nauseated. If this occurred, the experiment would be immediately discontinued and the participant will be asked to remain seated until they felt better. None of the participants stopped the study due to simulator sickness.

The participant began by completing the first portion of the SSQ, which is meant to be given before virtual environment exposure. Then, the participant was told that they were going to play *Minecraft* in Survival Mode, and that they were not to change any settings. Next, the participant played *Minecraft*, which had been pre-configured based on their condition. The participant played *Minecraft* for 30 minutes. In their flow studies, Jackson and Eklund (2002) found that their participants spent roughly 25 minutes engaging in their physical activity of choice. While playing *Minecraft* is not similar to the physical activities in their original studies, this number can be used as a loose guideline for the minimum time necessary to experience flow. Additionally, Brockmyer et al.'s (2009) participants played the game for 30 minutes, which seemed adequate for their validation study. Seemingly few studies have examined the amount of time required to achieve a flow state while playing video games. Rau et al. (2006) found that 60 minutes of gameplay in an online RPG was optimal for experiencing flow. Still, it is likely not the amount of time spent in the game that determines when flow occurs, if it even occurs at all, but may be related to the pace of in-game opportunities to perform tasks necessary to achieve goals, which allows the player to make internal observations as to the perceived challenge/skill balance. As many genres feature similar gameplay pacing, an interesting future study would be to examine the average time to experience flow across different genres. Given that this research has yet to occur, the 30-minute playtime found in this study was based on the following considerations: (1) The 25-minute average time found by Jackson and Eklund (2002) in their flow validation studies could serve as an acceptable minimum because it was at least long enough to allow flow to occur in a highly-involved, active, potentially fast-paced setting; (2) Brockmyer et al. (2009) had participants play a game for 30 minutes in their validation study;

(3) since online RPGs have a more story-focused, slower pace, flow may occur sooner than 60 minutes in *Minecraft*; and, (4) this shorter time will hopefully minimize the incidence of simulator sickness.

After gameplay, the participant completed several different surveys which were used to quantify their subjective gameplay experience. These surveys measured the participant's perceptions of attentional requirements, immersion, involvement, presence, and flow. They also completed additional surveys that provided further insight into their experience, to include measures of usability, simulator sickness, and a second willingness to suspend disbelief measure targeted at the current experience. Finally, a measure of cognitive absorption was administered to reveal further information about the revised game engagement model on a conceptual level.

After completing the final round of surveys, the participant received a post-experiment information sheet that explained the purpose of the study and provided contact details for the researcher and for the UCF IRB office. After the participant left, the research assistant reset the room and saved a copy of the participant's *Minecraft* world. See Figure 11 below for a graphical representation of the overall study procedure.

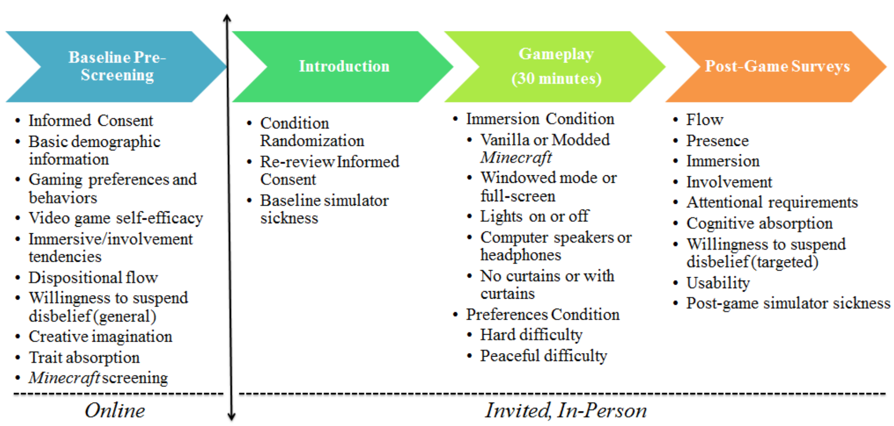


Figure 11. Study procedure outline.

Hypotheses and analysis plan

In summary, experienced *Minecraft* players were recruited for this study. Participants played *Minecraft,* either with strong immersive cues (high immersion) or with standard gameplay settings (low immersion). Whether flow is experienced was manipulated by changing the difficulty level to align with players' preferences. Therefore, there are two independent variables: immersion (low or high) and preferences (incongruent or congruent; see Table 3).

Table 3. Experimental condition description and expected game engagement states.

	Preferences	
Immersion	*Incongruent* Peaceful difficulty	*Congruent* Normal difficulty
Low • Vanilla *Minecraft* • Windowed mode • Computer speakers • Lights on, no curtain	*n* = 15 Will likely not experience presence or flow.	*n* = 15 Will likely experience flow, but not presence.
High • Modded *Minecraft* • Full screen mode • Headphones • Lights off, curtained area	*n* = 15 Will likely experience presence, but not flow.	*n* = 15 Will likely experience presence and flow.

While the independent variables are intended to encourage either immersion and presence or involvement and flow, it will not occur unless an individual similarly perceives these manipulations as either immersive or involving. Therefore, alongside condition, subjective ratings of both immersion and involvement were considered. The major dependent variables of interest are those stemming directly from the R-GEM (i.e., attentional requirements, immersion, involvement, presence, and flow), but several other variables were measured as the literature review has identified multiple points of conceptual overlap as well as several individual differences that influence the subjective experience (see Table 4).

Table 4. Measure scoring information.

Construct	Measure	Variables	Source
Demographic surveys			
Basic demographic information	RETRO Demographics Questionnaire	- age (years) - gender - ethnic background - year in school - primary language	N/A
Gaming preferences and behaviors	RETRO Demographics Questionnaire; Additional novel, targeted items	- computer comfort (1-7) - game comfort (1-7) - games played (hours / week) - all ways games are played - how the majority of games are played - top three favorite games (genres) - played a game with HMD (yes / no) - played *Minecraft* (yes / no) - hours playing *Minecraft* (hours total) - *Minecraft* with an HMD? (yes / no) - general game enjoyment (1-7) - *Minecraft* enjoyment (1-7) - *Minecraft* comfort (1-7)	N/A
VGSE	VGSES	- VGSE total score (10-40)	Pavlas et al., 2010
Minecraft Screening	*Minecraft* Screening	- all the ways they play *Minecraft* - how they prefer to play *Minecraft* - versions of *Minecraft* played - version of *Minecraft* they prefer to play - typical screen size - typical resolution - ratings of *Minecraft* modes - Survival Mode difficulty preference - patterns of *Minecraft* - *Minecraft* texture packs used - *Minecraft* mods used - *Minecraft* shaders used	N/A
Individual differences influencing game engagement			
Immersive/involvement tendencies	ITQ	- immersive/involvement tendencies total score (18-126)	Witmer & Singer, 1998
Dispositional flow	S DFS	- dispositional flow mean score (1-5)	Jackson et al., 2008
Willingness to suspend disbelief	N/A	- willingness to suspend disbelief item (1-7) - willingness to suspend disbelief mean (1-7)	Sas & O'Hare, 2004, some novel items
Creative imagination	CIS	- creative imagination *t*-score	Wilson & Barber, 1978
Trait absorption	TAS	- absorption total score (0-34)	Tellegen & Atkinson, 1974
Game engagement states			
Attentional requirements	PQ	- attentional requirements total score (7-70)	Witmer & Singer, 1998
Immersion	PQ	- immersion total score (7-112)	Witmer & Singer, 1998
Involvement	R PII; TAM; PQ	- involvement total score (21-105) - involvement categorization (low, medium, high)	Zaichkowsky, 1994; Davis, 1989; Witmer & Singer, 1998
Presence	Presence Equation	- presence count (0-9) - presence estimation - presence classification (yes / no)	Slater & Steed, 2000; Witmer et al., 2005
Flow	C FSS; FSS-2	- core flow mean score (1-5) - overall flow FFS-2 mean score (1-5) - challenge/skill balance FSS-2 mean (1-5) - merging of action & awareness FSS-2 mean (1-5) - clear goals FSS-2 mean (1-5) - unambiguous feedback FSS-2 mean (1-5) - concentration FSS-2 mean (1-5) - control FSS-2 mean (1-5) - loss of self-consciousness FSS-2 mean (1-5) - transformation of time FSS-2 mean (1-5) - autotelic experience FSS-2 mean (1-5)	Martin & Jackson, 2008; Jackson & Eklund, 2002
Model-validating variables			
Usability	SUS	- Usability percentile	Brooke, 1996
Simulator sickness	SSQ	- Simulator sickness total severity score	Kennedy et al., 1993
Cognitive absorption	CAS	- Mean cognitive absorption score (1-7) - Temporal dissociation CAS mean (1-7) - Focused immersion CAS mean (1-7) - Heightened enjoyment CAS mean (1-7) - Control CAS mean (1-7) - Curiosity CAS mean (1-7)	Agarwal & Karahanna, 2000

Analysis occurred in several stages, beginning with measurement evaluation to determine whether the proposed measures were valid, followed by manipulation checks to ensure that the imposed manipulations were effective. Then theoretical hypotheses derived from the eight assertions made about the R-GEM were evaluated, and, finally, additional analyses were conducted to address lingering questions about the game engagement construct with respect to prior work. Please note that the hypotheses listed correspond to the numbered list of assertions found in Table 2, and that they are all alternative hypotheses.

Measure evaluation. The first step was to ensure that the measures used were acceptable, especially since so many novel approaches were taken in this study. Three new items were created to augment Sas and O'Hare's (2004) single item measure of willingness to suspend disbelief. A very high correlation between these sets of items may be used to determine if these items are measuring similar constructs. Tabachnick and Fidell (2007) suggested that variables with correlations between $r = .70$ and $r = .90$ can be combined to create a composite score. If the correlation exceeds $r = .90$, however, these variables should not be combined as this may instead introduce inflated error. Essentially, rather than averaging out measurement error by combining two *related items* to assess a single construct, averaging two variables with correlations over $r = .90$ would instead be considered as including the *same item twice* and doubling the error. For this analysis, however, a less conservative correlation threshold is adopted of $r = .50$ because it is exploratory rather than a well-controlled study focused only on creating a new willingness to suspend disbelief measure. This is the same threshold value that Sas and O'Hare (2004) seemed to have used.

H0.1: The scores from the new willingness to suspend disbelief items are significantly, *strongly, and positively correlated (at least* r *= .50), but not redundant with (more than* r *= .90) the score from Sas and O'Hare's (2004) single item.*

Secondly, three novel subscales (immersion, involvement, and attentional requirements) were devised from the third version of the PQ (Witmer et al., 2005). Before these new subscales can be used, it is important to ensure that each set of items is measuring the proposed construct. The sample size was too low to conduct a formal CFA, which is a special case of SEM that examines correlational relationships. CFA is used to determine whether a proposed model fits a dataset's covariance. Instead, an Exploratory Factor Analysis (EFA) was used to determine if the items fell roughly into the proposed subscales.

H0.2: The items from the PQ are divided amongst the proposed subscales (immersion, involvement, and attentional requirements) as seen in an Exploratory Factor Analysis.

Additional items were added to the existing R PII (Zaichkowsky, 1994). As with the novel willingness to suspend disbelief items, it is of interest whether these items are complimentary to the existing ones. Therefore, a similar analysis was proposed:

H0.3: The scores from the TAM-influenced involvement items (Davis, 1989) as well as the PQ's involvement item (Witmer & Singer, 1998) are significantly, strongly, and positively correlated (at least r = .50), but not redundant with (more than r = .90) the total score from the Revised PII (Zaichkowsky, 1994).

Regarding presence measurement, Slater and Steed (2000) proposed that participants should estimate the amount of time they felt present in the VE, and if individuals reported percentages over 50%, they likely experienced presence. It was of interest if this estimation method aligned with the presence count data obtained in this study. This was examined in the high immersion and congruent preferences group. This group was selected because they should have had the richest gameplay experience and may have the most variance for all of the subjective state measures since all states were possible.

H0.4a: Presence count is significantly and positively correlated with the presence estimate in the high immersion and congruent preferences group.

H0.4b: Presence classification is significantly and positively predicted by presence count in the high immersion and congruent preferences group.

Regarding the multiple flow scales used in this study, the C FSS mean score and the overall FSS-2 mean score should be very highly correlated as they are both purported to measure general flow state experience. If this was the case, then the C FSS would be used for the measure of overall flow, while the individual flow dimension scales of the FSS-2 would be used to examine different aspects of the R-GEM. Jackson et al. (2010b) stated that the overall FSS-2 mean score was not recommended to measure the holistic experience of flow and that the dimension subscales should be used when possible. Fortunately, the C FSS was designed with this purpose in mind. The dimension subscales of the FSS-2, however, are still very valuable when examining different aspects of the flow experience. Therefore, it is important to establish the link between the C FSS and the FSS-2 in this study so that both scales can be used to measure the same construct of flow, although different dimensions will be examined across the analyses. A correlation threshold of $r = .70$ is adopted based on the recommendation from Tabachnick and Fidell (2007) that they may be combined into a composite score.

H0.5: The C FSS mean score and the overall FSS-2 mean score is significantly, strongly, and positively correlated ($r > .70$).

Manipulation check. Once the measures were established as useful, the next step was to ensure that the manipulations were effective. Those in the immersion condition, regardless of congruency with gameplay difficulty preference, were hypothesized to have higher immersion scores due to the presence of immersive cues (e.g., the wider field of view, more detailed textures and environments, etc.). Similarly, those who played at their desired difficulty level, regardless of immersion condition, were hypothesized to have had higher involvement challenge/skill balance scores.

H1: Ratings of immersion significantly increase when more immersive cues are present.

H2: Those in the congruent preferences group have significantly higher ratings of challenge/skill balance.

Even if the experimental manipulations did not work, it is important to remember that the R-GEM is based on subjective evaluation. Like Singer and Witmer (1999) explicitly stated: It may not matter if a VE has the most immersive technology in the world—if the individual does not perceive it as immersive, it will not be immersive to them. Therefore, if these manipulations have failed, the subjective scores for the different game engagement states can still be used to evaluate the relationships of the constructs in the R-GEM.

Theoretical hypotheses. The following measures were to be used for the rest of the proposed analyses:

- *Attentional requirements*: Attentional requirements total score from the revised PQ subscale (version 3; Witmer et al., 2005)
- *Immersion*: Immersion total score from the revised PQ subscale (Witmer et al., 2005)

- *Involvement*: Involvement total score from the three scales (Zaichkowsky, 1994; Davis et al., 1989; Witmer et al., 2005)
- *Presence*: Presence count (Slater & Steed, 2000; Witmer et al., 2005)
- *Flow*: C FSS mean score (Martin & Jackson, 2008)

The first aspect of the revised game engagement model evaluated was the importance of the attentional pre-requisites. Internal distractions can be quantified through usability, in that the more usable the individual perceived the game, the less distracted the player would be while playing. Additionally, external distractions can be minimized through the use of noise-cancelling head phones and a curtained-off area, which those in the high immersion condition used during gameplay. Therefore, it is hypothesized that:

H3a: Usability and whether the individual plays with external distractions blocked out significantly and positively predict the attentional requirements score.

Without focused attention and minimized internal and external distractions, it will be difficult to become either immersed or involved. For this next analysis, a low-level engagement score was created by averaging the immersion and involvement total scores. Then, it was examined whether attentional requirements score alone, or with the added aspects of usability and minimized distractions significantly predict low-level engagement scores. To examine the critical influence of the predictors in the model, the partial correlation coefficient was examined.

H3b: The attentional requirements score significantly and positively predicts low-level engagement scores.

As suggested by Witmer and Singer (1998), immersion and involvement were hypothesized to have a reciprocal relationship.

H4: The immersion total score and involvement total score are significantly and positively correlated.

Presence and flow both stem from low-level game engagement, however it was hypothesized that immersion is more influential in experiencing presence while involvement is more influential in flow. This can be examined using regression, where the relative contributions of immersion and involvement for predicting both presence and flow can be evaluated through the semi-partial and partial correlation coefficients, which examine both total and unique contribution of variance explained, respectively. Consideration of condition was not included in these analyses because conditions do not necessarily matter as long as a range of possible states are experienced:

H5a: The attentional requirements score, immersion total score, involvement total score, and the challenge/skill balance score are all significantly and positively predictive of the CORE flow mean score, however involvement is more influential than immersion.

H5b: The attentional requirements score, immersion score, and involvement score are all significantly and positively predictive of the presence count, however immersion is more influential than involvement. The challenge/skill balance score, while also included, is not a significant predictor of presence count.

It is hypothesized that presence occurs when an individual is highly immersed, or when they have high levels of individual differences such as willingness to suspend disbelief, creative imagination, or immersive tendencies to mitigate the immersive cue deficit. Level of experience should not matter, although subjective scores of involvement might given involvement's reciprocal relationship with immersion. Therefore, part and partial correlation coefficients were examined to see if the proposed individual differences explained overlapping or unique amounts of explained variance.

H6. The total presence count is significantly and positively predicted by high immersion total score, the presence of immersive cues, and high levels of individual differences (to include willingness to suspend disbelief, creative imagination, and immersive/ involvement tendencies). The level of challenge/skill balance does not matter, but involvement may.

Flow is hypothesized to occur when the individual is highly involved, likely due to their own level of motivation based on experience and/or the presence of goals and feedback, *and* there is balance struck between their skill level and the challenge provided by the game. This may be influenced by a number of individual differences, to include willingness to suspend disbelief, creative imagination, immersive/involvement tendencies, dispositional flow tendencies, trait absorption, age, and VGSE. Whether immersive cues were provided should not matter, although immersion may be somewhat influential given its reciprocal relationship with involvement and the conceptual overlap that exists between immersion and flow. Again, part and partial correlation coefficients were examined to see if the proposed individual differences explain overlapping or unique amounts of variance.

H7: The flow score is significantly and positively predicted by high involvement total score, playing on the congruent game difficulty, the presence of high flow mechanisms scores (feedback, goals, challenge/skill balance), and high levels of individual differences (to include willingness to suspend disbelief, creative imagination, and immersive/ involvement tendencies). The level of presence of immersive cues does not matter, but the *immersion score itself may.*

Finally, presence and flow are not mutually exclusive subjective states. It is possible to be both present and in flow, as well as only present or only in flow. It is also possible to not experience either presence or flow. This will depend on the inclusion of the mechanisms that promote either experience, as well as the individual player's subjective appraisal of those imposed mechanisms. This was examined by correlating presence count and flow score between the different conditions. This is similar to the approach taken by Murray, Fox, and Pettifer (2007, as cited in Brockmyer et al., 2009), who found that presence and flow were not correlated. Only in the high immersion and congruent preferences group would we expect to see a positive correlation between presence count and core flow mean score, as both their requirements are present and because correlation does not imply causation, but association. In all other conditions, correlations were not expected.

H8: Only in the instance where game feature mechanisms support both presence and flow are these states significantly associated, otherwise there is no significant association between presence and flow.

Additional questions. There is one final question raised in the literature review about Brockmyer et al.'s (2009) original game engagement model that can be addressed with the current study: is there a difference between the subjective states of flow and cognitive absorption? Recall that Brockmyer et al. (2009) proposed that cognitive absorption (Agarwal & Karahanna, 2000) and flow (Csikszentmihalyi, 1975a, 1975b, 1990) were both mind-altering, but opposing subjective states. They proposed that which state an individual experienced (i.e., either cognitive absorption or flow) was based on the individual's motivation and affect. This notion was debunked in a previous section as flow can involve both negative affect and extrinsic motivations, to an extent. Furthermore, a closer inspection of Agarwal and Karahanna's (2000) original theory revealed that cognitive absorption is grounded almost entirely in Csikszentmihalyi's (1975a, 1975b, 1990) flow theory, to the point that some items of the CAS were adapted from preexisting flow measures. In the CAS, the only distinct features are an additional immersion subscale as well as a subscale pertaining to curiosity, which also happened to be derived from a flow measure (Agarwal & Karahanna, 2000). This was examined by correlating scores from the theoretically-identical subscales as well as correlating the total flow and cognitive absorption scores. It was hypothesized that these correlations will be $r = .50$ or greater.

H9a: The mean cognitive absorption score and the CORE flow mean score is significantly, strongly, and positively correlated (at least r = .50).

H9b: The Temporal Dissociation CAS mean and the Transformation of Time FSS-2 mean is significantly, strongly, and positively correlated (at least r = .50).

H9c: The Control CAS mean and the Control FSS-2 mean is significantly, strongly, and positively correlated (at least r = .50).

H9d: The Heightened Enjoyment CAS mean and the Autotelic Experience FSS-2 mean is significantly, strongly, and positively correlated (at least r = .50).

CHAPTER FOUR: RESULTS

Initially, 1,210 undergraduate students completed the online screening portion of this study. Of those, 358 indicated that they had played *Minecraft* at least once. One-hundred and eighty one of those individuals were invited to participate in the in-person experimental session. Each of these invited participants was classified as a Novice, a Player, or an Expert. Novices reported playing less than 10 hours total of *Minecraft*, while Players reported playing between 10 and 29 hours. Participants were classified as an Expert if they: have played 30 or more hours of *Minecraft*; played the majority of *Minecraft* on PC; played *Minecraft* on a difficulty of Normal or higher; and, rated their overall enjoyment of *Minecraft* and of Survival Mode a 6 or higher out of 7. If they failed to meet any of these criteria, but still reported playing 30 or more hours of *Minecraft*, they were classified as a Player. See Table 5 below for the classification checklist.

Table 5. Minecraft classification checklist.

***Minecraft* Novice**	***Minecraft* Player**	***Minecraft* Expert**
❑ Played *Minecraft* less than 10 hours	❑ Played *Minecraft* between 10 – 29 hours ❑ Played *Minecraft* for 30 or more hours, but did not qualify as an Expert	❑ Played Minecraft 30 or more hours ❑ Played the majority of *Minecraft* on PC ❑ Played *Minecraft* on a difficulty of Normal or higher ❑ Rated their overall enjoyment of *Minecraft* a 6 or higher out of 7 ❑ Rated their overall enjoyment of Survival Mode in *Minecraft* a 6 or higher out of 7

These criteria were chosen because the study was conducted on computers, therefore players must be comfortable with the PC *Minecraft* controls to minimize potential distractions generated from unfamiliarity. Also, the experimental manipulation for flow required that participants were comfortable playing on the Normal difficulty. Additionally, thirty hours of gameplay was selected because it was a number of hours that was more than a trivial amount of gameplay. This was based on an analysis of early data, where this point emerged as a dividing line between those who may have tried playing *Minecraft* (e.g., those who reported playing 2, 8, 10 hours) and those who play it avidly (e.g., those who reported playing 30, 50, 100 hours).

One-hundred and one participants completed the in-person study. Of these, 3 never played Minecraft and were recruited for the pilot study. In total, there were 5 Novices, 41 Players, and 52 Experts. Only 93 Players and Experts were retained for final analysis reported here. There were four conditions based on the manipulations of immersion (low / high) and game difficulty (peaceful / normal). The numbers of Players and Experts by condition is reported in Table 6.

Table 6. Number of Players and Experts by condition.

	Low Immersion		**High Immersion**	
Difficulty	*Peaceful*	*Normal*	*Peaceful*	*Normal*
Player	10	13	9	9
Expert	11	11	15	15
Total	21	24	24	24

Data cleaning and screening

First, the data were screened for any errors, missing information, and outliers. Data were collected digitally, which eliminated input errors aside from responses requiring text field entry (e.g., age). For example, one participant reported that they were 9 years old. This was assumed to be an input error and that the individual was instead 19, especially since all participants had to certify that they were at least 18 years old to participate in the study.

The dataset was scanned for missing data. For constructs that relied on subscale calculation, any missing data were addressed using imputed means. This is a conservative approach that does not change the distribution and does not require guessing on the part of the experimenter, although one drawback is that the overall variance is reduced (Tabachnick & Fidell, 2007). This is only a problem when there is a large amount of missing data, which was not the case in this dataset. For missing data points, these imputed means were calculated for each condition rather than the overall mean.

One participant (#18) did not complete 40% of the CAS items. Since most other cases were instances of a single missing data point and because this error was unique, imputed means were not used for this participant for these CAS items. For other variables, missing values were not imputed. For example, any variable relating to presence was not imputed because it is unknown whether a non-response indicated no presence experienced or if it was accidentally skipped. Means were not imputed for the portion of the SSQ asking the participant to report the number of days since they participated in a given experience, again not knowing whether a non-response meant 0 days or if the question was accidentally skipped. Means were not imputed for free-response variables (e.g., screen resolution).

There were 340 variables collected for each of the 101 participants in the full dataset. Out of these, means were not imputed for 34 variables. Therefore, a total of 30,906 data points for 306 variables and 101 participants were assessed for missing data. Of these, 41 data points were missing (0.13% of all data) and the missing data were randomly distributed with the exception of the aforementioned CAS items for a single participant.

Witmer and Singer (1998) found that ratings of simulator sickness were negatively correlated with PQ scores. Given that the experience of simulator sickness should be distracting enough to cause disengagement with the game, those who experienced high levels of simulator sickness should be removed from the analysis. In their study, Balk et al. (2013) removed participants from their analysis whose SSQ scores were above 55 as they believed these participants decided to fight through the symptoms to complete the experiment. A similar approach was taken in this study, which led to the removal of Participants #70 and #100.

The full list of variables examined for outliers and for normality is found in Appendix F. Variables were converted to standardized scores and then placed in ascending order. Outliers were identified as those cases where scores on any of the examined variables exceeded +/- 3.29 (Tabachnick & Fidell, 2007). Ten outliers were identified across six participants. To address these outliers, the scores were changed so that they were one measurement unit beyond the case closest to the cut-off without exceeding it. Addressing outliers in this way helps to meet statistical analysis assumptions, while still retaining each case's deviance (Tabachnick & Fidell, 2007). Standardized scores were calculated after this adjustment to ensure that these individual cases were no longer outliers. See Table 7 for more information.

Table 7. Outlier adjustment information.

Variable	**#**	**Original Score**	**Adjusted Score**
SUS_Percentile	96	6.00	25.0
Usability	96	42.5	57.0
Flow_Concentration	32	2.25	2.75
CAS_Immersion	32	3.00	4.00
Flow_Control	53	2.75	3.25
VGSE	53	21.0	24.0
CAS_Curiosity	15	2.00	2.66
Flow_Time	15	1.25	1.75
CAS_Enjoyment	21	4.25	4.75
CAS_Enjoyment	29	4.25	4.75

Mahalanobis D^2 was calculated to identify multivariate outliers, but none were found. Skew and kurtosis for all of the variables listed in Appendix F were calculated by dividing the skew and kurtosis statistics by their respective standard errors. Values found to be +/- 3.3 are considered skewed or kurtotic (Tabachnick & Fidell, 2007). See Table 8 for details about this analysis.

Table 8. Skew and kurtosis analysis.

Variable	**Original Skew**	**Transformation**	**Highest Possible Value**	**Skew After Transform**
SSQ_Post_TS	7.298046665	Logarithmic	N/A	2.393675473
SUS_Percentile	-7.640638074	Reflect & Logarithmic	99.99	0.470774013
Flow_Merging	-3.715512111	Reflect & Logarithmic	5	1.849321193
Flow_Goals	-4.970331729	Reflect & Logarithmic	5	1.035830772
Flow_LossofSC	-3.958092609	Reflect & Logarithmic	5	0.935763085
Flow_Time	-4.191585477	Reflect & Logarithmic	5	1.80656132
CAS_Control	-4.68754851	Reflect & Logarithmic	7	1.43524268
CAS_Curiosity	-4.458653624	Reflect & Logarithmic	7	1.332263019
Variable	**Original Kurtosis**	**Transformation**	**Highest Value**	**Kurtosis After Transformation**
SSQ_Post_TS	6.683895516	Logarithmic	N/A	-2.29914773
SUS_Percentile	6.538846127	Logarithmic	N/A	-1.49714722

Of the affected variables, most were negatively skewed. This means there was a spike toward the high values and a long tail toward the negative values. The only variable that was positively skewed was the SSQ's Total Severity Score. This is expected, as simulator sickness should be very low in this population since these participants routinely play *Minecraft* and would not play as much if it usually made them sick. Transformation can be used to address these skewed and kurtotic variables. Transformation may improve the normality of the variable's distributions in that outliers will still be found at the edges of the distribution, but they will not violate assumptions of tests if the transformation successfully lowers the skew/kurtosis (Tabachnick & Fidell, 2007).

Due to the substantial amount of skew and kurtosis found, the logarithm transformation was selected. For those that were negatively skewed, the scores were reflected. After transformation, the skew and kurtosis of the affected variables was acceptable.

Based on this data cleaning and screening exercise, as well as examining experimenter logs and footage captured of gameplay, nine participants were removed from the study for the following reasons:

- #23: The wrong resource pack was used (experimenter error).
- #26: Participant had played with shaders in the past and was disqualified in case that prior experience affected game engagement scores.
- #37: Participant had played with shaders in the past and was disqualified in case that prior experience affected game engagement scores.
- #70: Experienced an amount of simulator sickness beyond the threshold and was removed in case the discomfort affected game engagement scores.
- #76: Participant failed to comply with the experimental protocol and muted the game. They played with the game muted for the entirety of the session.
- #77: The game was still muted when the participant completed the session (experimenter error).
- #78: The game was still muted when the participant completed the session (experimenter error).
- #79: Participant failed to comply with the experimental protocol and changed the difficulty half-way through the experimental session.

- #100: Experienced an amount of simulator sickness beyond the threshold and was removed in case the discomfort affected game engagement scores.

A total of 84 participants were retained in the final analysis. The final numbers of participants in each condition are listed below in Table 9.

Table 9. Number of Players and Experts by condition in the final dataset.

	Low Immersion		**High Immersion**	
Difficulty	*Peaceful*	*Normal*	*Peaceful*	*Normal*
Player	9	13	9	8
Expert	9	10	12	14
Total	18	23	21	22

Finally, a series of one-way ANOVAs was used to determine if there were any differences on baseline quantitative demographic variables between the four conditions. None of the tests were statistically significant. The means for each variable, regardless of condition, are found below (Table 10).

Table 10. Means and standard deviations of demographic variables (N = 84).

Variable	**Mean**	**Standard Deviation**	**Range**
Age	19.02	1.44	18 - 24
Hours Played of *Minecraft*	142.38	250.94	10 – 2,000
Enjoyment Rating of *Minecraft*	6.43	0.84	4 - 7
Rating of Minecraft's Survival Mode	6.49	0.87	4 - 7
Creative Imagination Scale *t*-score	52.89	10.40	27 - 83
Immersive / Involvement Tendencies	89.47	14.54	52.00 - 126.00
Dispositional Flow	4.13	0.55	2.56 - 5.00
General Willingness to Suspend Disbelief	5.58	1.16	2.00 - 7.00
Trait Absorption	20.58	7.05	3.00 - 34.00
Video Game Self-Efficacy	35.55	4.22	24.00 - 40.00

Two chi-square tests also were conducted for the categorical demographic variables (gender and expertise). These tests were not found to be statistically significant. There were 66 males in who participated in the in-person experimental session (78.6%). There also were 45 participants classified as Experts in this study (53.6%) and 39 Players (46.4%). Overall, the results of this initial analysis suggested that there were no baseline demographic differences between conditions.

The following analyses were conducted using IBM SPSS Statistics version 20, and were conducted at the .05 alpha level unless otherwise stated. Also, again unless otherwise stated, assumptions for each statistical test conducted was evaluated and satisfactorily met (see Appendix G).

Measure evaluation

The first set of hypotheses examined whether the many novel approaches for the subjective assessment of gameplay engagement used in this study were valid.

Hypothesis 0.1: *The scores from the new willingness to suspend disbelief items are significantly, strongly, and positively correlated (at least* r *= .50), but not redundant with (more than* r *= .90) the score from Sas and O'Hare's (2004) single item.*

Pearson's product moment correlation coefficient was computed between the three new willingness to suspend disbelief items and Sas and O'Hare's (2004) single item for the same construct. Since this was a directional hypothesis, it was a one-tailed test. See Table 11 below for the correlations. The hypothesis was only partially supported because while several significant, positive correlations were found, they did not all exceed the r = .50 threshold. Since this variable is not one of the four key game engagement states, this is acceptable and the mean of the scores were used to represent participants' willingness to suspend disbelief while playing *Minecraft* in this experiment.

Table 11. Correlations between Sas and O'Hare's (2004) willingness to suspend disbelief item with three new potential items.

	Sas & O'Hare Item	***New Willing Item 1***	***New Willing Item 2***
New Willing Item 1	-.187*		
New Willing Item 2	-.405**	.254**	
New Willing Item 3	.468**	-.011	-.237*

Notes: $*p < .05$ (one-tailed); $**p < .01$ (one-tailed); Items in italics are reverse-coded; $N = 84$, $df = 82$.

Hypothesis 0.2: *The items from the PQ are divided amongst the proposed subscales (immersion, involvement, and attentional requirements) as seen in an Exploratory Factor Analysis.*

An exploratory factor analysis (EFA) was conducted in order to determine the underlying factor structure of the PQ. The data were examined for the presence of outliers as well as the absence of multicolinearity, both of which were absent. Several correlations were above .30, suggesting that the correlation matrices were factorable. The sample size for an EFA was quite low, given that there were 27 items and Costello and Osborne (2005) recommended that should be at least 20 participants for each item for the most accurate results, which would mean the sample size should be $N = 540$, rather than $N = 84$. Since this is just an initial look at the factor structure of the PQ to determine if the proposed subscales can be used in the experimental analyses, the lack of adequate sample size is acceptable for this purpose.

Since the components underlying the questions were thought to have some overlap given their interrelated nature, principal axis factoring was used for extraction. Truncation made use of the Kaiser criterion (Eigenvalue > 1). Equamax rotation was applied to optimize the number of variables that loaded on each factor and to also minimize the number of factors. See Table 12 for the inter-correlations among study variables, which was composed of items selected the third version of the PQ, renamed to reflect proposed subscales (Witmer et al., 2005).

Table 12. Correlation matrix for exploratory factor analysis.

	Imm_01	Imm_02	Imm_03	Imm_04	Imm_05	Imm_06	Imm_07	Imm_08	PreReq_01	PreReq_02	PreReq_03	PreReq_04	PreReq_05	Imm_09	Imm_10	Imm_11	Imm_12	Imm_13	Imm_14	Imm_15	Imm_16	PreReq_06	PreReq_07	PreReq_08	PreReq_09	PreReq_10
Imm_02	.432**																									
Imm_03	.383**	.488**																								
Imm_04	.173	.415**	.192*																							
Imm_05	.387**	.342**	.531**	.256**																						
Imm_06	.183*	.334**	.503**	.244*	.532**																					
Imm_07	.396**	.329**	.535**	.265**	.627**	.537**																				
Imm_08	.550**	.309**	.239*	.182*	.266**	-.027	.241*																			
PreReq_01	-.026	.049	-.021	.110	-.065	.018	-.114	.160																		
PreReq_02	-.035	-.056	.184*	-.193*	.052	.188*	.096	-.148	.213*																	
PreReq_03	-.145	-.130	.090	-.051	-.024	.059	.063	-.208*	.018	.481**																
PreReq_04	.479**	.389**	.230*	.216*	.493**	.248*	.380**	.539**	.141	-.099	-.208*															
PreReq_05	.354**	.527**	.253*	.350**	.331**	.205*	.316**	.502**	-.024	-.232*	-.261**	.443**														
Imm_09	.227*	.308**	.233*	.132	.182*	.073	-.042	.335**	.113	-.136	-.244*	.203*	.335**													
Imm_10	.218*	.221*	.406**	.147	.414**	.250*	.331**	.231*	.145	.024	-.134	.223*	.166	.470**												
Imm_11	.294**	.292**	.301**	.074	.317**	.146	.212*	.210*	.170	.004	-.066	.203*	.181*	.486**	.493**											
Imm_12	.073	.338**	.465**	.276**	.322**	.312**	.349**	.060	.168	.172	.102	.092	.088	.149	.316**	.342**										
Imm_13	.250*	.186*	.149	.245*	.392**	.140	.194*	.346**	-.058	-.205*	-.106	.301**	.163	.369**	.272**	.095	.225*									
Imm_14	.057	.131	.101	.032	.219*	.135	.149	.100	.014	-.020	.053	.064	.097	.185*	.116	.170	.248*	.445**								
Imm_15	-.082	.268**	.338**	.219*	.190*	.238*	.285**	-.010	-.037	.167	.073	.025	.128	.137	.242*	.198*	.395**	.142	.173							
Imm_16	.009	.232*	.237*	.272**	.235*	.290**	.308**	.032	.045	.009	.107	.074	.034	.264**	.334**	.310**	.451**	.248*	.291**	.481**						
PreReq_06	.150	.143	-.066	.251*	.161	-.057	.119	.194*	-.086	-.434**	-.224*	.177	.203*	.234*	.069	.181	-.029	.258**	.133	.131	.199*					
PreReq_07	.193*	.341**	.029	.160	.279**	.143	.122	.308**	.115	-.299**	-.436**	.385**	.446**	.488**	.338**	.266**	.100	.326**	.110	.103	.129	.327**				
PreReq_08	-.106	-.013	-.061	-.045	-.069	.079	-.007	-.159	.420**	.089	.179	-.101	-.234*	-.161	.034	.012	.089	-.060	.261**	-.081	.092	.076	-.092			
PreReq_09	-.200*	-.150	.031	-.061	-.106	.141	-.081	-.169	.225*	.312**	.362**	-.334**	-.235*	.026	.024	.043	.045	.035	.131	.061	.153	-.180	-.275**	.269**		
PreReq_10	.206*	.241*	.011	.317**	.165	-.062	.149	.352**	-.036	-.414**	-.235*	.387**	.316**	.221*	.265**	.246*	.190*	.320**	.132	.069	.170	.482**	.284**	-.058	-.107	
Invol_15	-.003	.404**	.300**	.164	.166	.210*	.189*	.076	.103	-.061	-.028	.231*	.372**	.221*	.187*	.370**	.397**	.007	.104	.378**	.332**	.098	.192*	.060	-.021	.336**

Notes: *p < .05 (one-tailed); **p < .01 (one-tailed).

From the set of twenty-seven variables, eight factors emerged that together accounted for 68.17% of the variance. Examination of the Scree plot instead suggested the presence of three factors in the solution since the line flattens at four factors, which accounts for 43.47% of the variance. The purpose of this analysis was to determine if the PQ's items could be used to create the proposed, theoretically-driven subscales. See Table 13 for the rotated factor matrix with values below .40 suppressed, where the PQ items have been renamed to represent their proposed subscales.

Table 13. Rotated factor matrix for exploratory factor analysis.

	PQ Item	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Immersion 1	2		.651						
Immersion 2	3			.501					
Immersion 3	4	.507							
Immersion 4	6								
Immersion 5	7	.699							
Immersion 6	8	.656							
Immersion 7	14	.754							
Immersion 8	9		.755						
Immersion 9	10				.600	.411			
Immersion 10	15				.612				
Immersion 11	16				.723				
Immersion 12	5			.484					
Immersion 13	11							.834	
Immersion 14	12							.484	
Immersion 15	25			.555					
Immersion 16	32			.434					
Pre Req 1	22								.726
Pre Req 2	23						-.619		
Pre Req 3	19					-.564			
Pre Req 4	20		.566						
Pre Req 5	21		.490			.422			
Pre Req 6	24						.606		
Pre Req 7	17					.704			
Pre Req 8	26								.703
Pre Req 9	27								
Pre Req 10	31						.644		
Involvement	N/A			.631					

Note: PQ Item refers to the item numbering found in Witmer et al.'s (2005) version 3 of the PQ (Table 1, p. 302).

The factors were somewhat separated as hypothesized. For example, only immersion items loaded onto the first factor. The second factor was split between immersion and attentional requirements items. After that, the attentional requirements items largely loaded onto their own factors. The lone involvement item loaded onto a factor with immersion items, which makes sense given that the two states should be highly correlated. These factors were fragmented, but that may have been due to the nature of the rotation. These findings are somewhat supportive of the hypothesis, and the proposed subscales were used in the subsequent experimental analyses.

Hypothesis 0.3: *The scores from the TAM-influenced involvement items (Davis, 1989) as well as the PQ's involvement item (Witmer & Singer, 1998) are significantly, strongly, and positively correlated (at least r = .50), but not redundant with (more than r = .90) the total score from the Revised PII (Zaichkowsky, 1994).*

Pearson's product moment correlation coefficient was computed between the four TAM-influenced involvement items, the PQ's involvement item, and the total score from the Revised PII. Since this was a directional hypothesis, it was a one-tailed test. See Table 14 below for the correlations.

Table 14. Correlations among involvement items.

	Revised PII Total	***TAM Item 1***	**TAM Item 2**	**TAM Item 3**	***TAM Item 4***
TAM Item 1	-.321**				
TAM Item 2	.582**	-.369**			
TAM Item 3	.507**	-.509**	.626**		
TAM Item 4	-.126	.775**	-.258**	-.467**	
PQ Item	.500**	-.614**	.626**	.655**	-.407**

Note: *$p < .05$ (one-tailed); **$p < .01$ (one-tailed); Items in italics are reverse-coded; $N = 84$, $df = 82$.

Out of the fifteen correlations, nine of them were over $r = .50$ (60% of the items). The average of these correlations is .489. By examining Table 14, it appears that TAM Item 4 is the weakest. If an average of the correlations is found after removing TAM Item 4, the average is raised to .520. Still, even without dropping the weak item, this is adequate for the purposes of this study. A score for involvement was created by totaling all ten of the Revised PII items, the four new TAM items, and the lone PQ Item. The hypothesis was supported.

Hypothesis 0.4a: *Presence count is significantly and positively correlated with the presence estimate in the high immersion and congruent preferences group.*

Pearson's product moment correlation coefficient was computed between the presence count and the participant's presence estimate. This correlation was conducted in both the high immersion and congruent preferences group as well as in the full sample, especially since the sample size was too low to conduct this test without error just in the one condition. Since this was a directional hypothesis, it was a one-tailed test. In the high immersion and congruent preferences group, a significant and positive correlation was found, $r(18) = .891$, $p < .001$, $r^2 = .794$ and the effect was strong. In the full sample, a less strong, but still significant correlation was found, $r(77) = .749$, $p < .001$, $r^2 = .561$. The hypothesis was fully supported.

Hypothesis 0.4b: *Presence classification is significantly and positively predicted by presence count in the high immersion and congruent preferences group.*

A logistic regression was performed to determine whether presence count affected the likelihood that participants would have been classified as having experienced presence. This classification was based on whether they reported feeling present over 50% of the time during gameplay. This analysis was conducted in both the high immersion and congruent preferences group as well as in the full sample. As with Hypothesis 0.4a, the sample size was too low to conduct this test without error in only the high immersion and congruent preferences group.

In the high immersion and congruent preferences group, the logistic regression model was statistically significant, $\chi^2(1) = 9.221$, $p = .002$. The model explained between 34.2% and 45.8% of the variance in presence classification, as suggested by the Cox & Snell and Nagelkerke R^2, respectively. Furthermore, 86.4% of the cases were correctly classified. When the presence count score is raised by one point, the individual was 1.532 times more likely to experience presence. In the full sample, the logistic regression model was still statistically significant, $\chi^2(1) = 34.533$, $p < .001$. The model explained between 33.7% and 45.7% of the variance in presence classification, as suggested by the Cox & Snell and Nagelkerke R^2, respectively. Furthermore, 82.1% of the cases were correctly classified. When the presence count score is raised by one point, the individual was 1.646 times more likely to experience presence. The hypothesis was fully supported.

Hypothesis 0.5: *The C FSS mean score and the overall FSS-2 mean score is significantly, strongly, and positively correlated (r > .70).*

Pearson's product moment correlation coefficient was computed between the C FSS mean score and the FSS-2 mean score to determine whether the two flow scales were measuring the same construct. Since this was a directional hypothesis, it was a one-tailed test. A significant and positive correlation was found, $r(82) = .73$, $p < .001$, $r^2 = .53$, a strong effect. This suggests that the C FSS mean score can be used to represent overall flow, while the subscales of the FSS-2 can be used in conjunction with it to explore flow mechanisms. The hypothesis was supported.

Based on the results of the first hypotheses tested pertaining to measure evaluation, this study used the proposed new measures and subscales to conduct the experimental analysis.

Manipulation check

The next set of analyses focused on whether the experimental manipulations resulted in the hypothesized game engagement states. Specifically, those in the high immersion condition were hypothesized to have higher immersion scores due to the presence of immersive cues regardless of whether they were playing on their favored difficulty setting. Higher immersion should result in more presence experienced. Similarly, regardless of immersion condition, those who played on their favored difficulty setting were hypothesized to have higher ratings of challenge/skill balance. This is important as challenge/skill balance is a key mechanism of flow.

Hypothesis 1: *Ratings of immersion significantly increase when more immersive cues are present.*

A 2x2 factorial between-subjects ANOVA was conducted to determine how the immersion manipulation as well as the difficulty manipulation affected immersion scores. It was expected that there would be a main effect for the immersion manipulation, but not for the difficulty manipulation. An interaction effect would be indicative of an experimental confound and was not expected. Neither of the main effects were significant, and nor was the interaction effect, was significant (see Table 15). The experimental manipulations did not affect how much immersion was experienced. Overall, the average immersion experienced was 93.09 points (*SD* = 10.47) out of 112 possible points (see Table 16). The hypothesis was not supported.

Table 15. Analysis of immersion scores based on condition.

	F	***p*-level**	**Partial η^2**	**Power**
Immersion Condition (*df* = 1)	0.101	.751	.001	.061
Difficulty Condition (*df* = 1)	0.077	.782	.001	.059
Interaction (*df* = 1)	1.420	.237	.017	.218

Note: Total *df* = 84; Error *df* = 80.

Table 16. Means and standard deviations of immersion score by condition.

	Low Immersion	**High Immersion**	**Total**
Peaceful Difficulty	94.50 (6.84)	92.48 (12.06)	93.41 (9.93)
Normal Difficulty	91.10 (13.17)	94.60 (8.15)	92.81 (11.02)
Total	92.59 (10.87)	93.56 (10.18)	93.09 (10.47)

Note: Standard deviations are listed in parentheses.

Hypothesis 2: *Those in the congruent preferences group have significantly higher ratings of challenge/skill balance.*

A 2x2 factorial between-subjects ANOVA was conducted to determine how the immersion manipulation as well as the difficulty manipulation affected the challenge/skill balance scores. It was expected that would be a main effect for the difficulty manipulation, but not for the immersion manipulation. Again, an interaction effect would be indicative of an experimental confound and was not expected. See Table 17 for the results of the analysis.

Table 17. Analysis of challenge/skill balance scores based on condition.

	F	p-level	Partial η^2	Power
Immersion Condition (df = 1)	0.253	.616	.003	.079
Difficulty Condition (df = 1)	5.685	.019	.066	.654
Interaction (df = 1)	0.174	.678	.002	.070

Note: Total *df* = 84; Error *df* = 80.

The main effect for the immersion condition was not significant, and neither was the interaction effect. However, the main effect for the difficulty condition was significant. See Table 18 for the means and standard deviations and Figure 12 for a graphical representation of the relationship. Those who played on the Normal difficulty experienced more challenge/skill balance than those that played on the Peaceful difficulty. The hypothesis was fully-supported.

Table 18. Means and standard deviations of challenge/skill balance by condition.

	Low Immersion	High Immersion	Total
Peaceful Difficulty	4.01 (0.55)	3.89 (0.52)	3.95 (0.53)
Normal Difficulty	4.27 (0.71)	4.26 (0.58)	4.27 (0.65)
Total	4.16 (0.65)	4.08 (0.58)	4.12 (0.61)

Note: Standard deviations are listed in parentheses.

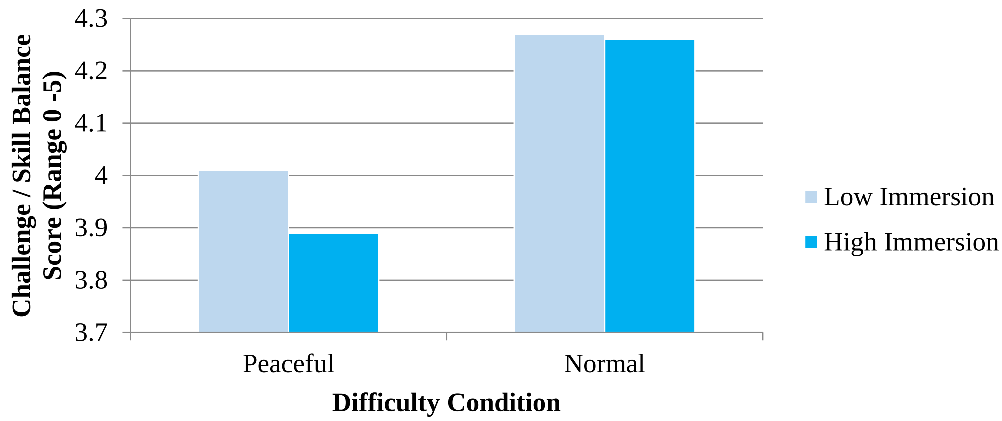


Figure 12. Challenge / skill balance scores by condition.

Based on the results for testing Hypotheses 1 and 2, the immersion manipulation did not have any effect on the amount of immersion experienced. The difficulty manipulation, however, did affect participant's reports of challenge/skill balance. Those who played on the Peaceful difficulty, which should have been very boring to these experienced players, did indeed experience a poor match between their skills and the challenge of the game. This is the mechanism for flow, therefore its manipulation affords the opportunity for a range of flow experiences. For example, a stepwise regression with all 9 FSS-2 subscales predicting the CORE flow mean score revealed that the three variables most influential in the score were the Concentration, Challenge/Skill Balance, and Feedback subscales. Also, a simple linear regression predicting the CORE flow score from the Challenge/Skill Balance score was significant: $F(1,83) = 52.91$, $p < .001$, CORE Flow = 2.12 + 0.547*(Challenge/Skill Balance).

The purpose of the experimental manipulation was to prevent range restriction. If only the optimal gameplay experienced was offered, players experiencing low amounts of immersion and flow may have been less likely. Even though the immersion condition did not affect immersion, a good range of scores for all four game engagement states was found (see Table 19):

Table 19. Means, standard deviations, and score ranges for game engagement variables.

	N	**Minimum**	**Maximum**	**Range**	**Mean**	**Std. Dev**	**Possible Scores**
Immersion	84	59	112	53 points	93.09	10.47	7 - 112
Presence Count	84	0	9	9 points	3.82	3.38	0 - 9
Presence %	79	0	95	95 points	58.65	32.47	0 - 100
Involvement	84	54	105	51 points	87.78	12.01	21 - 105
Flow (Core)	94	2.90	5.00	2.10 points	4.37	0.53	1 - 5

Theoretical hypotheses

Hypothesis 3a: *Usability and whether the individual plays with external distractions blocked out significantly and positively predict the attentional requirements score.*

A linear regression correlation analysis was conducted to predict participants' attentional requirements scores based on their overall rating of *Minecraft*'s usability. A combination of usability score and dummy-coded immersion condition (0 = low immersion; 1 = high immersion) significantly predicted participants' attentional requirements score, which accounted for R^2 = .165, which is 16.5% of the variance in the outcome score, $F(2,81) = 7.990$, $p = .001$. The coefficients are listed below in Table 20:

Table 20. Regression coefficients for predicting attentional requirements score.

	B	**Std. Error**	β	t	p	r	pr^2	sr^2
Constant	35.766	6.125		5.839	< .001			
Usability	0.275	0.070	.400	3.938	< .001	.403	.161	.160
Immersion Condition	0.733	1.452	.051	0.505	.615	.070	.003	.003

Examining the beta values, the immersion condition did not explain any meaningful variance in attentional requirements, however usability score did. As ratings of usability increased by one point, so did the ratings of how focused their attention was on playing the game by 0.275 points. The hypothesis was partially supported only because the immersion condition was not a significant predictor on its own.

Hypothesis 3b: *The attentional requirements score significantly and positively predicts low-level engagement scores.*

A hierarchical regression correlation analysis was conducted to predict participant's low-level game engagement scores based on their attentional requirements score. The total score for low-level game engagement was used. In addition to this, a second block was added to the hierarchical regression to examine how usability score contributed to the low-level game engagement score. This was done since testing Hypothesis 3a revealed that usability significantly and positively predicted the attentional requirements score.

In the first block, when attentional requirements was the sole predictor of low-level engagement score, the regression was not significant: $F(1,82) = 2.765$, $p = .100$, $R^2 = .003$. When usability was added in the second block, only to determine if usability mattered since attentional requirements did not, the combination of usability and attentional requirements significantly predicted low-level engagement score: $F(2,81) = 8.538$, $p < .001$, $R^2 = .174$. The addition of usability to the regression model was significant, $\Delta R^2 = .131$, $F(1,81) = 13.876$, $p < .001$. Coefficients are listed below in Table 21.

It does not appear that the attentional requirements score explains a meaningful amount of variance in low-level game engagement total score ($R^2 = .033$, or roughly 3.3% of the variance). Instead, perceptions of the game's usability predicts 17.4% of the variance in scores

(R^2 = .174). As ratings of usability increased by one point, so did the low-level game engagement total score by 0.746 points. This hypothesis was not supported because attentional requirements alone was not a significant predictor, however usability was.

Table 21. Regression coefficients for predicting low-level game engagement total score.

	B	Std. Error	β	*t*	*p*	*r*	pr^2	sr^2
BLOCK 1								
Constant	152.169	17.379		8.756	< .001			
Attentional Requirements	0.478	0.287	.181	1.663	.100	.181	.033	.033
BLOCK 2								
Constant	113.522	19.201		5.912	< .001			
Attentional Requirements	0.040	0.292	.015	0.137	.891	.181	< .001	< .001
Usability	0.746	0.200	.411	3.725	< .001	.417	.146	.141

Hypothesis 4: *The immersion total score and involvement total score are significantly and positively correlated.*

Pearson's product moment correlation coefficient was computed between immersion and involvement scores. Since this was a directional hypothesis, it was a one-tailed test. A significant and positive correlation was found, $r(82) = .426$, $p < .001$, $r^2 = .18$, a medium effect. This suggests that involvement and immersion are correlated, but not to the point that they are the same construct. The hypothesis was fully supported.

Hypothesis 5a: *The attentional requirements score, immersion total score, involvement total score, and the challenge/skill balance score are all significantly and positively predictive of the CORE flow mean score, however involvement is more influential than immersion.*

A multiple regression correlation analysis was conducted to predict the CORE flow score from the participant's attentional requirements, immersion, involvement, and challenge/skill balance scores. The four variables together significantly predicted flow scores and accounted for R^2 = .446, that is 44.6% of the variance in flow score: $F(4,79) = 15.901$, $p < .001$. Coefficients are listed below in Table 22.

Table 22. Regression coefficients for predicting flow score.

	B	**Std. Error**	***β***	***t***	***p***	***r***	***pr²***	***sr²***
Constant	1.967	0.533		3.689	< .001			
Attention Requirements	-0.011	0.007	-.149	-1.704	.092	-.096	.035	.020
Involvement	0.007	0.005	.153	1.497	.138	.462	.028	.016
Immersion	0.009	0.005	.182	1.752	.084	.446	.037	.022
Challenge / Skill Balance	0.391	0.096	.447	4.081	< .001	.626	.174	.117

The hypothesis tested states that "involvement will be more influential than immersion." To determine this, the zero-order, partial, and semi-partial correlations can be examined. Strictly looking at the zero-order (r) correlations, involvement has a higher correlation with flow than immersion does. However, examining the partial (pr^2) and semi-partial (sr^2) correlations reveals how much of the variance was explained uniquely due to the variable with and without considering the other variables in the set, respectively. Both the partial and semi-partial correlations suggested that immersion explained more variance in the flow score than involvement did, but only by a very small amount. It may be that this difference is not significantly different and they both roughly explain similar amounts of variance in the flow score. Furthermore, none of the variables aside from challenge/skill balance were uniquely significant. Because of this, the hypothesis was only partially supported.

Since the previous hypotheses found that usability score was more influential than the attention requirements score in predicting low-level game engagement, a follow-up analysis was conducted. This time, a hierarchical regression correlation analysis was conducted. In the first block, usability, involvement, and immersion were entered as variables. In the second block, challenge/skill balance was entered as a variable.

The first block was significant, and a combination of usability, involvement, and immersion predicted R^2 = .301, that is 30.1% of the variance in CORE flow score: $F(3, 80) = 11.470$, $p < .001$. Adding challenge/skill balance significantly increased the amount of explained variance ($\Delta R^2 = .131$, $F(1,79) = 18.293$, p < .001). The second block was also significant, and the combination of the four variables predicted R^2 = .432, that is 43.2% of the variance in flow score: $F(4, 79) = 15.035$, $p < .001$. Coefficients are listed below in Table 23.

Table 23. Regression coefficients for predicting flow score.

	B	**Std. Error**	***β***	***t***	***p***	***r***	***pr*^{2}**	***sr*^{2}**
BLOCK 1								
Constant	1.892	0.539		3.509	.001			
Usability	-0.006	0.005	-.119	-1.147	.255	.128	.016	.011
Involvement	0.016	0.005	.353	3.365	.001	.462	.124	.099
Immersion	.018	0.006	.343	3.154	.002	.446	.111	.087
BLOCK 2								
Constant	1.654	0.492		3.360	.001			
Usability	-0.005	0.005	-.090	-0.956	.342	.128	.011	.007
Involvement	0.007	0.005	.163	1.556	.124	.462	.030	.017
Immersion	0.009	0.005	.167	1.561	.122	.446	.030	.017
Challenge / Skill Balance	0.410	0.096	.469	4.277	< .001	.626	.188	.132

Interestingly, not much changed between the analyses. Challenge/skill balance is still extremely influential in the CORE flow score. Whether involvement or immersion has more of an influence on the CORE flow score flips between blocks and the scores remain incredibly close, which continues to suggest that they explain roughly the same amount of variance. However, most telling is that including usability rather than attentional requirements did not seem to matter. As seen in Hypothesis 3b, usability is a strong, positive predictor of low-level game engagement, but when the variables for low-level game engagement are included alongside it, the potential significance of the contribution is lost.

Hypothesis 5b: *The attentional requirements score, immersion score, and involvement score are all significantly and positively predictive of the presence count, however immersion is* *more influential than involvement. The challenge/skill balance score, while also included, is not a significant predictor of presence count.*

To test the above hypothesis, a hierarchical regression correlation analysis was conducted. In the first block, attention requirements, involvement, and immersion were entered as variables. In the second block, challenge/skill balance was entered as a variable.

The first block was significant, and a combination of attention requirements, involvement, and immersion predicted $R^2 = .310$, that is 31.0% of the variance in presence count: $F(3, 80) = 12.005$, $p < .001$. As predicted, adding challenge/skill balance as a predictor did not significantly increase the amount of explained variance: $\Delta R^2 = .010$, $F(1,79) = 1.193$, $p = .278$. Coefficients are listed below in Table 24. Note that the coefficients for the second block have been omitted from this table since the addition of challenge/skill balance was not significant:

Table 24. Regression coefficients for predicting presence count.

BLOCK 1	**B**	**Std. Error**	***β***	***t***	***p***	***r***	***pr²***	***sr²***
Constant	-13.343	3.689		-3.617	.001			
Attentional Requirements	-0.023	0.045	-.049	-0.511	.611	.072	.003	.002
Involvement	0.062	0.029	.221	2.152	.034	.402	.055	.040
Immersion	0.141	0.034	.435	4.129	< .001	.517	.176	.147

Only involvement and immersion were significantly and positively predictive of presence count, but this time it does appear that immersion was more influential than involvement. Also, as predicted, challenge/skill balance did not matter for presence count like it had for flow. Because of this, the hypothesis was only partially supported.

A follow-up hierarchical regression correlation analysis, where attentional requirements was replaced by usability, was also conducted. In the first block, usability, involvement, and immersion were entered as variables. In the second block, challenge/skill balance was entered as a variable. The first block was significant, and a combination of usability, involvement, and immersion predicted R^2 = .363, that is 36.3% of the variance in presence count: $F(3, 80)$ = 15.195, p < .001. Once again, adding challenge/skill balance did not significantly increase the amount of explained variance: ΔR^2 = .008, $F(1,79)$ = 1.042, p = .310. Coefficients are listed below in Table 25, and again the second block is omitted due to lack of significant improvement over the first block:

Table 25. Regression coefficients for predicting presence count.

BLOCK 1	**B**	**Std. Error**	***β***	***t***	***p***	***r***	***pr²***	***sr²***
Constant	-10.751	3.256		-3.301	.001			
Usability	-0.084	0.032	-.259	-2.623	.010	.027	.079	.055
Involvement	0.075	0.028	.266	2.661	.009	.402	.081	.056
Immersion	0.164	0.034	.508	4.892	< .001	.517	.230	.191

In this case, replacing attentional requirements with usability was meaningful. All three predictors, usability included, were positively significant with respect to predicting presence count. Furthermore, immersion continued to be more influential than involvement. If the hypothesis had stated usability instead of attentional requirements, it would have been fully supported.

Hypothesis 6: *The total presence count is significantly and positively predicted by high immersion total score, the presence of immersive cues, and high levels of individual differences (to include willingness to suspend disbelief, creative imagination, and immersive/involvement tendencies). The level of challenge/skill balance does not matter, but involvement may.*

To test the above hypothesis, a hierarchical regression correlation analysis was conducted. For this analysis, the sample size requirement was violated. Since there were 7 predictors, a sample size of at least 106 was required, which was 22 participants above the number of participants found in this analysis (N = 84). The analysis was still conducted, with the caveat that it may be inaccurate after the third block (which has 5 predictors and required N = 90). The results of the analysis are found below in Table 26 with the coefficients listed in Table 27.

Table 26. Summary of regression analysis for predicting presence count.

Block	Predictors Added	R^2	F-test	ΔR^2	ΔR^2 F-Test
1	Immersion Condition	.001	$F(1,82) = 0.090, p = .765$		
2	Immersion	.268	$F(2,81) = 14.816, p < .001$	.267	$F(1,81) = 29.511, p < .001$
3	Willingness to Suspend Disbelief; Creative Imagination; Immersive / Involvement Tendencies	.331	$F(5,78) = 7.734, p < .001$	.064	$F(3,78) = 2.473, p = .068$
4	Challenge / Skill Balance	.339	$F(6,77) = 6.592, p < .001$	.008	$F(1,77) = 0.921, p = .340$
5	Involvement	.355	$F(7,76) = 5.972, p <. 001$	.016	$F(1,76) = 1.827, p = .181$

Adding in the individual differences at the third block approached significance, but fell short: $F(3,78) = 2.473$, $p = .068$. Neither challenge/skill balance nor involvement was a significant predictor of presence count. Furthermore, condition continued to not have an impact on presence count scores. Therefore, this hypothesis was only partially supported.

Upon examination of the regression coefficients in the third block, it appears that creative imagination is the least important predictor in the set of three individual differences explored. A follow-up hierarchical regression correlation analysis was conducted to determine if a combination of immersion, willingness to suspend disbelief, and immersive/involvement tendencies, and then the addition of creative imagination, significantly predicted presence count. The results of the analysis are found below in Table 28 with the coefficients listed in Table 29.

The combination of immersion, willingness to suspend disbelief, and immersive/involvement tendencies significantly predicted presence count. Adding in creative imagination as a predictor did not explain any additional variance. While immersion continued to be the strongest, most influential predictor, immersive/involvement tendencies also was relevant, especially since it was nearly significant by itself.

Table 27. Regression coefficients for predicting presence count.

	B	Std. Error	*β*	*t*	*p*	*r*	*pr²*	*sr²*
BLOCK 1								
Constant	3.707	0.531		6.977	< .001			
Immersion Condition	0.223	0.743	.033	0.300	.765	.033	.001	.001
BLOCK 2								
Constant	-11.764	2.884		-4.078	< .001			
Immersion Condition	0.061	0.640	.009	0.095	.924	.033	< .001	< .001
Immersion	0.167	0.031	.517	5.432	< .001	.517	.267	.266
BLOCK 3								
Constant	-15.871	3.190		-4.975	< .001			
Immersion Condition	0.408	0.652	.061	0.626	.533	.033	.005	.003
Immersion	0.128	0.034	.397	3.791	< .001	.517	.155	.123
Willingness to SD	0.495	0.335	.151	1.480	.143	.324	.027	.019
Creative Imagination	0.032	0.036	.097	0.881	.381	.290	.010	.007
Imm / Involv Tendencies	0.033	0.025	.144	1.345	.182	.326	.023	.016
BLOCK 4								
Constant	-16.033	3.196		-5.016	< .001			
Immersion Condition	0.450	0.654	.067	0.688	.494	.033	.006	.004
Immersion	0.113	0.037	.351	3.034	.003	.517	.107	.079
Willingness to SD	0.429	0.342	.131	1.254	.214	.324	.020	.013
Creative Imagination	0.033	0.036	.101	0.918	.361	.290	.011	.007
Imm / Involv Tendencies	0.024	0.027	.104	0.908	.367	.326	.011	.007
Challenge / Skill Balance	0.651	0.678	.118	0.960	.340	.425	.012	.008
BLOCK 5								
Constant	-17.191	3.293		-5.221	< .001			
Immersion Condition	0.472	0.651	.070	0.726	.470	.033	.007	.004
Immersion	0.102	0.038	.315	2.672	.009	.517	.086	.061
Willingness to SD	0.470	0.342	.143	1.375	.173	.324	.024	.016
Creative Imagination	0.037	0.036	.113	1.027	.308	.290	.014	.009
Imm / Involv Tendencies	0.015	0.027	.064	0.548	.585	.326	.004	.003
Challenge / Skill Balance	0.342	0.712	.062	0.480	.632	.425	.004	.002
Involvement	0.044	0.033	.157	1.352	.181	.402	.023	.016

Note: Willingness to SD = Willingness to Suspend Disbelief; Imm / Involv Tendencies = Immersive / Involvement Tendencies.

Table 28. Summary of regression analysis for predicting presence count.

Block	Predictors Added	R^2	*F*-test	ΔR^2	ΔR^2 *F*-Test
1	Immersion; Willingness to Suspend Disbelief; Immersive / Involvement Tendencies	.323	$F(3,80) = 12.743$, $p < .001$		
2	Creative Imagination	.328	$F(4,79) = 9.644$, $p < .001$	.005	$F(1,79) = 0.558$, $p = .457$

Table 29. Regression coefficients for predicting presence count.

	B	Std. Error	β	t	p	r	pr^2	sr^2
BLOCK 1								
Constant	-15.334	3.118		-4.917	< .001			
Immersion	0.139	0.032	.429	4.347	< .001	.517	.191	.160
Willingness to SD	0.436	0.327	.133	1.335	.186	.324	.022	.015
Imm / Involv Tendencies	0.042	0.023	.178	1.828	.071	.326	.040	.028
BLOCK 2								
Constant	-15.584	3.145		-4.955	< .001			
Immersion	0.132	0.033	.408	3.963	< .001	.517	.166	.134
Willingness to SD	0.464	0.330	.142	1.407	.163	.324	.024	.017
Imm / Involv Tendencies	0.034	0.025	.148	1.393	.167	.326	.024	.017
Creative Imagination	0.026	0.034	.079	0.747	.457	.290	.007	.005

Note: Willingness to SD = Willingness to Suspend Disbelief; Imm / Involv Tendencies = Immersive / Involvement Tendencies.

Hypothesis 7: *The flow score is significantly and positively predicted by high involvement total score, playing on the congruent game difficulty, the presence of high flow mechanisms scores (feedback, goals, challenge/skill balance), and high levels of individual differences (to include willingness to suspend disbelief, creative imagination, and immersive/ involvement tendencies). The level of presence of immersive cues does not matter, but the immersion score itself may.*

To test the above hypothesis, a hierarchical regression correlation analysis was conducted. For this analysis, the sample size requirement was violated. Since there were 14 predictors, a sample size of at least 162 was required, which was a number far higher than the number of participants in this study ($N = 84$). The analysis was still conducted, with the caveat that it may be inaccurate after the third block (which has 5 predictors and requires $N = 90$). The results of the analysis are found below in Table 30 with the coefficients listed in Table 31.

Table 30. Summary of regression analysis for predicting flow.

Block	Predictors Added	R^2	*F*-test	ΔR^2	ΔR^2 *F*-Test
1	Involvement condition	.006	$F(1,81) = 0.528$, $p = .470$		
2	Involvement	.216	$F(2,80) = 14.003$, $p < .001$	.209	$F(1,80) = 21.345$, $p < .001$
3	Feedback; Goals; Challenge / Skill Balance	.539	$F(5,77) = 18.025$, $p < .001$	.324	$F(3,77) = 18.024$, $p < .001$
4	Dispositional flow; Trait Absorption; Age; VGSE	.571	$F(9,73) = 10.791$, $p < .001$	.032	$F(4,73) = 1.345$, $p = .262$
5	Willingness to Suspend Disbelief; Creative Imagination; Immersive / Involvement Tendencies	.585	$F(12,70) = 8.239$, $p < .001$	.015	$F(3,70) = 0.820$, $p = .487$
6	Immersion	.588	$F(13,69) = 7.583$, $p < .001$	.003	$F(1,69) = 0.467$, $p = .497$
7	Immersion condition	.590	$F(14,68) = 7.002$, $p < .001$	.002	$F(1,68) = 0.363$, $p = .549$

The hypothesis was partially supported. None of the proposed individual differences were predictors, and neither was the involvement condition. The key predictors of flow were involvement and then two of the flow mechanisms: challenge/skill balance and feedback. The flow subscale for goals was not significant. This may have been because *Minecraft* itself does not have goals other than the ones the player imposes on gameplay. The FSS-2's Clear Goal's subscale may not have captured this particular aspect of gameplay.

Hypothesis 8: *Only in the instance where game feature mechanisms support both presence and flow are these states significantly associated, otherwise there is no significant association between presence and flow.*

Pearson's product moment correlation coefficient was computed between the CORE flow mean score and the presence count. For testing this hypothesis, the correlations are calculated by condition (see Table 32). Since this was a directional hypothesis, these were all one-tailed tests.

Table 31. Regression coefficients for predicting flow.

	B	Std. Error	*β*	*t*	*p*	*r*	pr^2	sr^2
BLOCK 1								
Constant	4.333	0.086		50.299	< .001			
Difficulty Condition	0.086	0.118	.080	0.726	.470	.080	.006	.006
BLOCK 2								
Constant	2.557	0.392		6.523	< .001			
Difficulty Condition	0.064	0.106	.060	0.608	.545	.080	.005	.004
Involvement	0.020	0.004	.458	4.620	< .001	.461	.211	.209
BLOCK 3								
Constant	0.715	0.485		1.475	.144			
Difficulty Condition	-0.102	0.093	-.095	-1.097	.276	.080	.015	.007
Involvement	0.005	0.004	.108	1.150	.254	.461	.017	.008
Challenge / Skill Balance	0.360	0.092	.413	3.915	< .001	.624	.166	.092
Goals (transformed)	0.247	0.277	.087	0.891	.375	-.334	.010	.005
Feedback	0.399	0.090	.433	4.435	< .001	.621	.203	.118
BLOCK 4								
Constant	1.543	0.765		2.018	.047			
Difficulty Condition	-0.141	0.095	-.132	-1.489	.141	.080	.030	.013
Involvement	0.004	0.004	.094	0.995	.323	.461	.013	.006
Challenge / Skill Balance	0.412	0.095	.471	4.320	< .001	.624	.203	.110
Goals (transformed)	0.253	0.276	.089	0.916	.363	-.334	.011	.005
Feedback	0.394	0.097	.428	4.044	< .001	.621	.183	.096
Dispositional Flow	0.126	0.107	.128	1.173	.245	.414	.018	.008
Trait Absorption	-0.007	0.006	-.091	-1.098	.276	.187	.016	.007
Age	-0.031	0.029	-.083	-1.053	.296	-.012	.015	.007
VGSE	-0.021	0.012	-.163	-1.691	.095	.174	.038	.017
BLOCK 5								
Constant	1.423	0.812		1.753	.084			
Difficulty Condition	-0.144	0.095	-.135	-1.516	.134	.080	.032	.014
Involvement	0.003	0.004	.067	0.692	.491	.461	.007	.003
Challenge / Skill Balance	0.399	0.103	.457	3.890	< .001	.624	.178	.089
Goals (transformed)	0.202	0.279	.071	0.724	.471	-.334	.007	.003
Feedback	0.373	0.100	.405	3.735	< .001	.621	.166	.082
Dispositional Flow	0.143	0.110	.146	1.304	.197	.414	.024	.010
Trait Absorption	-0.012	0.007	-.152	-1.607	.113	.187	.036	.015
Age	-0.028	0.030	-.077	-0.956	.342	-.012	.013	.005
VGSE	-0.020	0.012	-.157	-1.615	.111	.174	.036	.015
Willing to SD	-0.034	0.046	-.066	-.746	.458	.188	.008	.003
Creative Imagination	0.003	0.005	.055	0.610	.544	.256	.005	.002
Imm / Involv Tend	0.004	0.004	.115	1.119	.267	.395	.018	.007
BLOCK 6								
Constant	1.454	0.816		1.782	.079			
Difficulty Condition	-0.158	0.097	-.148	-1.618	.110	.080	.036	.016
Involvement	0.003	0.004	.079	0.797	.428	.461	.009	.004
Challenge / Skill Balance	0.416	0.106	.476	3.928	< .001	.624	.182	.092
Goals (transformed)	0.216	0.281	.076	0.770	.444	-.334	.008	.004
Feedback	0.404	0.110	.438	3.674	< .001	.621	.164	.081
Dispositional Flow	0.144	0.110	.147	1.307	.195	.414	.024	.010
Trait Absorption	-0.012	0.007	-.159	-1.669	.100	.187	.039	.017
Age	-0.026	0.030	-.070	-0.855	.396	-.012	.010	.004
VGSE	-0.020	0.012	-.161	-1.646	.104	.174	.038	.016
Willing to SD	-0.027	0.047	-.051	-0.565	.574	.188	.005	.002
Creative Imagination	0.004	0.005	.075	0.786	.435	.256	.009	.004
Imm / Involv Tend	0.004	0.004	.104	0.993	.324	.395	.014	.006
Immersion	-0.004	0.006	-.076	-0.684	.497	.446	.007	.003
BLOCK 7								
Constant	1.474	0.821		1.796	.077			
Difficulty Condition	-0.146	0.100	-.137	-1.467	.147	.080	.031	.013
Involvement	0.003	0.004	.079	0.794	.430	.461	.009	.004
Challenge / Skill Balance	0.413	0.106	.473	3.880	< .001	.624	.181	.091
Goals (transformed)	0.150	0.303	.053	0.493	.624	-.334	.004	.001
Feedback	0.395	0.111	.428	3.541	.001	.621	.156	.076
Dispositional Flow	0.137	0.111	.140	1.233	.222	.414	.022	.009
Trait Absorption	-0.012	0.007	-.152	-1.569	.121	.187	.035	.015
Age	-0.025	0.030	-.067	-0.815	.418	-.012	.010	.004
VGSE	-0.020	0.012	-.159	-1.619	.110	.174	.037	.016
Willing to SD	-0.023	0.047	-.045	-0.493	.623	.188	.004	.001
Creative Imagination	0.005	0.005	.089	0.904	.369	.256	.012	.005
Imm / Involv Tend	0.004	0.004	.098	0.925	.358	.395	.012	.005
Immersion	-0.004	0.006	-.084	-0.745	.459	.446	.008	.003
Immersion Condition	0.057	0.094	.053	0.602	.549	.074	.005	.002

Note: Willingness to SD = Willingness to Suspend Disbelief; Imm / Involv Tendencies = Immersive / Involvement Tendencies; VGSE = Video game self-efficacy.

Table 32. Correlations between presence and flow by condition.

Condition	**Immersion**	**Difficulty**	**Correlation**	r^2	***p*-level**
A	Low	Peaceful	$r(16) = .236, p = .172$	.056	.172
B	Low	Normal	$r(21) = .416, p = .024$	.173	.024
C	High	Peaceful	$r(19) = .460, p = .018$	.212	.018
D	High	Normal	$r(20) = .541, p = .005$	.293	.005

The most significant and strongest correlation was in Condition D, which supported both flow and presence experiences, and was non-significant and weakest in Condition A, which encouraged neither state. This partially supports Hypothesis 8.

Hypothesis 8 revisited: To revisit Hypothesis 8, individual differences, experimental conditions, and the other game engagement states were entered into a stepwise regression to determine how they predicted immersion, presence, involvement, and flow. See Appendix F for the full list of the predictor variables used in these four stepwise regressions (*pin* = .05, *pout* = .10). For their respective regressions, the predictor variable that also happened to be the outcome variable was removed. Several assumptions were not met for this test (see Appendix G). For example, given that each analysis has 16 predictor variables, the sample size assumption is not met since this requires $N = 178$ and this dataset only had 84 participants. Since this is not connected to any particular hypothesis and is just an exploration of the data, these analyses were conducted despite the violations. See Table 33 for the results of these analyses and Table 34 for the regression coefficients.

Table 33. Summary of stepwise regression analyses for predicting game engagement states.

Outcome Variable	R^2	***F*-test**
Immersion (total score)	.561	$F(6,76) = 16.200, p < .001$
Presence (count)	.380	$F(4,78) = 11.938, p < .001$
Involvement (total score)	.449	$F(5,77) = 12.559, p < .001$
CORE Flow	.352	$F(3,79) = 14.303, p < .001$

Table 34. Regression coefficients for predicting immersion.

Immersion	**B**	**Std. Error**	***β***	***t***	***p***	***r***	***pr²***	***sr²***
Constant	-11.929	16.269		-0.733	.466			
Presence (count)	1.021	0.270	.327	3.782	< .001	.519	.158	.082
Usability	0.389	0.078	.388	5.011	< .001	.401	.248	.145
Creative Imagination	0.243	0.083	.240	2.938	.004	.330	.102	.050
MC Rating – Survival	2.257	0.930	.187	2.426	.018	.257	.072	.034
Flow (core)	3.829	1.684	.195	2.274	.026	.446	.064	.030
Age	1.210	0.559	.166	2.166	.033	.127	.058	.027
Presence (count)	**B**	**Std. Error**	***β***	***t***	***p***	***r***	***pr²***	***sr²***
Constant	-12.627	3.490		-3.618	.001			
Immersion	0.147	0.035	.459	4.163	< .001	.519	.181	.138
Flow (core)	1.295	0.630	.206	2.056	.043	.416	.052	.033
Usability	-0.076	0.032	-.236	-2.381	.020	.028	.068	.045
Willing to SD	0.639	0.311	.196	2.052	.044	.319	.051	.033
Involvement	**B**	**Std. Error**	***β***	***t***	***p***	***r***	***pr²***	***sr²***
Constant	-12.716	13.790		-0.922	.359			
Flow (core)	6.625	2.095	.294	3.162	.002	.461	.115	.071
Imm / Involv Tendencies	0.241	0.076	.292	3.162	.002	.425	.115	.071
Usability	0.372	0.101	.325	3.683	< .001	.310	.150	.097
Hours Playing Minecraft	-0.011	0.004	-.227	-2.597	.011	-.162	.081	.048
MC Rating – Survival	2.949	1.169	.214	2.522	.014	.244	.076	.045
CORE Flow	**B**	**Std. Error**	***β***	***t***	***p***	***r***	***pr²***	***sr²***
Constant	1.956	0.471		4.156	< .001			
Involvement	0.013	0.005	.283	2.789	.007	.461	.089	.064
Dispositional Flow	0.283	0.093	.288	3.039	.003	.414	.104	.076
Presence (count)	0.039	0.016	.246	2.482	.015	.416	.072	.051

Note: MC Rating – Survival is the rating that the participant assigned to Minecraft's Survival Mode, which is what they played during the experiment; Willing to SD = willingness to suspend disbelief; Imm / Involv Tendencies = immersive / involvement tendencies.

Additional questions

To determine whether cognitive absorption was the same theoretical construct as flow, the following four sub-hypotheses were tested.

Hypothesis 9a: *The mean cognitive absorption score and the CORE flow mean score is significantly, strongly, and positively correlated (at least r = .50).*

Pearson's product moment correlation coefficient was computed between the CORE flow mean score and the CAS score. Since this was a directional hypothesis, it was a one-tailed test.

A significant and positive correlation was found, $r(81) = .673$, $p < .001$, $r^2 = .45$, a strong effect. This suggests that the flow and cognitive absorption scores are very highly correlated and may indeed be a part of the same construct. The hypothesis was supported.

Hypothesis 9b: *The Temporal Dissociation CAS mean and the Transformation of Time FSS-2 mean is significantly, strongly, and positively correlated (at least r = .50).*

Pearson's product moment correlation coefficient was computed between the time subscales from the FSS-2 and CAS. Since this was a directional hypothesis, it was a one-tailed test. A significant and negative correlation was found, $r(81) = -.693$, $p < .001$, $r^2 = .48$, a strong effect. The temporal variable for flow was transformed, which involved reflecting the scores, thus making the correlation negative. This suggests that temporal subscales for both the flow and cognitive absorption scores are very highly correlated and may be a part of the same construct. The hypothesis was supported.

Hypothesis 9c: *The Control CAS mean and the Control FSS-2 mean is significantly, strongly, and positively correlated (at least r = .50).*

Pearson's product moment correlation coefficient was computed between the control subscales for the FSS-2 and CAS. Since this was a directional hypothesis, it was a one-tailed test. The scatterplot of the two variables suggested a very weak linear relationship, which may be a violation of the assumptions. A significant and negative correlation was found, $r(81) = -.576$, $p < .001$, $r^2 = .33$, a strong effect. The CAS Control variable was transformed, which involved reflecting the scores, thus making the correlation negative. This suggested that control subscales for both the flow and cognitive absorption scores are highly correlated and may be a part of the same construct. The hypothesis was supported.

Hypothesis 9d: *The Heightened Enjoyment CAS mean and the Autotelic Experience FSS-2 mean is significantly, strongly, and positively correlated (at least r = .50).*

Pearson's product moment correlation coefficient was computed between the FSS-2 Autotelic Experience and the CAS Enjoyment subscales. Since this was a directional hypothesis, it was a one-tailed test. The Durbin-Watson statistic suggested that independence of observations assumption may have been violated. A significant and positive correlation was found, $r(81) = .660$, $p < .001$, $r^2 = .44$, a strong effect. This suggests that the autotelic, motivating, and enjoyable aspects of flow and cognitive absorption scores are very highly correlated and may indeed be a part of the same construct. The hypothesis was supported.

Summary

The results of the statistical analyses are summarized below in Table 35:

Table 35. Results of hypothesis testing.

#	Hypothesis	Support
0.1	The scores from the new willingness to suspend disbelief items are significantly, strongly, and positively correlated (at least $r = .50$), but not redundant with (more than $r = .90$) the score from Sas and O'Hare's (2004) single item.	Partially Supported
0.2	The items from the PQ are divided amongst the proposed subscales (immersion, involvement, and attentional requirements) as seen in an Exploratory Factor Analysis.	Partially Supported
0.3	The scores from the TAM-influenced involvement items (Davis, 1989) as well as the PQ's involvement item (Witmer & Singer, 1998) are significantly, strongly, and positively correlated (at least $r = .50$), but not redundant with (more than $r = .90$) the total score from the Revised PII (Zaichkowsky, 1994).	Supported
0.4a	Presence count is significantly and positively correlated with the presence estimate in the high immersion and congruent preferences group.	Supported
0.4b	Presence classification is significantly and positively predicted by presence count in the high immersion and congruent preferences group.	Supported
0.5	The C FSS mean score and the overall FSS-2 mean score is significantly, strongly, and positively correlated ($r > .70$).	Supported
1	Ratings of immersion significantly increase when more immersive cues are present.	Not Supported
2	Those in the congruent preferences group have significantly higher ratings of challenge/skill balance.	Supported

3a	Usability and whether the individual plays with external distractions blocked out significantly and positively predict the attentional requirements score.	Partially Supported
3b	The attentional requirements score significantly and positively predicts low-level engagement scores.	Not Supported
4	The immersion total score and involvement total score are significantly and positively correlated.	Supported
5a	The attentional requirements score, immersion total score, involvement total score, and the challenge/skill balance score are all significantly and positively predictive of the CORE flow mean score, however involvement is more influential than immersion.	Partially Supported
5b	The attentional requirements score, immersion score, and involvement score are all significantly and positively predictive of the presence count, however immersion is more influential than involvement. The challenge/skill balance score, while also included, is not a significant predictor of presence count.	Partially Supported
6	The total presence count is significantly and positively predicted by high immersion total score, the presence of immersive cues, and high levels of individual differences (to include willingness to suspend disbelief, creative imagination, and immersive/involvement tendencies). The level of challenge/skill balance does not matter, but involvement may.	Partially Supported
7	The flow score is significantly and positively predicted by high involvement total score, playing on the congruent game difficulty, the presence of high flow mechanisms scores (feedback, goals, challenge/skill balance), and high levels of individual differences (to include willingness to suspend disbelief, creative imagination, and immersive/involvement tendencies). The level of presence of immersive cues does not matter, but the immersion score itself may.	Partially Supported
8	Only in the instance where game feature mechanisms support both presence and flow are these states significantly associated, otherwise there is no significant association between presence and flow.	Partially Supported
9a	The mean cognitive absorption score and the CORE flow mean score is significantly, strongly, and positively correlated (at least r = .50).	Supported
9b	The Temporal Dissociation CAS mean and the Transformation of Time FSS-2 mean is significantly, strongly, and positively correlated (at least r = .50).	Supported
9c	The Control CAS mean and the Control FSS-2 mean is significantly, strongly, and positively correlated (at least r = .50).	Supported
9d	The Heightened Enjoyment CAS mean and the Autotelic Experience FSS-2 mean is significantly, strongly, and positively correlated (at least r = .50).	Supported

CHAPTER FIVE: DISCUSSION

This study of 84 *Minecraft* players provided support for the proposed R-GEM. The twenty hypotheses were divided across four broad categories and the results of which are reviewed below.

Hypothesis 0.1 was concerned with whether the newly proposed willingness to suspend disbelief items could be combined with the single item proposed by Sas and O'Hare (2004) to create the Willingness to Suspend Disbelief Scale. The items were somewhat correlated, although the New Willing Item 1 was weakly correlated in particular. A follow-up analysis not reported here re-computed the willingness to suspend disbelief mean score removing the weak item and re-analyzed the affected hypothesis. The results did not come out any differently. Therefore, this newly proposed Willingness to Suspend Disbelief Scale was valid for the purposes of the game engagement analysis.

Hypothesis 0.2 was tested to determine if the PQ (version 3; Witmer et al., 2005) could be divided into subscales for immersion, attentional requirements, and involvement. As a cursory look, the immersion and attentional requirements items form their own subscales, although the single involvement item loaded on a factor with other immersion items. This could be due to the positively reciprocal relationship between immersion and involvement. Still, as a preliminary analysis, these findings are promising and it appears that dividing the PQ into these subscales is a step in the right direction.

The purpose of testing Hypothesis 0.3, like for Hypothesis 0.1, was to determine if additional items influenced by the TAM (Davis, 1989) could be added to the R PII (Zaichkowsky, 1994) to create a more reliable measure tailored to game involvement, the Game

Involvement Scale. The correlations were quite high, on average. Still, one item, the TAM Item 4, was weakly correlated compared to the others. A second version of this measure was created by finding the total without TAM Item 4. A follow-up analysis not reported here did not find the removal of TAM Item 4 to change any of the statistical analysis results, therefore all items could be used to generate the new Game Involvement Scale score.

Hypotheses 0.4a and 0.4b were focused on the measurement of presence. Slater and Steed (2000) suggested that one way to measure presence was to ask participants to report the percentage of time they felt present while in the virtual environment. The reported presence estimates percentages were very strongly, positively correlated with the presence count method proposed by Slater and Steed (2000). Presence count can also be used to correctly classify ~ 80% of the cases where a participant experienced presence over half of the time. These measures are a good starting point for measuring presence and should be used in other studies, especially instead of the PQ.

Hypothesis 0.5 explored whether the C FSS's (Martin & Jackson, 2008) CORE flow mean score could be used in lieu of the mean overall score from the FSS-2 (Jackson & Eklund, 2002). Recall that Jackson et al. (2010b) did not advise using the mean FSS-2 score for measuring flow, and Procci et al. (2012) did not find its counterpart, the DFS-2, to be a good fit for measuring the flow experience of gamers. Furthermore, the shorter version of the FSS-2, the S FSS, was found to not be as reliable as the S DFS (Jackson et al., 2008). Therefore, a suitable replacement for measuring flow in the gaming context is needed. This study examined whether a different type of flow scale had potential. Given the very strong, positive correlation between C FSS CORE flow mean and FSS-2 mean score for overall flow experienced, the C FSS CORE flow mean score can be used.

Overall, the above findings suggested that the proposed methods of measuring the different constructs related to game engagement had merit. The analysis of the experimental and theoretical hypotheses could proceed.

Hypothesis 1 and 2 were used to determine if the experimental manipulations were successful. The immersion manipulation did not affect participants' ratings of immersion, however the game difficulty manipulation did affect ratings of challenge/skill balance. Challenge/skill balance is a key mechanism for flow, so this was the desired effect. Still, the goal of the manipulation was to create a range of possible scores so that the regression analyses found in later hypothesis were not negatively affected by range restriction. A range of experiences was achieved, so the failure of the immersion condition to produce different subjective experiences between conditions is acceptable in this case.

Regardless of the experimental manipulation, the participants' immersion scores were very high, where the average immersion score was roughly 83% of the total possible score. This is interesting since *Minecraft* itself is such a low-resolution game, which does not align with the hypotheses of Slater and Wilbur (1997) and Witmer and Singer (1998), where more vivid, detailed aspects of virtual environments should result in more immersion. It may be that these experienced *Minecraft* players, since they have a history of choosing to play this particular game, were more accepting of and less distracted by the low fidelity models and textures, and instead focused more on actual gameplay. This high level of game-specific involvement had a positive influence on immersion, where experienced players mitigated the lack of immersive characteristics, perhaps purely out of familiarity, acceptance, and desire to play. Essentially, these individuals had a history of overlooking the graphical shortcomings of *Minecraft* and played it on a regular basis, likely becoming immersed anyway. *Minecraft* novices, who are unfamiliar with the game and are perhaps used to playing other high-quality games may not be able to ignore the fidelity deficit, and their level of immersion may have been lower in the low immersion condition than in the high immersion condition. This finding supports the notion that it is not just the aspects of the technology that encourage immersion, as Slater and Wilbur (1997) initially suggested. Instead, immersion is the result of a combination of other factors, such as experience and involvement, in addition to the technology (Singer & Witmer, 1999).

Hypotheses 3 through 8 all examined theoretical aspects of the R-GEM. Hypothesis 3a and 3b were concerned with the attentional requirements aspect. Usability strongly predicted the attentional requirements score, but the attentional requirements score did not significantly contribute to explaining variance of low-level game engagement scores. This is surprising given the number of researchers who highlighted the importance of attentional requirements (Draper et al., 1998; Witmer & Singer, 1998; Wirth et al., 2007). Usability, however, did. Usability should take the place of attentional requirements in the R-GEM, until future studies can further explore attentional requirements as a construct related to game engagement.

Hypothesis 4 examined the relationship between immersion and involvement. They were highly correlated, but not so much so that they were the same construct. This supports the notion that they are related, but not the same, game engagement states (Witmer & Singer, 1998).

Hypothesis 5a examined the significant predictors of flow. The analysis revealed that usability, involvement, and immersion significantly predicted flow, and that adding in challenge/skill balance, a mechanism of flow, further increased the predictive power of the regression. This supports the proposed R-GEM given that attentional requirements construct is replaced by usability. Challenge/skill balance emerged as the strongest, most influential predictor of flow, as hypothesized. Interestingly, immersion and involvement influenced flow equally. This may be due to the conceptual overlap between some mechanisms of flow and immersion, such as the aspects of control.

Hypothesis 5b examined the significant predictors of presence. The analysis revealed that usability, involvement, and immersion significantly predicted presence, and that challenge/skill balance did not affect the relationship. It further establishes that presence and flow are different game engagement states given that the key flow mechanism of challenge/skill balance was not a significant predictor of presence. Also, usability, involvement, and immersion were all uniquely significant. Since both involvement and immersion were both unique predictors, this is suggestive of the highly-related yet distinct nature of the constructs. Still, immersion was a stronger predictor than involvement, as hypothesized. It is also interesting that usability was a predictor of presence, but it was not for flow. It could have to do with the damaging nature of distractions on immersion and presence. Again, this supports the proposed model when the attentional requirements construct is replaced by usability.

Hypothesis 6 explored how individual differences influenced presence. In addition to presence being the epitome of immersion, it was originally hypothesized that the individual's willingness to suspend disbelief, creative imagination, and immersive / involvement tendencies also were critical. For presence, immersion was the strongest predictor as hypothesized. Interestingly, willingness to suspend disbelief and immersive/involvement tendencies were significant predictors of presence, but creative imagination was not. Sas and O'Hare (2004) reported a correlation between presence and willingness to suspend disbelief of $r = .53$, $p < .01$ (p. 532). The correlation found in this study was positive, but was not as high as the one found by the other researchers ($r = .324$, found in Block 3 of Table 27).

Hypothesis 7 examined the predictors of flow. Involvement, challenge/skill balance, and feedback were the significant predictors, as hypothesized. The many other proposed individual differences did not explain a significant amount of remaining variance in the flow score.

Hypothesis 8 examined the correlational relationship between presence and flow across the four conditions. In the condition that supported both presence and flow, the correlation between the two states was the strongest. In the condition that supported the experiences the least, the correlation was the weakest. This lends support to the idea that, given the right circumstances, either, both, or none of these game engagement states can be experienced.

The purpose of Hypothesis 9a, 9b, 9c, and 9d was to determine how the CAS (Agarwal & Karahanna, 2000) was correlated with the CORE flow mean score (Martin & Jackson, 2008) and selected subscales from the FSS-2 (Jackson & Eklund, 2002). The underlying assumption was that they are the same constructs given their similar theoretical underpinnings. The analysis revealed that there was not a difference between cognitive absorption and flow as measured in this study. They should not be considered opposing, separate states as Brockmyer et al. (2009) suggested in the original game engagement model. It is important to clarify here that this finding pertains to the state measure of absorption (Agarwal & Karahanna, 2000), not the established trait measure (Tellegen & Atkinson, 1974).

Conclusion

As a final look at the R-GEM, Hypothesis 8 was revisited with a series of four stepwise regressions to determine what variables were the most relevant predictors for each of the game engagement states. Table 36 below summarizes which predictor variables were significant at the $p < .01$ and $p < .05$ levels. Examining the predictors significant at $p < .01$, it is clear that several of the hypothesized constructs align with the game engagement states as originally outlined in Figure 8.

Table 36. Significant predictors of game engagement states.

	***p* < .01**	***p* < .05**
Immersion	Presence (+) Usability (+) Creative Imagination (+)	Survival Rating (+) Flow (+) Age (+)
Presence	Immersion (+)	Flow (+) Usability (-) Willing to SD (+)
Involvement	Flow (+) Imm / Involv Tendencies (+) Usability (+)	Hours Playing Minecraft (-) Survival Rating (+)
Flow	Involvement (+) Dispositional Flow (+)	Presence (+)

Initially, the attentional requirements construct was predicted to be the precursor to immersion and involvement, however the analysis for Hypothesis 3b revealed that usability was a better predictor than attentional requirements score. This is interesting because usability was a hypothesized predictor the attentional requirements subscale score, which was supported by Hypothesis 3a. It appears that including attentional requirements as a construct between usability and low-level game engagement is unnecessary, or it may be that the way attentional requirements was measured needs to be refined. Either way, usability is a positive predictor of both immersion and involvement, but not presence and flow. This supports the model proposed in Figure 8 if the construct of attentional requirements is replaced with usability.

Given that immersion and involvement are highly correlated as seen in Hypothesis 4, but not so much that they are predictors for one-another, this further establishes that the constructs are related but separate, again supporting the R-GEM as proposed in Figure 8. Also, the key predictors for presence and flow are immersion and involvement, respectively, as suggested in the R-GEM.

The interrelated nature of the game engagement states is seen in the list of predictors that were significant at the $p < .05$ level. For example, both presence and flow are both predictors of one-another. What is interesting is that immersion and involvement did not also have this relationship. Perhaps these high-level game engagement states are very related, even more so than immersion and involvement. Still, they are not predictors of one-another at the $p < .01$ level and their correlations are not excessively high as seen in Hypothesis 8. This continues to establish that they are indeed separate states.

The exploration of individual differences also revealed how these game engagement states may be influenced by other factors. Immersion is positively predicted by creative imagination, which is extremely interesting because Hypothesis 5b found that presence, which is the outcome of high amounts of immersion, was not predicted by that individual difference. Instead, willingness to suspend disbelief and immersive/involvement tendencies were predictors of presence. The results suggest that creative imagination positively predicts immersion, while immersion and willingness to suspend disbelief together positively predict presence. Regarding involvement, it was positively predicted by immersive/involvement tendencies, while flow was positively predicted by involvement and dispositional tendency to experience flow.

There are several other variables that emerged as significant predictors, however these will not be added to the R-GEM at this time. These predictors are rating of the Survival Mode of

Minecraft, which is the type of game played in this experiment, age, and hours spent playing *Minecraft*. Each of these variables suffers from range restriction. For example, roughly half of the participants were classified as Experts, and these individuals had to have rated Survival Mode a 6 or 7 out of 7 possible points. The age range was also very limited to a span of seven years. Finally, the participants were recruited due to their experience playing the game, with a mean amount of hours reaching into the hundreds. Given this range restriction, further study is needed in a much more varied sample to determine if these are indeed significant predictors.

With the exception of replacing attention with usability and highlighting the even more strongly interrelated nature of flow and presence, the R-GEM proposed in Figure 8 was supported by the findings of this study. Additionally, the eight assertions of game engagement from Table 2 also were supported by these results. The R-GEM is updated below based on the findings of the experiment in Figure 13.

High ratings of usability positively influence low-level game engagement, which is both immersion and involvement. For immersion, it may be that good usability reduces distractions so that the experience seems as natural as possible. For involvement, good usability may be indicative of a higher-quality gameplay experience, which would encourage the player to want to play it, hence increased involvement. Immersion is further increased when the individual has a higher amount of creative imagination, while involvement increases when the individual has high immersive/involvement tendencies. As suggested by Witmer and Singer (1998), immersion and involvement are separate states but do have a reciprocal relationship with one-another.

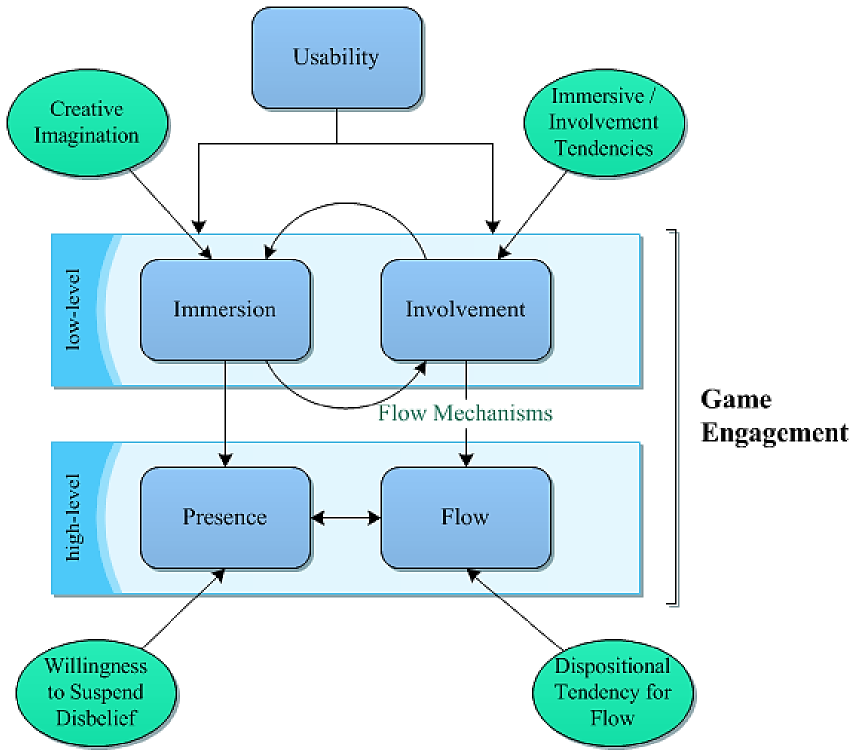


Figure 13. Finalized revised game engagement model.

Both presence and flow are also unique game engagement states. Presence is the result of a large amount of immersion, which is further positively influenced by high amounts of willingness to suspend disbelief. Flow results from involvement as well as the presence of flow mechanisms, such as achieving a balance between the challenge of the game and the skill of the player. Whether an individual experiences flow is also positively influenced by his or her dispositional tendency to experience flow. Not represented here is the finding that immersion and involvement were equally influential in flow. This did not emerge in the final analysis, and may have been an artifact of the high immersion scores, which may be specific to *Minecraft*.

Definitions revisited. Based on the above results, the earlier proposed definitions for immersion, involvement, presence, and flow have been refined to specify the specific individual differences and other mechanisms that were found to be most influential.

Immersion is “a [subjective] state characterized by perceiving oneself to be enveloped by, included in, and interacting with [a video game] that provides a continuous stream of stimuli and experiences” (Witmer & Singer, 1998, p. 227). Immersion requires focused attention on a limited stimulus field and minimized distractions (Csikszentmihalyi, 1990; Sheridan, 1992, 1994; Witmer & Singer, 1998), which can be promoted by the video game system itself. Immersion may be enhanced by the capability of the video game’s technology to provide the player immersive cues. This includes the ability to interact with the video game through a virtual representation of the player (Slater & Wilbur, 1997). Interaction must seem natural with regard to the input mechanisms and the game’s response to the player (Witmer & Singer, 1998; IJsselsteijn et al., 2000). Immersive cues are also strengthened by increasing the “extent and fidelity [and resolution] of sensory information” (Slater & Wilbur, 1997; IJsselsteijn et al., 2000; Wirth et al., 2007). Lacking immersive cues, involvement and individual differences such as creative imagination can mitigate the deficit, thus helping the player to experience immersion (Witmer & Singer, 1998; Sas & O’Hare, 2004; Wirth et al., 2007).

Involvement is a motivational factor (Wirth et al., 2007) regarding gameplay that is “experienced as a consequence of focusing one’s energy and attention on a coherent set of stimuli or meaningfully related activities and events. Involvement depends on the degree of [perceived relevance] that the individual attaches to the stimuli, activities, or events” (Witmer & Singer, 1998, p. 227). Involvement is “increased by [playing video games] that stimulate, challenge, and engage the user either cognitively, physically, or emotionally” (Witmer et al., 2005, p. 299). Involvement has a reciprocal relationship with immersion, where increasing a sense of immersion similarly increases a sense of involvement, and vice-versa (Witmer & Singer, 1998). Involvement is also positively influenced by the general tendency to become immersed/involved.

Presence is a state of "conviction of being located in [the game] environment" (Wirth et al., 2007, p. 495). It is "a binary experience, during which perceived self-location and…perceived action possibilities are connected to [the game environment], and mental capacities are bound by the [the game environment] instead of reality" (Wirth et al., 2007, p. 497). Presence occurs when the individual is experiencing a large amount of immersion, but it is also positively influenced by the individual's willingness to suspend disbelief.

Flow is the "optimal experience" (Csikszentmihalyi, 1990, p. 39), in which attention is focused on a limited stimulus field provided by the video game and where outside distractions have been minimized, resulting in a merging of action and awareness where "that person's attention is completely absorbed" (Csikszentmihalyi, 1990, p. 53) and playing the game becomes "seemingly effortless" (Csikszentmihalyi, 1990, p. 54). This occurs in an activity with clear goals, immediate feedback as one works toward those goals, and a progressive level of challenge which encourages an individual to increase their skills (challenge/skill balance) so that the activity is not boring or frustrating. This results in a state of concentration, in which only the activity matters, a loss of self-consciousness where an individual loses themselves in the activity "with no room for self-scrutiny" (Csikszentmihalyi, 1990, p. 63), and a sense of control in which an individual feels confident in their abilities and does not worry about failure. The individual experiences a sense of time distortion, in which time either slows down or speeds up. The end result is the euphoric, intrinsically-motivating autotelic experience which individuals seek to recreate “simply because the doing itself is the reward” (Csikszentmihalyi, 1990, p. 67). The likelihood of experiencing flow is increased when flow mechanisms are present, such as challenge/skill balance, and when an individual has a high dispositional tendency toward flow.

Measurement recommendations. Based on the findings from the study, tentative guidelines as how to best measure the different game engagement states, and what other measures should be also considered, are listed below in Table 37.

Table 37. R-GEM measurement recommendations.

	State Measure	**Supporting Measure**
Immersion	PQ (version 3) - Immersion subscale	Creative Imagination Scale *t*-score
Involvement	Game Involvement Scale	Immersive Tendencies Questionnaire
Presence	Presence Count Survey	Willingness to Suspend Disbelief Scale
Flow	CORE Flow State Scale	SHORT Dispositional Flow Scale

Immersion can be measured using items from Witmer et al.’s (2005) version 3 of the Presence Questionnaire (PQ). The specific item numbers used in this study to formulate the new immersion subscale are listed in Table 13. The PQ produces total scores rather than means, and therefore total scores were used in this study. Future work could focus on further refining the items for this new immersion subscale of the PQ and determining if a mean or a total score should be used. Creative imagination as measured by Wilson and Barber’s (1978) Creative Imagination Scale’s *t*-score should also be collected since it is also a predictor of immersion and may explain why (or why not) players experienced this state.

Involvement can be measured by totaling the item scores from the R PII (Zaichkowsky, 1994), four of the new TAM-influenced items introduced in this study (Davis, 1989), and item 18 from the PQ version 3 (Witmer et al., 2005, see p. 302). This combination of items could be simply called the Game Involvement Scale. Involvement is also influenced by immersive/involvement tendencies, so it is also beneficial to collect these data using the ITQ (Witmer & Singer, 1998).

Presence can be measured using the presence count method proposed by Slater and Steed (2000), in which participants rate items on a 7-point scale related to experiencing presence, and for every item that is endorsed a 6 or higher, a point is added to the total count. The items used in this study included the five items from Slater and Steed's (2000) experiment and the five items Witmer et al. (2005) provided to augment the PQ (items A1, A2, A3, and A4, see p. 311). These nine items create the Presence Count Survey. At the completion of the survey, it is also useful to ask the participant to report the percentage of time they felt that they were present (Slater & Steed, 2000). Given the high correlation between presence estimate and presence count, the presence estimate may be used in lieu of completing the nine other items if time and/or number of items the participant has to complete is a concern. Willingness to suspend disbelief appears to be an important individual difference to consider when measuring presence, so using the single item from Sas and O'Hare (2004) averaged with the three new items proposed in this work can be used to create a mean score. These items comprise the new Willingness to Suspend Disbelief Scale.

Flow can be measured with the 10-item C FSS (Martin & Jackson, 2008), which is a mean score. This is especially meaningful as a viable replacement for the FSS-2 (Jackson & Eklund, 2002), and S FSS (Jackson et al., 2008) was needed for measuring flow in the gaming context. Given its role as a significant predictor, dispositional tendency to experience flow should be measured by the S DFS (Jackson et al., 2008), which is also a mean score. Still, if the purpose of the study is to examine the individual mechanisms of flow rather than the flow experience itself, the FSS-2 can still be used (Jackson & Eklund, 2002), but perhaps with caution.

Finally, if interested in usability, the SUS (Brooke, 1996) raw score can also be collected to measure the precursor to low-level game engagement.

Limitations

This study had its limitations. For example, the immersion manipulation could have been made stronger through the use of an HMD and immersive controller. When the experiment was being designed, two options were available to fit this need: A developer's version of the Oculus Rift (model DK1), which is a commercial-grade HMD, and a Razer Hydra, a controller which allows the player to interact with the game in naturalistic ways (e.g., swing the controller to swing an in-game axe). At the time, an unlicensed mod to *Minecraft* was created to allow the game to be played with the Oculus Rift and Razer Hydra. This would have provided a very immersive experience as the HMD offered a 3D view of the virtual environment across a 110 degree field of view (Popa, 2014), and the controller allowed for highly naturalistic inputs. However, this version of *Minecraft* was not able to be modded to have shaders. Therefore, the only way to further increase immersion would be to use higher resolution textures. Also, the modded version of *Minecraft* that worked with the system immersion-promoting technologies had a flaw where if the player died, the game appeared to freeze. To proceed with the game required the player to break immersion.

This early version of the Oculus Rift suffered from some design flaws as well. For example, even with the lenses set the furthest away from the eye as possible, eye lashes would smudge the display. This was extremely distracting during gameplay. Also, the Razer Hydra was somewhat difficult to use, especially for a player who had never used it before. Learning to use it would require actually looking at the controller at first, which would have not been possible had they been wearing the HMD. Therefore, at the time of this study, the available technology was not mature enough to be used in an experimental setting since all of these issues amounted to distraction. Given that distracting equipment (Slater & Wilbur, 1997) and increased interface awareness (Witmer & Singer, 1998) are damaging to game engagement, other ways of encouraging the experience of immersion were selected, such as the use of high-resolution textures and advanced shaders.

Another limitation of this work was that some of the measures were previously unexamined. Even though the PQ (Witmer & Singer, 1998; Witmer et al., 2005) had been used extensively in past research, the subscales used to represent immersion and attentional requirements were entirely new. Additionally, the measures for willingness to suspend disbelief and for game involvement combined both new and old items. While these measures were grounded in theory, they would all benefit from further, more focused validation studies.

Finally, this study only examined one game–*Minecraft*. The results may only generalize to this very specific genre, which is a sandbox adventure game that features heavy elements of crafting and requires some amount of creativity. It is possible that the results may have been different in a goal-driven game, or if the game involved elements of multiplayer-based competition. Even having participants play different modes in *Minecraft* may have resulted in a different pattern of results. It may be that different types of games (e.g., goal-driven competitive games, sandbox-style god-games, story-focused gaming experiences) may result in different permutations of the R-GEM. The theoretical framework itself should not change, but the strengths of the relationships may, and new relationships between constructs may form while others may fade away.

Future Directions

As mentioned above, there were limitations with this study that can be addressed with follow-up work. As the technology improves, this study can be replicated using a better HMD and more naturalistic control inputs. For example, the Oculus Rift has been recently updated to the second version of the development kit (model DK2), which features a higher resolution and a faster screen refresh rate. The DK1 featured orientation tracking alongside the on-board gyroscope, accelerometer, and magnetometer. The DK2 included this in addition to the newly-added positional tracking. Also, the DK2 added the ability for users to wear eye glasses while using the Oculus Rift (Popa, 2014). For comparison, the DK1 had interchangeable lenses since it was not compatible with eye glasses. This further limited its usefulness in an experimental setting.

Aside from the Oculus Rift, more HMD technologies are becoming commercially available. Project Morpheus is Sony's response to the Oculus Rift. This virtual reality system is geared toward consoles, the PlayStation 4 specifically, and has many of the same features as the DK2, with the added benefit of built-in 3D audio (Gilbert, 2014). Using the Project Morpheus HMD could be another alternative to replicating this study since *Minecraft* has been deployed to consoles.

Microsoft recently bought Mojang, the developers of *Minecraft*. Microsoft also has announced its new project, the HoloLens, which is a wearable augmented reality device. The earliest demos of the device featured a *Minecraft*-like game that was projected into the actual environment—which was, in this demo, the wearer's living room—from a top-down view (Plante, 2015). This stable version of *Minecraft* coupled with virtual reality technologies that may be better than the current Oculus Rift may make it possible to replicate this study with technologies that support more system immersion. Given that immersion scores were not affected by the manipulation in this study, a follow-up study using an improved HMD will determine whether this was the fault of the experimental manipulation, or if immersion truly is influenced by so many other factors that simply providing the opportunity via technology does not guarantee that immersion is experienced.

An aspect of this study worth revisiting is the selection of the sample: These participants were all experienced *Minecraft* players. This was done for two reasons: to minimize the training needed for the completion of the experiment, which reduced the overall session time to prevent fatigue and ensured participants were not distracted by needing to learn how to play the game; and, to increase the likelihood that high-level game engagement would occur, just as Brockmyer et al. (2009) had done in their study.

Game-specific familiarity and expertise may have shaped the results found in this study, and should be explored in more detail regarding the construct of game engagement. For example, Mizobata, Silpasuwanchai, and Ren (2014) found that general gaming experience and immersion, as measured by the Game Engagement Questionnaire—which was a different measure than the one developed by Brockmyer et al. (2009) so it is unknown whether immersion was truly being measured according to the definition in this work—was significantly, positively, and strongly correlated for one Kinect game (r = .68; p. 61), but not for the other motion-based games explored in the study. Participants in the present study became highly immersed in

Minecraft in spite of its graphical limitations. It was discussed previously that familiarity and self-selection may have increased immersion in the non-immersive virtual environment through desire to become involved, which has a positive reciprocal relationship with immersion.

Along these lines, it may be that these *Minecraft* ‘experts’ needed fewer details to achieve the same level of immersion since they were so involved. For example, Sims, Moshell, Hughes, Cotton, and Xiao (2001) explored participants’ ability to recreate virtual trees from memory. They found that participants were most accurate when recreating symmetrical virtual trees. The researchers supposed that this was “because they could use a pre-existing ‘tree schema’ to fill in forgotten or distorted information” (p. 1938). This is similar to the idea that having a greater amount of creative immersion will result in more immersion; if the individual is able to fill in missing details from existing schemata, the immersion decrement will be mitigated. Perhaps since these experienced players were so willing to play, they focused more on gameplay than the virtual environment, and let their own previous knowledge of how the virtual environment should look fill in any gaps—and arguably this is something that they do on a regular basis since they choose to play this low-fidelity game—which resulted in the high immersion scores found across conditions. Therefore, the role of expertise as it influences game engagement should be examined in future studies.

Another avenue for continued work is measure refinement. As mentioned above, several of the scales—specifically the newly-minted Game Involvement Scale and the Willingness to Suspend Disbelief Scale—used in this study were composed of new items. For each of these scales, and their related constructs, follow-up studies focused on the construct itself as well as item refinement is warranted.

Two full subscales for immersion and attentional requirements, as well as an item donated to the Game Involvement Scale, were derived from the PQ (version 3; Witmer et al., 2005). The immersion items could be refined in a study that makes use of more advanced HMD technology. The attentional requirements score was not a significant predictor of low-level game engagement as originally hypothesized. It may be that this measure needs improvement. For example, a much larger item pool can be created and then a study focused on attentional requirements alone will help refine the measure. This study, while it sought to minimize distractions, did not introduce them. The refined measure can be validated with a study that introduces distractions. It also may be that, for experienced *Minecraft* players, the attentional requirements factor of the game engagement model was less important. These players habitually play, by choice. They may be better at blocking out distractions. Therefore, this study could also be replicated with a greater range of experience levels (e.g., more novices and non-players) to determine if this was a factor.

Once the measures have been refined and validated, additional studies could compare the R-GEM and its measurement set against other measures as a means for additional validation. For example, one follow-up study could explore how different aspects of the R-GEM, such as the presence count and presence estimate, correlate with physiological measures. This could be the objective measure that Sheridan (1992) called for. Once relationships are established with a theoretically-driven, well-validated model, physiological measurement can be used to track game engagement in real-time (e.g., Nacke, Stellmach, Sasse, & Lindley, 2009; Levillain, Orero, Rifqi, & Bouchon-Meunier, 2010). Better understanding of each construct in the model could then mean that a game—in conjunction with a set of physiological sensors—can be developed to recognize and target game engagement states by manipulating game features. For example, increasing a sense of immersion in virtual reality exposure therapy may make treatment more effective, although the relationship has not yet been established (Krijn, Emmelkamp, Olafsson, & Biemond, 2004). A meta-analysis conducted by Parsons and Rizzo (2008) was unable to draw any conclusions about immersion's role in virtual reality exposure therapy effectiveness due to inconsistent reporting and even hinted that the topic of immersion itself was "beset by methodological controversy" (p. 258). Hopefully this work will help alleviate some of that controversy.

The proposed measures and revised game engagement model can also be tested against other up and coming subjective measures, such as the User Engagement Scale (Wiebe et al., 2014) and even the original Game Engagement Questionnaire (Brockmyer et al., 2009). Another scale to compare the R-GEM and its measured against is the Play Experience Scale developed by Pavlas et al. (2012). The R-GEM should also be examined in populations with more variety for the range-restricted demographic variables found in this study (e.g., age, experience).

As mentioned in the limitations section, the R-GEM was only examined with one game, *Minecraft*. It is incorrect to say that the R-GEM is validated simply because the hypothesized relationships held true in this singular instance. Once the measures are refined, a large-scale, naturalistic comparison study could be conducted between *Minecraft*, the single-player sandbox set on Survival Mode, and Riot Games' *League of Legends*, which is a highly-competitive, multiplayer, goal-oriented online game. Given the important role of goals in flow theory (Csikszentmihalyi, 1990), and since the Clear Goals subscale of the FSS-2 was not a significant predictor of CORE flow in Hypothesis 7, exploring the R-GEM using a highly goal-driven game will contribute additional information about the model and the significant predictors of flow. It may be that game genre needs to be added to the R-GEM, or that there might be different versions of the R-GEM that apply to different types of games. Therefore, the R-GEM should be explored in a variety of game types to determine how genre influences the model.

Once further validated, other potential variables could be explored to see how they are related to game engagement. For example, Canossa (2012) explored how player behaviors in *Minecraft* could potentially be linked to the 16 basic human desires as measured by the Reiss Motivational Profile Estimator. Three experienced game designers assigned *Minecraft* player behaviors to the measured desires. A strong association was found between a player's level of curiosity and how far they traveled in-game. A similar analysis comparing in-game behaviors to game engagement states could also be conducted.

Finally, once the R-GEM is even more well-established, future work can also examine how game engagement relates to other outcomes, such as aggression and learning. The influence of gameplay on aggression has been a long-discussed, controversial topic with much of the research failing to establish a relationship between violent video game and aggressive acts. Indeed, this topic was the impetus for one of the more influential papers in this work (Brockmyer et al., 2009). So far, the scientific literature has not established an irrefutable link between gaming and aggression, and the debate continues (Anderson et al., 2010; Ferguson, 2013; Ferguson et al., 2013).

Rather than focusing on aggression due to prolonged play of violent video games, Przybylski, Deci, Rigby, and Ryan (2014) instead sought to determine how playing a game may make an individual more aggressive in the short-term. Przybylski et al.'s (2014) approach was grounded in self-determination theory (SDT). According to SDT, human behavior is motivated by three core needs: competency, autonomy, and relatedness (c.f., Ryan & Deci, 2000). The researchers' earlier work found that violent content was not found to be the source of an individual's motivation to play a game, however need satisfaction was (Przybylski, Ryan, & Rigby, 2009; Przybylski, Rigby, & Ryan, 2010)

Przybylski et al. (2014) argued that a lack of competence, which they defined as "the experience of efficacy" (p. 441), should result in higher levels of aggression when games are manipulated to be too difficult, as this difficulty threatens their sense of game-specific competency. Essentially, this is the interplay of VGSE and usability: When a game is unusable, it lowers game-specific VGSE after play. This results in frustration, which is, of course, purported to be the opposite of flow. Przybylski et al. (2014) conducted seven separate studies that spanned different levels of violent and aggressive game content, as well as manipulated level of control complexity and examined how game experience affected game enjoyment and short-term aggression. Essentially, they found that poor usability lowered their participants' perceived competence while playing, which resulted in short-term increases in aggression. The level of aggression was not related to the level of violence and opportunities for aggressive acts present in the game. Furthermore, overall enjoyment of the game was negatively associated with aggression.

Przybylski et al.'s (2014) findings about usability are interesting. They also reaffirm usability's place as important to game engagement. It has interesting implications for VGSE and how that may affect other parts of the revised game engagement model. For example, VGSE was found to be a positive predictor of flow by Pavlas (2010). This may be especially relevant, as

Przybylski et al.'s (2014) work is based on SDT, which evolved from the foundational work of intrinsic motivation, which directly shaped Csikszentmihalyi's (1990) theory of flow. Incorporating aspects of SDT as they relate to involvement and flow into the R-GEM should prove to be an interesting avenue for future research.

As mentioned in the introduction to this paper, a well-designed serious game is an effective learning tool (Connolly et al., 2012). Several researchers are interested in how different elements of game engagement make for more effective serious games (see Procci and Bowers, 2014). For example, Pavlas (2010) found that flow was an independent predictor of game-based learning. Still, it is impossible to draw accurate conclusions about the relationship between the different constructs of game engagement and learning outcomes in serious games without first thoroughly validating the game engagement model.

An example given earlier in this work was from Admiraal et al. (2011). The researchers were interested in how flow impacted student learning outcomes. Any casual examiner of the literature, who may not be doing anything more than skimming abstracts, might only read the statement that "Flow was shown to have an effect on their game performance, but not on their learning outcome" (Admiraal et al., 2011, p. 1185). This would be misleading, as they defined flow in their study as "engagement of each team in the game play itself" and measured it by observing whether the participant was merely "working on the completion of the assignment" or "using the tools" (Admiraal et al., 2011, p. 1189). They did this despite their clear understanding of the construct as evidenced by their literature review. Their operationalization of flow was different from what is generally accepted in the literature. If this study was included in a meta-analysis on flow's effect on learning, the findings could be confounded due to the improper use of the word ‘flow’. Now, with the strict definitions and measurement guidelines provided by the R-GEM, a study like theirs could be replicated and more accurate conclusions drawn.

The goal of this dissertation was to be a starting point for researchers interested in the constructs under the umbrella of game engagement, to include immersion, involvement, presence, and flow. Proposed here is a framework to explain these oft confused and misused constructs. Like Sheridan (1992) originally stated, progress regarding any of these states is impossible without strong operational definitions and reliable methods of measurement. This revised game engagement model was based strictly on theory and leveraged the work of researchers over the past two and half decades. Definitions were distilled and incorporated into an overarching model, which was then tested. Even with newly-proposed measures for some of the constructs, support for the R-GEM was found. Moving forward, these clear definitions, suggested measurement techniques, and model of game engagement interactions will ensure consistent usage of these constructs, which will improve the scientific rigor of gaming research.

APPENDIX A: OUTLINE OF IMMERSION, INVOLVEMENT, PRESENCE, AND FLOW

Immersion: A subjective feeling of being enveloped by the games' stimuli and experiences.

Requirements of immersion

- Focused attention on a limited stimulus field
    - Can be fostered by minimizing internal distractions (ability to master controls and basics of gameplay, understandable interfaces, lack of glitches) and external (e.g., background noises) distractions through high-quality, usable designs and game peripherals that limit external sensory interference (e.g., HMDs, headphones)

Immersion is supported by a balance of…

- The technological capabilities of the system
    - Realistic interactivity and input mechanisms
        - The player can control the senses of an in-game representation to explore the game
        - The player can interact with the game, which responds realistically in-kind
        - There is a perfect match between the player's input and the in-game response (e.g., proper optic flow when moving, no lag)
        - Naturalistic inputs (the player turns their head in real-life, mirrored in-game)
    - Higher fidelity and number of senses provided by the game
- Motivation and personal *involvement* in wanting to be immersed in a game, reciprocal relationship
- Individual differences
    - Willingness to suspend disbelief
    - Creative imagination
    - Immersive tendencies

Involvement: Personal motivation to play the game.

- Motivation and personal *involvement* in wanting to be immersed in a game, reciprocal relationship

Requirements of involvement…

- Focused attention on a limited stimulus field
    - Sheer determination on the part of the player (if the player wants to play, they will be motivated to focus attention on their own)
- Goals
- Immediate feedback
- Individual differences
    - Motivation to play
    - Willingness to suspend disbelief
    - Creative imagination

Presence: Accepts the game environment as their primary egocentric reference frame (PERF), meaning that the player believes they are physically located in the game; is a binary state.

Presence is caused by…

- Very, very high levels of overall immersion, a combination of:
    - Immersive cues provided by the video game (e.g., high-fidelity, increased field of view, HMD, naturalistic inputs, etc.)
        - Lots of immersive cues makes it easier to confirm and needs more distractions to take away
        - Even if immersive cues are low, can be mitigated by other factors
    - Personal involvement and motivation to be present
    - Creative imagination and willingness to suspend disbelief

Flow: The optimal experience of being 'in the zone', a goal-driven high where time is distorted.

- Shares in common with immersion: Merging of Action and Awareness (effortlessly absorbed) , Concentration, Control, Loss of Self-Consciousness
- Shares in common with involvement: Goals and Immediate Feedback
- Unique to the flow experience: Autotelic Experience (intrinsically rewarding and enjoyable, 'in the zone'); A distorted perception of time

Requirements of flow…

- Shares with immersion: Need for focused attention, limited stimulus field, minimized distractions
- Shares with involvement: Need for goals and immediate feedback
- Unique to flow: Subjective appraisal of a Challenge/Skill Balance
- Individual differences
    - Motivation to play
    - Willingness to suspend disbelief
    - Creative imagination
    - Dispositional flow
    - Trait absorption
    - Age
    - Video game self-efficacy

APPENDIX B: MEASURES

Demographic Surveys

- *Basic demographic information* – pre-existing RETRO questionnaire
    - Age (years)
    - Gender (male / female / other / no response)
    - Ethnic background (White / Black / Asian / Hispanic / Other)
    - Year in school (Freshmen / Sophomore / Junior / Senior / 5th + Year Senior)
    - Primary language (English / Spanish / Other)
- *Gaming preferences and behaviors*- pre-existing RETRO questionnaire with new items
    - Comfort with computers (single item score, 1-7)
    - Comfort with games (single item score, 1-7)
    - Games played (number of hours / week)
    - How games are played (computer / console / cell phone / tablet / handheld gaming device / N/A)
    - How the majority of games are played (computer / console / cell phone / tablet / handheld gaming device / N/A)
    - List top three favorite games (qualitative, will be coded by genre)
    - Ever played a game using an HMD? (yes / no)
    - Ever played Minecraft? (yes / no)
    - Hours spent playing Minecraft (hours total)
    - Ever played Minecraft with an HMD? (yes / no)
    - Rating of general game enjoyment (single item score, 1-7)
    - Rating of Minecraft enjoyment (single item score, 1-7)
    - Comfort with Minecraft (single item score, 1-7)
- *Video game self-efficacy*
    - Video Game Self-Efficacy Scale (VGSES; Pavlas, Heyne, Bedwell, Lazzara, & Salas, 2010): total video game self-efficacy (VGSE) score (10-40)
- *Minecraft screening*
    - How Minecraft is played (computer, console, cell phone, tablet)
    - How the majority of Minecraft is played (computer, console, cell phone, tablet)
    - What versions of Minecraft have been played
    - What version of Minecraft they prefer to play
    - Normal screen size
    - Normal resolution
    - Rating of Creative Mode (single item score, 1-7)
    - Rating of Survival Mode (single item score, 1-7)
    - Rating of Hardcore mode (single item score, 1-7)
    - Preferred level of difficulty (peaceful, easy, normal, hard, N/A)
    - If they play Minecraft in windowed mode (yes, no, not sure)
    - If they play Minecraft in full-screen mode (yes, no, not sure)
    - If they play in Survival Mode (yes, no, not sure)
    - If they play in Creative Mode (yes, no, not sure)
    - If they to play in Hardcore Mode (yes, no, not sure)
    - If they play with the sound off (yes, no, not sure)

- If they play with headphones (yes, no, not sure)
- If they prefer to play alone (yes, no, not sure)
- If they played with Minecraft texture packs (yes/no, list)
- If they played with Minecraft mods (yes/no, list)
- If they played with Minecraft shaders (yes/no, list)

Individual Differences Influencing Game Engagement

- *Immersive and involvement tendencies*
    - Immersive Tendencies Questionnaire (ITQ; Witmer & Singer, 1998): total immersive and involvement tendencies score (18-126)
- *Dispositional flow*
    - SHORT Dispositional Flow Scale (S DFS; Jackson, Martin, & Eklund, 2008): mean dispositional flow score (1-5)
- *Willingness to suspend disbelief*
    - Willingness to suspend disbelief item (Sas & O'Hare, 2004): item score (1-7)
    - Willingness to suspend disbelief Scale (Sas & O'Hare, 2004 + novel items): mean score (1-7)
- *Creative imagination*
    - Creative Imagination Scale (CIS; Wilson & Barber, 1978): CIS total *t*-score
- *Trait absorption*
    - Tellegen Absorption Scale (TAS; Tellegen & Atkinson, 1974): total score (0-34)

Game Engagement States

- *Attentional requirements*
    - Presence Questionnaire (PQ; Witmer & Singer, 1998): Attentional Requirements subscale total score (7-70)
- *Immersion*
    - Presence Questionnaire (PQ; Witmer & Singer, 1998): Immersion subscale total score (7-112)
- *Involvement*
    - Revised Personal Involvement Inventory (Revised PII; Zaichkowsky, 1994), the Involvement item from the Presence Questionnaire (PQ; Witmer & Singer, 1998), and items influenced from the TAM (Davis, 1989): Total involvement score (21-105)
    - Involvement categorization (based on Zaichkowsky, 1994): low (21-48), medium (49-77), or high (78-105) involvement
- *Presence*
    - Items from Slater and Steed (2000) and Witmer, Jerome, and Singer (2005) using Slater and Steed's presence count method: total presence count (0-9)
    - Presence estimation (Slater & Steed, 2000): presence percentage

    - Presence classification (Slater & Steed, 2000): no (0% - 50%) or yes (51% - 100%)
- *Flow*
    - CORE Flow State Scale (C FSS; Martin & Jackson, 2008): Mean core flow score (1-5)
    - Flow State Scale-2 (FSS-2; Jackson & Eklund, 2002): Mean total flow score (1-5)
        - Challenge/Skill Balance FSS-2 subscale mean (1-5)
        - Merging of Action and Awareness FSS-2 subscale mean (1-5)
        - Clear Goals FSS-2 subscale mean (1-5)
        - Unambiguous Feedback FSS-2 subscale mean (1-5)
        - Concentration on the Task at Hand FSS-2 subscale mean (1-5)
        - Sense of Control FSS-2 subscale mean (1-5)
        - Loss of Self-Consciousness FSS-2 subscale mean (1-5)
        - Transformation of Time FSS-2 subscale mean (1-5)
        - Autotelic Experience FSS-2 subscale mean (1-5)

Model-validating variables

- *Usability*
    - System Usability Scale (SUS; Brooke, 1996): Usability percentile
- *Simulator sickness*
    - Simulator Sickness Questionnaire (SSQ, Kennedy, Lane, Berbaum, & Lilienthal, 1993): Simulator Sickness Total Severity score percentile
- *Cognitive absorption*
    - Cognitive Absorption Scale (CAS; Agarwal & Karahanna, 2000): Mean cognitive absorption (1-7)
        - Temporal Dissociation CAS subscale mean (1-7)
        - Focused Immersion CAS subscale mean (1-7)
        - Heightened Enjoyment CAS subscale mean (1-7)
        - Control subscale CAS mean (1-7)
        - Curiosity subscale CAS mean (1-7)

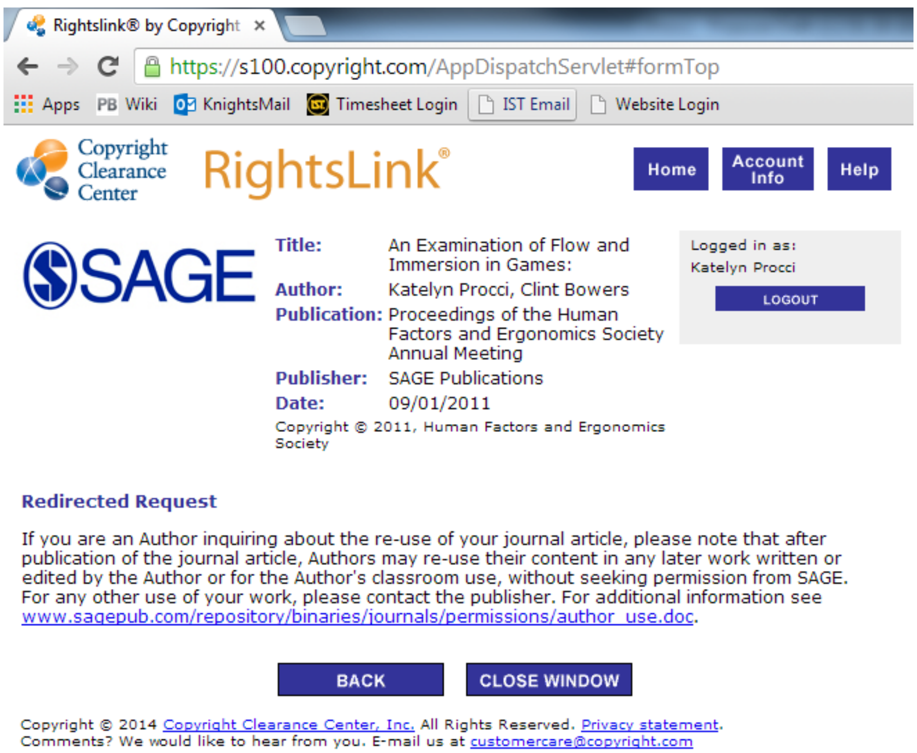
APPENDIX C: COPYRIGHT PERMISSIONS

The use of multiple text excerpts warranted securing usage permissions so that they may be included in this work. Items requiring copyright permission for use in this work included:

- Excerpt from Procci & Bowers, 2011, p. 2183
- Excerpt from Csikszentmihalyi, 1975b
- Excerpt from Csikszentmihalyi, 1990
- Excerpts from Slater & Wilbur, 1997
- Excerpts from Witmer & Singer, 1998
- Excerpts from Brockmyer et al., 2009
- Usage of the Trait Absorption Scale
- Usage of the Flow State Scale – 2, Short Dispositional Flow Scale, and Core Flow Scale

---

Permission for an Excerpt from Procci and Bowers, 2011


Rightslink® by Copyright

https://s100.copyright.com/AppDispatchServlet#formTop

Apps PB Wiki KnightsMail Timesheet Login IST Email Website Login

Copyright Clearance Center RightsLink®

Home | Account Info | Help

SAGE

Title: An Examination of Flow and Immersion in Games:

Author: Katelyn Procci, Clint Bowers

Publication: Proceedings of the Human Factors and Ergonomics Society Annual Meeting

Publisher: SAGE Publications

Date: 09/01/2011

Copyright © 2011, Human Factors and Ergonomics Society

Logged in as:
Katelyn Procci

LOGOUT

Redirected Request

If you are an Author inquiring about the re-use of your journal article, please note that after publication of the journal article, Authors may re-use their content in any later work written or edited by the Author or for the Author's classroom use, without seeking permission from SAGE. For any other use of your work, please contact the publisher. For additional information see www.sagepub.com/repository/binaries/journals/permissions/author_use.doc.

BACK | CLOSE WINDOW

Copyright © 2014 Copyright Clearance Center, Inc. All Rights Reserved. Privacy statement.
Comments? We would like to hear from you. E-mail us at customercare@copyright.com

Permission for an Excerpt from Csikszentmihalyi, 1975b


Rightslink® by Copyright

https://s100.copyright.com/AppDispatchServlet#formTop

Apps PB Wiki KnightsMail Timesheet Login IST Email Website Login

Copyright Clearance Center RightsLink®

Home | Account Info | Help

SAGE

Title: Play and Intrinsic Rewards:

Publication: Journal of Humanistic Psychology

Publisher: SAGE Publications

Date: 07/01/1975

Copyright © 1975, SAGE Publications

Logged in as:
Katelyn Procci
LOGOUT

Gratis

Permission is granted at no cost for sole use in a Master's Thesis and/or Doctoral Dissertation. Additional permission is also granted for the selection to be included in the printing of said scholarly work as part of UMI's "Books on Demand" program. For any further usage or publication, please contact the publisher.

BACK | CLOSE WINDOW

Copyright © 2014 Copyright Clearance Center, Inc. All Rights Reserved. Privacy statement.
Comments? We would like to hear from you. E-mail us at customercare@copyright.com

Permission for an Excerpt from Csikszentmihalyi, 1990

Adult Trade
Subsidiary Rights Dept.

10 East 53rd Street
New York, New York
10022-5299

Telephone 212 207-7000
Fax 212 207-7902

HarperCollins*Publishers*

MARCH 14, 2014

KATELYN PROCCI
3342 MISSION BAY BLVD. APT 148
ORLANDO, FL 32817

Re: MIHALY CSIKSZENTMIHALYI/ FLOW: THE PSYCHOLOGY OF OPTIMAL EXPERIENCE

Dear KATELYN PROCCI,

Permission is granted for the use of material from page 49 as specified in your application for use in your dissertation. The credit line should include title, author, copyright year and copyright holder(s) and notice that the material has been reproduced by permission of HarperCollins Publishers. Permission is granted for a maximum of five copies. Permission is not granted for *publication* in any other medium.

For electronic rights, please contact: BROCKMAN INC. 5 E 59TH STREET, NEW YORK, NY 10022 (212) 935-8900 http://www.brockman.com

Sincerely,

HarperCollins Publishers
Permissions Dept.

Permission for Excerpts from Slater and Wilbur, 1997

RE: MIT Press Journals Permissions

Nicholas Yeaton <nyeaton@MIT.EDU>

Thu 2/13/2014 11:08 AM

To:Katelyn Procci <kprocci@knights.ucf.edu>;

Dear Katelyn Procci,

We are happy to grant non-exclusive, one time use permission to reprint this excerpt in your dissertation. As this is a non-commercial/academic publication, there is no fee in this case. Please be sure to cite the original source of this excerpt in your dissertation.

Please feel free to contact me with any additional questions.

Sincerely,
-Nick Yeaton

Nick Yeaton
Permissions Assistant
The MIT Press
55 Hayward Street
Cambridge, MA 02142
nyeaton@mit.edu l (617)-258-0591

From: Katelyn Procci [mailto:kprocci@knights.ucf.edu]
Sent: Tuesday, February 11, 2014 3:18 PM
To: journals-rights@mit.edu
Subject: MIT Press Journals Permissions

User entered information:

A. Permissions license should be granted and invoiced to the following:

Requestors (Your) Name: Katelyn Procci
Institution/Organization/Business: University of Central Florida
Mailing Address: UCF Dept. of Psychology, 4000 Central Florida Blvd
City and State: Orlando, FL
Postal/Zip Code: 32816
Phone number: 540-842-3924
Fax number:

E-mail address: kprocci@knights.ucf.edu

B. Full Citation of MIT Journal material to be used:

MIT Press Journal Title: Presence
Article: A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments
Author(s): Slater, M., & Wilbur, S.
Volume/Issue/Year: 1996, 6(6)
Page Range of full article: 603–616
If Excerpt or Figure only, please specify: Short text excerpts (~25 words)

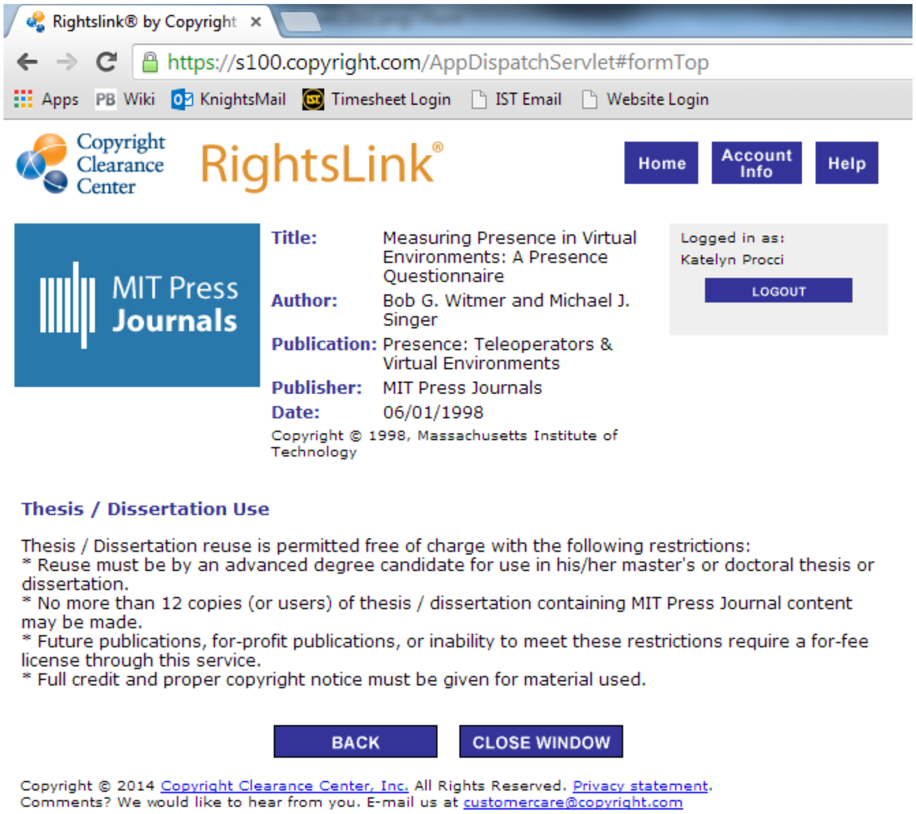
C. FOR REPUBLICATION USE: New Publication Information (if relevant):

Request Reference Code, if any:
Title of New Book/Journal:
Title of New Chapter/Article: The subjective gameplay experience: An examination of the revised game engagement model
Anticipated date of publication: August 2014
Initial Estimated Print Run:
Publisher: University of Central Florida (dissertation)
Editors/Authors: Katelyn Procci
Will the new work be in print format? no
Will the new work be in electronic format? yes

D. FOR CLASSROOM/PHOTOCOPYING USE (if relevant):

Request Reference Code, if any:
Professor Name:
Course Name and/or Number:
Institution/School:
Number of Copies:
Semester to be used:
Will the material be available in print format? no
Will the material be available in electronic format? no

Permission for Excerpts from Witmer and Singer, 1998



Rightslink® by Copyright ×

https://s100.copyright.com/AppDispatchServlet#formTop

Apps PB Wiki KnightsMail Timesheet Login IST Email Website Login

Copyright Clearance Center RightsLink®

Home | Account Info | Help

MIT Press Journals

Title: Measuring Presence in Virtual Environments: A Presence Questionnaire

Author: Bob G. Witmer and Michael J. Singer

Publication: Presence: Teleoperators & Virtual Environments

Publisher: MIT Press Journals

Date: 06/01/1998

Copyright © 1998, Massachusetts Institute of Technology

Logged in as:
Katelyn Procci

LOGOUT

Thesis / Dissertation Use

Thesis / Dissertation reuse is permitted free of charge with the following restrictions:
* Reuse must be by an advanced degree candidate for use in his/her master's or doctoral thesis or dissertation.
* No more than 12 copies (or users) of thesis / dissertation containing MIT Press Journal content may be made.
* Future publications, for-profit publications, or inability to meet these restrictions require a for-fee license through this service.
* Full credit and proper copyright notice must be given for material used.

BACK | CLOSE WINDOW

Copyright © 2014 Copyright Clearance Center, Inc. All Rights Reserved. Privacy statement.
Comments? We would like to hear from you. E-mail us at customercare@copyright.com

Permission for Excerpts from Brockmyer et al., 2009

view.info.copyright.com/?j=fe5c10757264047570128&m=fecd15717160

Apps | PB Wiki | KnightsMail | Timesheet Login | IST Email | Website Login

RightsLink

Copyright Clearance Center

Thank You For Your Order!

Dear Ms. Katelyn Procci,

Thank you for placing your order through Copyright Clearance Center's RightsLink service. Elsevier has partnered with RightsLink to license its content. This notice is a confirmation that your order was successful.

Your order details and publisher terms and conditions are available by clicking the link below:
http://s100.copyright.com/CustomerAdmin/PLF.jsp?ref=5a30ecb6-9ce9-48b6-93ee-e0907d6788c4

Order Details
Licensee: Katelyn Procci
License Date: Feb 11, 2014
License Number: 3326060547471
Publication: Journal of Experimental Social Psychology
Title: The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing
Type Of Use: reuse in a thesis/dissertation
Total: 0.00 USD

To access your account, please visit https://myaccount.copyright.com.

Please note: Online payments are charged immediately after order confirmation; invoices are issued daily and are payable immediately upon receipt.

To ensure we are continuously improving our services, please take a moment to complete our customer satisfaction survey.

B.1:v4.2

+1-877-622-5543 / Tel: +1-978-646-2777

This email was sent to: **kprocci@knights.ucf.edu**

Please visit Copyright Clearance Center for more information.

This email was sent by Copyright Clearance Center
222 Rosewood Drive Danvers, MA 01923 USA

To view the privacy policy, please go here.

Permission for Usage of the Trait Absorption Scale

UNIVERSITY OF MINNESOTA

University of Minnesota Press

Request for License to Reproduce a Portion of a Test Instrument in Research

Date: January 16, 2014

Organization/Name of individual requesting license: Katelyn Procci

Complete Mailing address: Department of Psychology, University of Central Florida, 4000 Central Florida Boulevard., Orlando, FL 32826

Phone: 540-842-3924 E-Mail: kprocci@knights.ucf.edu

Investigator:

Name Katelyn Procci
Affiliation University of Central Florida
Degree MS, Modeling & Simulation, and a doctoral student in Human Factors

Advisor (if investigator is student):

Name Clint Bowers
Affiliation University of Central Florida
Degree PhD, Clinical Psychology

Nature of License

Instrument: MPQ Standard (Tellegen)

Full description of requested material to be reproduced (If relevant, please provide scale name or individual test booklet item numbers):

- Instructions, items, and scoring information for the Absorption scale (primary trait scale AB, 34 items total) from the Multidimensional Personality Questionnaire Standard

This material to be reproduced in (please check all that apply):

_X_ Research Instrument

_X_ Dissertation/Thesis -- Will be used FOR a dissertation, but items will not be reproduced IN the actual text of the dissertation.

__ Other (include electronic or alternative medium, such as the Internet). Explain below.

Please Note: The Press *does not* allow for the reproduction of test items in dissertations.

Mill Place, Suite 290
111 Third Avenue South
Minneapolis, MN 55401-2520
Office: 612-627-1970
http://www.upress.umn.edu

University of Minnesota Press

Nature of Research

Title of Study: An Examination of Game Engagement

Detailed description of study, including major hypotheses and rationale for use of the instrument (approximately 200 words):

The purpose of the study is to examine constructs related to the subjective experience of gameplay. A model of game engagement has been developed that encompasses these constructs (specifically, attention, immersion, involvement, presence, and flow). The purpose of this work is to test the efficacy of this model. Participants who are both novices and experienced gamers will be recruited. They will all play the same game, however half of the participants will play the game using an HMD and an immersive controller, while the other half will play the game normally. Lots of data will be collected pertaining to the different experiential constructs. Based on the model, specific relationships between constructs will be tested. Additionally, it is hypothesized that individual differences, to include trait absorption, influence these experiential states.

Detailed description of subjects to be tested, including number. age, gender, marital status, if a member of a specific group (for example, nonclinical community or student population; clinical inpatient or outpatient):

We will recruit, at maximum, 400 undergraduate students from the University of Central Florida through the Department of Psychology's participant pool. The typical age range for these participants is between 18 and 22 years, and most of who are single. We will recruit both males and females.

Study Site(s): University of Central Florida

Length/Duration of study:

1-2 hour sessions, study will take course over the Spring 2014 – Fall 2014 semester.

Please explain in detail how do you intend to score your results:

I believe this scale returns a score based on 34 endorsements. I could use this as a rate.

Ideally, I would use the absorption score in a regression analysis, along with other variables such as video game self-efficacy, to determine if and how scores for the experiential constructs of interest can be predicted.

Mill Place, Suite 290
111 Third Avenue South
Minneapolis, MN 55401-2520
Office: 612-627-1970
http://www.upress.umn.edu

UNIVERSITY OF MINNESOTA

University of Minnesota Press

For internal use only

____ Permission for the above request is denied.

__x__ One-time, nonexclusive permission for the above request is granted with the following special provisions:

This constitutes nonexclusive, one-time permission to use the items of the MPQ Tellegen Absorption Scale (TAS) as a research instrument in the study described in this request (item numbers will be used in dissertation as opposed to item content). Any additional use or reproduction of MPQ items (such as journal articles, alternative media, any online use or distribution/display) would require separate permission. All MPQ materials are currently unpublished and copyrighted by Auke Tellegen.

Please be sure the following copyright line is used: Multidimensional Personality Questionnaire™ (MPQ™). Copyright © 1995, 2003 by Auke Tellegen. Unpublished test. Used by permission of the University of Minnesota Press. All Rights Reserved.

Tami Brown — January 22, 2014

Tami Brown
Contracts/Permissions Coordinator

Date

Mill Place, Suite 290
111 Third Avenue South
Minneapolis, MN 55401-2520
Office: 612-627-1970
http://www.upress.umn.edu

Permission for Usage of the FSS 2, S DFSS, and C FSS

For use by Katelyn Procci only. Received from Mind Garden, Inc. on March 28, 2014

mind garden

www.mindgarden.com

To whom it may concern,

This letter is to grant permission for the above named person to use the following copyright material in his/her research:

Flow Scales:

LONG Dispositional Flow Scale-2 (DFS-2) –Physical, Copyright © 1996, 2001 by S.A. Jackson. All rights reserved.

LONG Dispositional Flow Scale-2 (DFS-2) - General, Copyright © 2009 by S.A. Jackson. All rights reserved.

LONG Flow State Scale-2 (FSS-2) - Physical, Copyright © 1996, 2001 by S.A. Jackson. All rights reserved.

LONG Flow State Scale (FSS-2) - General, Copyright © 2009 by S.A. Jackson. All rights reserved.

SHORT Dispositional Flow Scale (S DFS), Copyright © 2002, 2009 by S.A. Jackson. All rights reserved.

SHORT Flow State Scale (S FSS), Copyright © 2002, 2009 by S.A. Jackson. All rights reserved.

CORE Dispositional Flow Scale (C DFS), Copyright © 2006, 2009 by S.A. Jackson and A. J. Martin. All rights reserved.

CORE Flow State Scale (C FSS), Copyright © 2006, 2009 by S.A. Jackson and A. J. Martin. All rights reserved.

Five sample items in total may be reproduced for inclusion in a proposal, thesis, or dissertation.

Any or all of these entire instruments may not be included or reproduced at any time in any published material.

Sincerely,

Robert Most
Mind Garden, Inc.
www.mindgarden.com

APPENDIX D: IRB APPROVAL LETTER

University of Central Florida Institutional Review Board
Office of Research & Commercialization
12201 Research Parkway, Suite 501
Orlando, Florida 32826-3246
Telephone: 407-823-2901 or 407-882-2276
www.research.ucf.edu/compliance/irb.html

Approval of Human Research

From: **UCF Institutional Review Board #1**
FWA00000351, IRB00001138

To: **Katelyn C. Procci** and Co-PIs: **Clint A. Bowers, Florian G. Jentsch, Thomas R. McDaniel, Valerie K. Sims**

Date: **March 27, 2014**

Dear Researcher:

On 3/26/2014 the IRB approved the following human participant research until 3/26/2015 inclusive:

Type of Review:	Submission Correction for UCF Initial Review Submission Form Expedited Review
Project Title:	Exploring Minecraft Gameplay
Investigator:	Katelyn C Procci
IRB Number:	SBE-14-10135
Funding Agency:	
Grant Title:	
Research ID:	N/A

The scientific merit of the research was considered during the IRB review. The Continuing Review Application must be submitted 30days prior to the expiration date for studies that were previously expedited, and 60 days prior to the expiration date for research that was previously reviewed at a convened meeting. Do not make changes to the study (i.e., protocol, methodology, consent form, personnel, site, etc.) before obtaining IRB approval. A Modification Form **cannot** be used to extend the approval period of a study. All forms may be completed and submitted online at https://iris.research.ucf.edu .

If continuing review approval is not granted before the expiration date of 3/26/2015, approval of this research expires on that date. When you have completed your research, please submit a Study Closure request in iRIS so that IRB records will be accurate.

Use of the approved, stamped consent document(s) is required. The new form supersedes all previous versions, which are now invalid for further use. Only approved investigators (or other approved key study personnel) may solicit consent for research participation.

In the conduct of this research, you are responsible to follow the requirements of the Investigator Manual.

On behalf of Sophia Dziegielewski, Ph.D., L.C.S.W., UCF IRB Chair, this letter is signed by:

Signature applied by Patria Davis on 03/27/2014 12:31:58 PM EST

IRB Coordinator

Page 1 of 1

APPENDIX E: INFORMED CONSENT DOCUMENT

Exploring *Minecraft* Gameplay

Informed Consent Document

Principal Investigator(s):	Katelyn Procci, MS
Faculty Supervisor:	Clint Bowers, PhD
Co-Investigator(s):	Florian Jentsch, PhD; Valerie Sims, PhD; Rudy McDaniel, PhD;
Research Associate(s):	David Garcia; Christopher Bratta; Cierra Godwin; Stephanie Formanek; and, Christine Kreutzer.
Investigational Site(s):	University of Central Florida, Department of Psychology.

Introduction: Researchers at the University of Central Florida (UCF) study many topics. To do this we need the help of people who agree to take part in a research study. You are being invited to take part in a research study which will include about 2,000 people from UCF and the surrounding Orlando, FL area. You have been asked to take part in this research study because you are either a student in a psychology class, or have heard about the study and are interested in participating. You must be 18 years of age or older to be included in the research study.

The person doing this research is Katelyn Procci of the University of Central Florida Psychology Department. Because the researcher is a graduate student, she is being guided by Dr. Clint Bowers, a UCF faculty supervisor in the Psychology Department. UCF students learning about research are helping to do this study as part of the research team. Their names are: David Garcia, Christopher Bratta, Cierra Godwin, Stephanie Formanek, and Christine Kreutzer.

What you should know about a research study:

- Someone will explain this research study to you.
- A research study is something you volunteer for.
- Whether or not you take part is up to you.
- You should take part in this study only because you want to.
- You can choose not to take part in the research study.
- You can agree to take part now and later change your mind.
- Whatever you decide it will not be held against you.
- Feel free to ask all the questions you want before you decide.

University of Central Florida IRB
IRB NUMBER: SBE-14-10135
IRB APPROVAL DATE: 12/24/2014
IRB EXPIRATION DATE: 3/26/2015

Purpose of the research study: The purpose of this study is to gather data about how people experience the game *Minecraft*.

What you will be asked to do in the study: This study will be conducted in two stages. All participants will complete an online survey, which will take one hour. During the online survey:

- You will receive this Informed Consent Document and be asked to confirm your acceptance to be in this study.
- You will complete a brief demographics survey.
- You will complete four additional surveys about your general tendency to experience immersion and absorption, your willingness to suspend disbelief, and your level of creative imagination.
- You will be asked about your gameplay experiences related to the game *Minecraft.*

Some participants (roughly 156) will be selected to participate in an in-person experimental study. They will be selected on the basis of their responses from the online survey. These individuals will be contacted via email (either through Sona or through an email address that they provide) to participate. This session will take roughly one hour to complete. If you are invited to participate in this in-person experimental session, the following will occur:

- Upon arrival to the session, you will once again receive this Informed Consent Document, which you will read. The purpose of the experiment will also be explained to you. You will be able to ask questions or voice any concerns you may have.
- Immediately after the consent process, you may complete a brief spatial ability task.
- You will play the game *Minecraft.*
- After gameplay, you will complete eight surveys about your experience.
- You will receive additional information about the purpose of this study and have a chance to ask any questions you may have.

You do not have to answer every question or complete every task.

You will not lose any benefits if you skip questions or tasks.

Location: Psychology Building, University of Central Florida

Time required: The online survey will take 1 hour to complete. Additionally, the invitation-only in-person experimental session will also take 1 hour to complete.

Risks: There is a small risk that people who take part will develop what is ordinarily referred to as simulator sickness. It occurs once in awhile to people who are exposed to prolonged continuous testing in simulated environments, such as a game. Symptoms consist of nausea and a feeling of being light-headed. The risk is minimized as a result of the short duration of each session. If you experience any of the symptoms mentioned, please tell the researcher and remain seated until the symptoms disappear.

University of Central Florida IRB
IRB NUMBER: SBE-14-10135
IRB APPROVAL DATE: 12/24/2014
IRB EXPIRATION DATE: 3/26/2015

Benefits: We cannot promise any benefits to you or others from your taking part in this research. However, possible benefits include gaining insight into how psychological research studies are conducted.

Compensation or payment: If you have signed up through the Sona system, you will be awarded 0.5 Sona credits for completing the online survey. If you are selected for participation in the in-person study and complete it, you will be awarded an additional 1.0 Sona credits for your participation. If you choose not to participate, you may notify your instructor and ask for an alternative assignment of equal effort for equal credit. There will be no penalty. If you did not participate in this study for Sona credit, you will not be compensated for completing the online study, however if you are invited to *and* complete the in-person experimental study, you will be compensated with a $20 Amazon.com gift card. You may only receive one form of compensation (i.e., you cannot complete this study for both Sona credit and the gift card).

Confidentiality: We cannot promise complete secrecy. As such, we will limit your personal data collected in this study to people who have a need to review this information. Also, all paper-records of data will be destroyed 5 years after it has been collected and entered into a secure electronic database. Organizations that may inspect and copy your information include the IRB and other representatives of UCF.

Study contact for questions about the study or to report a problem: If you have questions, concerns, or complaints, would like to know the results, or think the research has hurt you, contact Katelyn Procci, Applied Experimental Human Factors Program, Department of Psychology, College of Sciences, by email at kprocci@knights.ucf.edu, or Dr. Clint Bowers, Faculty Supervisor, Department Psychology by email at clint.bowers@ucf.edu.

IRB contact about your rights in the study or to report a complaint: Research at the University of Central Florida involving human participants is carried out under the oversight of the Institutional Review Board (UCF IRB). This research has been reviewed and approved by the IRB. For information about the rights of people who take part in research, please contact: Institutional Review Board, University of Central Florida, Office of Research & Commercialization, 12201 Research Parkway, Suite 501, Orlando, FL 32826-3246 or by telephone at (407) 823-2901. You may also talk to them for any of the following:

- Your questions, concerns, or complaints are not being answered by the research team.
- You cannot reach the research team.
- You want to talk to someone besides the research team.
- You want to get information or provide input about this research.

University of Central Florida IRB
IRB NUMBER: SBE-14-10135
IRB APPROVAL DATE: 12/24/2014
IRB EXPIRATION DATE: 3/26/2015

APPENDIX F: EXPERIMENTAL VARIABLE LISTS

Variable Codebook

Variable in SPSS	Description	Source
ID	Study-assigned participant ID	Participant ID Log
Condition	Condition	Participant ID Log
Researcher	Research Assistant	Participant ID Log
Date	Date	Participant ID Log
Time	Time	Participant ID Log
Age	Age in years	Demographics Questionnaire
Gender	Gender	Demographics Questionnaire
Race	Race	Demographics Questionnaire
Year	Year in school	Demographics Questionnaire
PrimaryLangauge	Primary Language	Demographics Questionnaire
ComputerComfort	Average comfort with computers	Demographics Questionnaire
GameComfort	Average comfort with games	Demographics Questionnaire
HoursGaming	Hours / week playing games	Demographics Questionnaire
PlayComputer	Do they play games on a computer?	Demographics Questionnaire
PlayConsole	Do they play games on a console?	Demographics Questionnaire
PlayCellphone	Do they play games on a cellphone?	Demographics Questionnaire
PlayTablet	Do they play games on a tablet?	Demographics Questionnaire
PlayHandheld	Do they play games on a handheld?	Demographics Questionnaire
PlayNothing	Do they not play games?	Demographics Questionnaire
PlayMajority	How do the play the majoirty of games?	Demographics Questionnaire
FaveGame1	What is their favorite game?	Demographics Questionnaire
FaveGame2	What is their second favorite game?	Demographics Questionnaire
FaveGame3	What is their third favorite game?	Demographics Questionnaire
HMD	Ever played a game with an HMD?	Demographics Questionnaire
VGSE_01	Video Game Self-Efficacy - Item 01	VGSE - Pavlas, 2010
VGSE_02	Video Game Self-Efficacy - Item 02	VGSE - Pavlas, 2010
VGSE_03	Video Game Self-Efficacy - Item 03	VGSE - Pavlas, 2010
VGSE_04	Video Game Self-Efficacy - Item 04	VGSE - Pavlas, 2010
VGSE_05	Video Game Self-Efficacy - Item 05	VGSE - Pavlas, 2010
VGSE_06	Video Game Self-Efficacy - Item 06	VGSE - Pavlas, 2010
VGSE_07	Video Game Self-Efficacy - Item 07	VGSE - Pavlas, 2010
VGSE_08	Video Game Self-Efficacy - Item 08	VGSE - Pavlas, 2010
VGSE_09	Video Game Self-Efficacy - Item 09	VGSE - Pavlas, 2010
VGSE_10	Video Game Self-Efficacy - Item 10	VGSE - Pavlas, 2010
ITQ_01	Immersive / Involvement General Tendencies - Item 01	ITQ - Witmer & Singer, 1998

ITQ_02	Immersive / Involvement General Tendencies - Item 02	ITQ - Witmer & Singer, 1998
ITQ_03	Immersive / Involvement General Tendencies - Item 03	ITQ - Witmer & Singer, 1998
ITQ_04	Immersive / Involvement General Tendencies - Item 04	ITQ - Witmer & Singer, 1998
ITQ_05	Immersive / Involvement General Tendencies - Item 05	ITQ - Witmer & Singer, 1998
ITQ_06	Immersive / Involvement General Tendencies - Item 06	ITQ - Witmer & Singer, 1998
ITQ_07	Immersive / Involvement General Tendencies - Item 07	ITQ - Witmer & Singer, 1998
ITQ_08	Immersive / Involvement General Tendencies - Item 08	ITQ - Witmer & Singer, 1998
ITQ_09	Immersive / Involvement General Tendencies - Item 09	ITQ - Witmer & Singer, 1998
ITQ_10	Immersive / Involvement General Tendencies - Item 10	ITQ - Witmer & Singer, 1998
ITQ_11	Immersive / Involvement General Tendencies - Item 11	ITQ - Witmer & Singer, 1998
ITQ_12	Immersive / Involvement General Tendencies - Item 12	ITQ - Witmer & Singer, 1998
ITQ_13	Immersive / Involvement General Tendencies - Item 13	ITQ - Witmer & Singer, 1998
ITQ_14	Immersive / Involvement General Tendencies - Item 14	ITQ - Witmer & Singer, 1998
ITQ_15	Immersive / Involvement General Tendencies - Item 15	ITQ - Witmer & Singer, 1998
ITQ_16	Immersive / Involvement General Tendencies - Item 16	ITQ - Witmer & Singer, 1998
ITQ_17	Immersive / Involvement General Tendencies - Item 17	ITQ - Witmer & Singer, 1998
ITQ_18	Immersive / Involvement General Tendencies - Item 18	ITQ - Witmer & Singer, 1998
WillingGeneral_01	Willingness to Suspend Disbelief - Item 01	Willingness to Suspend Disbelief - Sas & O'Hare, 2004
WillingGeneral_02	Willingness to Suspend Disbelief - Item 02	Willingness to Suspend Disbelief - new item
WillingGeneral_03	Willingness to Suspend Disbelief - Item 03	Willingness to Suspend Disbelief - new item
WillingGeneral_04	Willingness to Suspend Disbelief - Item 04	Willingness to Suspend Disbelief - new item
TAS_01	Trait Absorption - Item 01	TAS - Tellegen & Atkinson, 1974
TAS_02	Trait Absorption - Item 02	TAS - Tellegen & Atkinson, 1974
TAS_03	Trait Absorption - Item 03	TAS - Tellegen & Atkinson, 1974
TAS_04	Trait Absorption - Item 04	TAS - Tellegen & Atkinson, 1974
TAS_05	Trait Absorption - Item 05	TAS - Tellegen & Atkinson, 1974
TAS_06	Trait Absorption - Item 06	TAS - Tellegen & Atkinson, 1974
TAS_07	Trait Absorption - Item 07	TAS - Tellegen & Atkinson, 1974

TAS_08	Trait Absorption - Item 08	TAS - Tellegen & Atkinson, 1974
TAS_09	Trait Absorption - Item 09	TAS - Tellegen & Atkinson, 1974
TAS_10	Trait Absorption - Item 10	TAS - Tellegen & Atkinson, 1974
TAS_11	Trait Absorption - Item 11	TAS - Tellegen & Atkinson, 1974
TAS_12	Trait Absorption - Item 12	TAS - Tellegen & Atkinson, 1974
TAS_13	Trait Absorption - Item 13	TAS - Tellegen & Atkinson, 1974
TAS_14	Trait Absorption - Item 14	TAS - Tellegen & Atkinson, 1974
TAS_15	Trait Absorption - Item 15	TAS - Tellegen & Atkinson, 1974
TAS_16	Trait Absorption - Item 16	TAS - Tellegen & Atkinson, 1974
TAS_17	Trait Absorption - Item 17	TAS - Tellegen & Atkinson, 1974
TAS_18	Trait Absorption - Item 18	TAS - Tellegen & Atkinson, 1974
TAS_19	Trait Absorption - Item 19	TAS - Tellegen & Atkinson, 1974
TAS_20	Trait Absorption - Item 20	TAS - Tellegen & Atkinson, 1974
TAS_21	Trait Absorption - Item 21	TAS - Tellegen & Atkinson, 1974
TAS_22	Trait Absorption - Item 22	TAS - Tellegen & Atkinson, 1974
TAS_23	Trait Absorption - Item 23	TAS - Tellegen & Atkinson, 1974
TAS_24	Trait Absorption - Item 24	TAS - Tellegen & Atkinson, 1974
TAS_25	Trait Absorption - Item 25	TAS - Tellegen & Atkinson, 1974
TAS_26	Trait Absorption - Item 26	TAS - Tellegen & Atkinson, 1974
TAS_27	Trait Absorption - Item 27	TAS - Tellegen & Atkinson, 1974
TAS_28	Trait Absorption - Item 28	TAS - Tellegen & Atkinson, 1974
TAS_29	Trait Absorption - Item 29	TAS - Tellegen & Atkinson, 1974
TAS_30	Trait Absorption - Item 30	TAS - Tellegen & Atkinson, 1974
TAS_31	Trait Absorption - Item 31	TAS - Tellegen & Atkinson, 1974
TAS_32	Trait Absorption - Item 32	TAS - Tellegen & Atkinson, 1974
TAS_33	Trait Absorption - Item 33	TAS - Tellegen & Atkinson, 1974
TAS_34	Trait Absorption - Item 34	TAS - Tellegen & Atkinson, 1974
CIS_Check1	Did they complete the entire exercise?	CIS - Wilson & Barber, 1978
CIS_01	Creative Imagination Scale - Item 01	CIS - Wilson & Barber, 1978
CIS_02	Creative Imagination Scale - Item 02	CIS - Wilson & Barber, 1978
CIS_03	Creative Imagination Scale - Item 03	CIS - Wilson & Barber, 1978
CIS_04	Creative Imagination Scale - Item 04	CIS - Wilson & Barber, 1978
CIS_05	Creative Imagination Scale - Item 05	CIS - Wilson & Barber, 1978
CIS_06	Creative Imagination Scale - Item 06	CIS - Wilson & Barber, 1978
CIS_07	Creative Imagination Scale - Item 07	CIS - Wilson & Barber, 1978
CIS_08	Creative Imagination Scale - Item 08	CIS - Wilson & Barber, 1978
CIS_09	Creative Imagination Scale - Item 09	CIS - Wilson & Barber, 1978
CIS_10	Creative Imagination Scale - Item 10	CIS - Wilson & Barber, 1978
CIS_Check2	Did they get the passcode right?	CIS - Wilson & Barber, 1978

Enjoyment_Games	General game enjoyment	Minecraft Screener
Enjoyment_MC	Minecraft game enjoyment	Minecraft Screener
MinecraftComfort	Minecraft game comfort	Minecraft Screener
PlayMinecraft	Have they ever played Minecraft?	Minecraft Screener
HoursMinecraft	Hours they have played Minecraft	Minecraft Gameplay Survey
MinecraftHMD	Have they ever played Minecraft with an HMD?	Minecraft Gameplay Survey
PlayMCComputer	How do they play Minecraft?: Computer	Minecraft Gameplay Survey
PlayMCConsole	How do they play Minecraft?: Console	Minecraft Gameplay Survey
PlayMCCellphone	How do they play Minecraft?: Cellphone	Minecraft Gameplay Survey
PlayMCTablet	How do they play Minecraft?: Tablet	Minecraft Gameplay Survey
PlayMCMajority	How do they play Minecraft?: Majority	Minecraft Gameplay Survey
MCVer_176	Have the played Minecraft ver. 1.7.6 and later	Minecraft Gameplay Survey
MCVer_175	Have the played Minecraft ver. 1.7.5	Minecraft Gameplay Survey
MCVer_174	Have the played Minecraft ver. 1.7.4	Minecraft Gameplay Survey
MCVer_172	Have the played Minecraft ver. 1.7.2	Minecraft Gameplay Survey
MCVer_164	Have the played Minecraft ver. 1.6.4	Minecraft Gameplay Survey
MCVer_162	Have the played Minecraft ver. 1.6.2	Minecraft Gameplay Survey
MCVer_161	Have the played Minecraft ver. 1.6.1	Minecraft Gameplay Survey
MCVer_152	Have the played Minecraft ver. 1.5.2	Minecraft Gameplay Survey
MCVer_151	Have the played Minecraft ver. 1.5.1	Minecraft Gameplay Survey
MCVer_15	Have the played Minecraft ver. 1.5 and below	Minecraft Gameplay Survey
MCVer_147	Have the played Minecraft ver. 4.7	Minecraft Gameplay Survey
MCVer_DontKnow	Have the played Minecraft ver. Don't know	Minecraft Gameplay Survey
MCVer_Prefer	What version of Minecraft do they prefer?	Minecraft Gameplay Survey
MCScreenSize	What size screen do they usually play Minecraft on?	Minecraft Gameplay Survey
MCResolution	What resolution do they usually play Minecraft on?	Minecraft Gameplay Survey
MCRating_Survival	Enjoyment of Surival Mode (rating)	Minecraft Gameplay Survey
MCRating_Creative	Enjoyment of Creative Mode (rating)	Minecraft Gameplay Survey
MCRating_Hardcore	Enjoyment of Hardcore Mode (rating)	Minecraft Gameplay Survey
MCSurvival_Difficulty	Preferred difficulty level in Survival	Minecraft Gameplay Survey
MCWindowed	Play Minecraft in windowed mode	Minecraft Gameplay Survey
MCFullscreen	Play Minecraft in full-screen mode	Minecraft Gameplay Survey
MCLikeSurvival	Like to play in Survival Mode	Minecraft Gameplay Survey
MCLikeCreative	Like to play in Creative Mode	Minecraft Gameplay Survey
MCLikeHardcore	Like to play in Hardcore Mode	Minecraft Gameplay Survey
MCSoundOff	Play with the sound off	Minecraft Gameplay Survey
MCHeadphones	Play with headphones	Minecraft Gameplay Survey

MCPlayAlone	Prefer to play alone	Minecraft Gameplay Survey
MCTexturePacks	Play with texture packs	Minecraft Gameplay Survey
MCTexturePacks_List	List of texture packs (if applicable)	Minecraft Gameplay Survey
MCMods	Play with mods	Minecraft Gameplay Survey
MCMods_List	List of mods (if applicable)	Minecraft Gameplay Survey
MCShaders	Play with shaders	Minecraft Gameplay Survey
MCShaders_List	List of shaders (if applicable)	Minecraft Gameplay Survey
SDFS_01	Short Dispositional Flow Scale - Item 01	S-DFS-2 - Jackson et al., 2008
SDFS_02	Short Dispositional Flow Scale - Item 02	S-DFS-2 - Jackson et al., 2008
SDFS_03	Short Dispositional Flow Scale - Item 03	S-DFS-2 - Jackson et al., 2008
SDFS_04	Short Dispositional Flow Scale - Item 04	S-DFS-2 - Jackson et al., 2008
SDFS_05	Short Dispositional Flow Scale - Item 05	S-DFS-2 - Jackson et al., 2008
SDFS_06	Short Dispositional Flow Scale - Item 06	S-DFS-2 - Jackson et al., 2008
SDFS_07	Short Dispositional Flow Scale - Item 07	S-DFS-2 - Jackson et al., 2008
SDFS_08	Short Dispositional Flow Scale - Item 08	S-DFS-2 - Jackson et al., 2008
SDFS_09	Short Dispositional Flow Scale - Item 09	S-DFS-2 - Jackson et al., 2008
Days_Sim	SSQ - Pre-Game Survey	SSQ - Kennedy et al., 1993
Days_Flight	SSQ - Pre-Game Survey	SSQ - Kennedy et al., 1993
Days_Sea	SSQ - Pre-Game Survey	SSQ - Kennedy et al., 1993
Days_VE	SSQ - Pre-Game Survey	SSQ - Kennedy et al., 1993
Unusual	SSQ - Pre-Game Survey	SSQ - Kennedy et al., 1993
Fitness	SSQ - Pre-Game Survey	SSQ - Kennedy et al., 1993
FitnessReason	SSQ - Pre-Game Survey	SSQ - Kennedy et al., 1993
Sick	SSQ - Pre-Game Survey	SSQ - Kennedy et al., 1993
SickReason	SSQ - Pre-Game Survey	SSQ - Kennedy et al., 1993
SickRating	SSQ - Pre-Game Survey	SSQ - Kennedy et al., 1993
SickDays	SSQ - Pre-Game Survey	SSQ - Kennedy et al., 1993
SickSymptoms	SSQ - Pre-Game Survey	SSQ - Kennedy et al., 1993
Recovered	SSQ - Pre-Game Survey	SSQ - Kennedy et al., 1993
Beer	SSQ - Pre-Game Survey	SSQ - Kennedy et al., 1993
Wine	SSQ - Pre-Game Survey	SSQ - Kennedy et al., 1993
Liquor	SSQ - Pre-Game Survey	SSQ - Kennedy et al., 1993
None	SSQ - Pre-Game Survey	SSQ - Kennedy et al., 1993
Sedatives	SSQ - Pre-Game Survey	SSQ - Kennedy et al., 1993
Headache	SSQ - Pre-Game Survey	SSQ - Kennedy et al., 1993
Antihistamine	SSQ - Pre-Game Survey	SSQ - Kennedy et al., 1993
Decongestant	SSQ - Pre-Game Survey	SSQ - Kennedy et al., 1993
OtherDrug	SSQ - Pre-Game Survey	SSQ - Kennedy et al., 1993
SleepHours	SSQ - Pre-Game Survey	SSQ - Kennedy et al., 1993

SleepRating	SSQ - Pre-Game Survey	SSQ - Kennedy et al., 1993
SSQPre_01	SSQ Pre-Game Symptom Rating - General Discomfort	SSQ - Kennedy et al., 1993
SSQPre_02	SSQ Pre-Game Symptom Rating - Fatigue	SSQ - Kennedy et al., 1993
SSQPre_03	SSQ Pre-Game Symptom Rating - Headache	SSQ - Kennedy et al., 1993
SSQPre_04	SSQ Pre-Game Symptom Rating - Eye strain	SSQ - Kennedy et al., 1993
SSQPre_05	SSQ Pre-Game Symptom Rating - Difficulty focusing	SSQ - Kennedy et al., 1993
SSQPre_06	SSQ Pre-Game Symptom Rating - Salivation increased	SSQ - Kennedy et al., 1993
SSQPre_07	SSQ Pre-Game Symptom Rating - Sweating	SSQ - Kennedy et al., 1993
SSQPre_08	SSQ Pre-Game Symptom Rating - Nausea	SSQ - Kennedy et al., 1993
SSQPre_09	SSQ Pre-Game Symptom Rating - Difficulty concentrating	SSQ - Kennedy et al., 1993
SSQPre_10	SSQ Pre-Game Symptom Rating - "Fullness of the head"	SSQ - Kennedy et al., 1993
SSQPre_11	SSQ Pre-Game Symptom Rating - Blurred vision	SSQ - Kennedy et al., 1993
SSQPre_12	SSQ Pre-Game Symptom Rating - Dizziness with eyes open	SSQ - Kennedy et al., 1993
SSQPre_13	SSQ Pre-Game Symptom Rating - Dizziness with eyes closed	SSQ - Kennedy et al., 1993
SSQPre_14	SSQ Pre-Game Symptom Rating - Vertigo	SSQ - Kennedy et al., 1993
SSQPre_15	SSQ Pre-Game Symptom Rating - Stomach awareness	SSQ - Kennedy et al., 1993
SSQPre_16	SSQ Pre-Game Symptom Rating - Burping	SSQ - Kennedy et al., 1993
CFSS_01	Core Flow State Scale - Item 01	C FSS - Martin & Jackson, 2008
CFSS_02	Core Flow State Scale - Item 02	C FSS - Martin & Jackson, 2008
CFSS_03	Core Flow State Scale - Item 03	C FSS - Martin & Jackson, 2008
CFSS_04	Core Flow State Scale - Item 04	C FSS - Martin & Jackson, 2008
CFSS_05	Core Flow State Scale - Item 05	C FSS - Martin & Jackson, 2008
CFSS_06	Core Flow State Scale - Item 06	C FSS - Martin & Jackson, 2008
CFSS_07	Core Flow State Scale - Item 07	C FSS - Martin & Jackson, 2008
CFSS_08	Core Flow State Scale - Item 08	C FSS - Martin & Jackson, 2008
CFSS_09	Core Flow State Scale - Item 09	C FSS - Martin & Jackson, 2008
CFSS_10	Core Flow State Scale - Item 10	C FSS - Martin & Jackson, 2008
FSS_01	Flow State Scale 2 - Item 01	FSS-2 - Jackson & Eklund, 2002
FSS_02	Flow State Scale 2 - Item 02	FSS-2 - Jackson & Eklund, 2002
FSS_03	Flow State Scale 2 - Item 03	FSS-2 - Jackson & Eklund, 2002
FSS_04	Flow State Scale 2 - Item 04	FSS-2 - Jackson & Eklund, 2002
FSS_05	Flow State Scale 2 - Item 05	FSS-2 - Jackson & Eklund, 2002
FSS_06	Flow State Scale 2 - Item 06	FSS-2 - Jackson & Eklund, 2002
FSS_07	Flow State Scale 2 - Item 07	FSS-2 - Jackson & Eklund, 2002

FSS_08	Flow State Scale 2 - Item 08	FSS-2 - Jackson & Eklund, 2002
FSS_09	Flow State Scale 2 - Item 09	FSS-2 - Jackson & Eklund, 2002
FSS_10	Flow State Scale 2 - Item 10	FSS-2 - Jackson & Eklund, 2002
FSS_11	Flow State Scale 2 - Item 11	FSS-2 - Jackson & Eklund, 2002
FSS_12	Flow State Scale 2 - Item 12	FSS-2 - Jackson & Eklund, 2002
FSS_13	Flow State Scale 2 - Item 13	FSS-2 - Jackson & Eklund, 2002
FSS_14	Flow State Scale 2 - Item 14	FSS-2 - Jackson & Eklund, 2002
FSS_15	Flow State Scale 2 - Item 15	FSS-2 - Jackson & Eklund, 2002
FSS_16	Flow State Scale 2 - Item 16	FSS-2 - Jackson & Eklund, 2002
FSS_17	Flow State Scale 2 - Item 17	FSS-2 - Jackson & Eklund, 2002
FSS_18	Flow State Scale 2 - Item 18	FSS-2 - Jackson & Eklund, 2002
FSS_19	Flow State Scale 2 - Item 19	FSS-2 - Jackson & Eklund, 2002
FSS_20	Flow State Scale 2 - Item 20	FSS-2 - Jackson & Eklund, 2002
FSS_21	Flow State Scale 2 - Item 21	FSS-2 - Jackson & Eklund, 2002
FSS_22	Flow State Scale 2 - Item 22	FSS-2 - Jackson & Eklund, 2002
FSS_23	Flow State Scale 2 - Item 23	FSS-2 - Jackson & Eklund, 2002
FSS_24	Flow State Scale 2 - Item 24	FSS-2 - Jackson & Eklund, 2002
FSS_25	Flow State Scale 2 - Item 25	FSS-2 - Jackson & Eklund, 2002
FSS_26	Flow State Scale 2 - Item 26	FSS-2 - Jackson & Eklund, 2002
FSS_27	Flow State Scale 2 - Item 27	FSS-2 - Jackson & Eklund, 2002
FSS_28	Flow State Scale 2 - Item 28	FSS-2 - Jackson & Eklund, 2002
FSS_29	Flow State Scale 2 - Item 29	FSS-2 - Jackson & Eklund, 2002
FSS_30	Flow State Scale 2 - Item 30	FSS-2 - Jackson & Eklund, 2002
FSS_31	Flow State Scale 2 - Item 31	FSS-2 - Jackson & Eklund, 2002
FSS_32	Flow State Scale 2 - Item 32	FSS-2 - Jackson & Eklund, 2002
FSS_33	Flow State Scale 2 - Item 33	FSS-2 - Jackson & Eklund, 2002
FSS_34	Flow State Scale 2 - Item 34	FSS-2 - Jackson & Eklund, 2002
FSS_35	Flow State Scale 2 - Item 35	FSS-2 - Jackson & Eklund, 2002
FSS_36	Flow State Scale 2 - Item 36	FSS-2 - Jackson & Eklund, 2002
Pres_01	Presence Count Survey - Item 01	Presence Survey - Slater & Steed, 2000
Pres_02	Presence Count Survey - Item 02	Presence Survey - Slater & Steed, 2000
Pres_03	Presence Count Survey - Item 03	Presence Survey - Slater & Steed, 2000
Pres_04	Presence Count Survey - Item 04	Presence Survey - Slater & Steed, 2000
Pres_05	Presence Count Survey - Item 05	Presence Survey - Slater & Steed, 2000
Pres_06	Presence Count Survey - Item 06	PQ Rev 3 - Witmer et al., 2005
Pres_07	Presence Count Survey - Item 07	PQ Rev 3 - Witmer et al., 2005

Pres_08	Presence Count Survey - Item 08	PQ Rev 3 - Witmer et al., 2005
Pres_09	Presence Count Survey - Item 09	PQ Rev 3 - Witmer et al., 2005
Pres_Percent	Estimated time Present during Gameplay	Presence Survey - Slater & Steed, 2000
Imm_01	Immersion - Item 01	PQ Rev 3 - Witmer et al., 2005
Imm_02	Immersion - Item 02	PQ Rev 3 - Witmer et al., 2005
Imm_03	Immersion - Item 03	PQ Rev 3 - Witmer et al., 2005
Imm_04	Immersion - Item 04	PQ Rev 3 - Witmer et al., 2005
Imm_05	Immersion - Item 05	PQ Rev 3 - Witmer et al., 2005
Imm_06	Immersion - Item 06	PQ Rev 3 - Witmer et al., 2005
Imm_07	Immersion - Item 07	PQ Rev 3 - Witmer et al., 2005
Imm_08	Immersion - Item 08	PQ Rev 3 - Witmer et al., 2005
PreReq_01	Pre-Requisite - Item 01	PQ Rev 3 - Witmer et al., 2005
PreReq_02	Pre-Requisite - Item 02	PQ Rev 3 - Witmer et al., 2005
PreReq_03	Pre-Requisite - Item 03	PQ Rev 3 - Witmer et al., 2005
PreReq_04	Pre-Requisite - Item 04	PQ Rev 3 - Witmer et al., 2005
PreReq_05	Pre-Requisite - Item 05	PQ Rev 3 - Witmer et al., 2005
Imm_09	Immersion - Item 09	PQ Rev 3 - Witmer et al., 2005
Imm_10	Immersion - Item 10	PQ Rev 3 - Witmer et al., 2005
Imm_11	Immersion - Item 11	PQ Rev 3 - Witmer et al., 2005
Imm_12	Immersion - Item 12	PQ Rev 3 - Witmer et al., 2005
Imm_13	Immersion - Item 13	PQ Rev 3 - Witmer et al., 2005
Imm_14	Immersion - Item 14	PQ Rev 3 - Witmer et al., 2005
Imm_15	Immersion - Item 15	PQ Rev 3 - Witmer et al., 2005
Imm_16	Immersion - Item 16	PQ Rev 3 - Witmer et al., 2005
PreReq_06	Pre-Requisite - Item 06	PQ Rev 3 - Witmer et al., 2005
PreReq_07	Pre-Requisite - Item 07	PQ Rev 3 - Witmer et al., 2005
PreReq_08	Pre-Requisite - Item 08	PQ Rev 3 - Witmer et al., 2005
PreReq_09	Pre-Requisite - Item 09	PQ Rev 3 - Witmer et al., 2005
PreReq_10	Pre-Requisite - Item 10	PQ Rev 3 - Witmer et al., 2005
Invol_01	Involvement - Item 01	Revised PII - Zaichkowsky, 1994
Invol_02	Involvement - Item 02	Revised PII - Zaichkowsky, 1994
Invol_03	Involvement - Item 03	Revised PII - Zaichkowsky, 1994
Invol_04	Involvement - Item 04	Revised PII - Zaichkowsky, 1994
Invol_05	Involvement - Item 05	Revised PII - Zaichkowsky, 1994
Invol_06	Involvement - Item 06	Revised PII - Zaichkowsky, 1994
Invol_07	Involvement - Item 07	Revised PII - Zaichkowsky, 1994
Invol_08	Involvement - Item 08	Revised PII - Zaichkowsky, 1994
Invol_09	Involvement - Item 09	Revised PII - Zaichkowsky, 1994

Invol_10	Involvement - Item 10	Revised PII - Zaichkowsky, 1994
Invol_11	Involvement - Item 11	Novel involvement item
Invol_12	Involvement - Item 12	Novel involvement item
Invol_13	Involvement - Item 13	Novel involvement item
Invol_14	Involvement - Item 14	Novel involvement item
Invol_15	Involvement - Item 15	PQ Rev 3 - Witmer et al., 2005
CAS_01	Cognitive Absorption - Item 01	CAS - Agarwal & Karahanna, 2000
CAS_02	Cognitive Absorption - Item 02	CAS - Agarwal & Karahanna, 2000
CAS_03	Cognitive Absorption - Item 03	CAS - Agarwal & Karahanna, 2000
CAS_04	Cognitive Absorption - Item 04	CAS - Agarwal & Karahanna, 2000
CAS_05	Cognitive Absorption - Item 05	CAS - Agarwal & Karahanna, 2000
CAS_06	Cognitive Absorption - Item 06	CAS - Agarwal & Karahanna, 2000
CAS_07	Cognitive Absorption - Item 07	CAS - Agarwal & Karahanna, 2000
CAS_08	Cognitive Absorption - Item 08	CAS - Agarwal & Karahanna, 2000
CAS_09	Cognitive Absorption - Item 09	CAS - Agarwal & Karahanna, 2000
CAS_10	Cognitive Absorption - Item 10	CAS - Agarwal & Karahanna, 2000
CAS_11	Cognitive Absorption - Item 11	CAS - Agarwal & Karahanna, 2000
CAS_12	Cognitive Absorption - Item 12	CAS - Agarwal & Karahanna, 2000
CAS_13	Cognitive Absorption - Item 13	CAS - Agarwal & Karahanna, 2000
CAS_14	Cognitive Absorption - Item 14	CAS - Agarwal & Karahanna, 2000
CAS_15	Cognitive Absorption - Item 15	CAS - Agarwal & Karahanna, 2000
CAS_16	Cognitive Absorption - Item 16	CAS - Agarwal & Karahanna, 2000
CAS_17	Cognitive Absorption - Item 17	CAS - Agarwal & Karahanna, 2000
CAS_18	Cognitive Absorption - Item 18	CAS - Agarwal & Karahanna, 2000
CAS_19	Cognitive Absorption - Item 19	CAS - Agarwal & Karahanna, 2000
CAS_20	Cognitive Absorption - Item 20	CAS - Agarwal & Karahanna, 2000
Willing_01	Willingness to Suspend Disbelief - Item 01	Sas & O'Hare, 2004
Willing_02	Willingness to Suspend Disbelief - Item 02	New item
Willing_03	Willingness to Suspend Disbelief - Item 03	New item
Willing_04	Willingness to Suspend Disbelief - Item 04	New item
SUS_01	Usability - Item 01	SUS - Brooke, 1986, reprinted 1996
SUS_02	Usability - Item 02	SUS - Brooke, 1986, reprinted 1996
SUS_03	Usability - Item 03	SUS - Brooke, 1986, reprinted 1996
SUS_04	Usability - Item 04	SUS - Brooke, 1986, reprinted 1996
SUS_05	Usability - Item 05	SUS - Brooke, 1986, reprinted 1996
SUS_06	Usability - Item 06	SUS - Brooke, 1986, reprinted 1996
SUS_07	Usability - Item 07	SUS - Brooke, 1986, reprinted 1996
SUS_08	Usability - Item 08	SUS - Brooke, 1986, reprinted 1996
SUS_09	Usability - Item 09	SUS - Brooke, 1986, reprinted 1996

SUS_10	Usability - Item 10	SUS - Brooke, 1986, reprinted 1996
SSQ_01	SSQ Post-Game Symptom Rating - General Discomfort	SSQ - Kennedy et al., 1993
SSQ_02	SSQ Post-Game Symptom Rating - Fatigue	SSQ - Kennedy et al., 1993
SSQ_03	SSQ Post-Game Symptom Rating - Headache	SSQ - Kennedy et al., 1993
SSQ_04	SSQ Post-Game Symptom Rating - Eye strain	SSQ - Kennedy et al., 1993
SSQ_05	SSQ Post-Game Symptom Rating - Difficulty focusing	SSQ - Kennedy et al., 1993
SSQ_06	SSQ Post-Game Symptom Rating - Salivation increased	SSQ - Kennedy et al., 1993
SSQ_07	SSQ Post-Game Symptom Rating - Sweating	SSQ - Kennedy et al., 1993
SSQ_08	SSQ Post-Game Symptom Rating - Nausea	SSQ - Kennedy et al., 1993
SSQ_09	SSQ Post-Game Symptom Rating - Difficulty concentrating	SSQ - Kennedy et al., 1993
SSQ_10	SSQ Post-Game Symptom Rating - "Fullness of the head"	SSQ - Kennedy et al., 1993
SSQ_11	SSQ Post-Game Symptom Rating - Blurred vision	SSQ - Kennedy et al., 1993
SSQ_12	SSQ Post-Game Symptom Rating - Dizziness with eyes open	SSQ - Kennedy et al., 1993
SSQ_13	SSQ Post-Game Symptom Rating - Dizziness with eyes closed	SSQ - Kennedy et al., 1993
SSQ_14	SSQ Post-Game Symptom Rating - Vertigo	SSQ - Kennedy et al., 1993
SSQ_15	SSQ Post-Game Symptom Rating - Stomach awareness	SSQ - Kennedy et al., 1993
SSQ_16	SSQ Post-Game Symptom Rating - Burping	SSQ - Kennedy et al., 1993
SSQ_Motion	SSQ Post-Game Survey	SSQ - Kennedy et al., 1993
PerformanceRating	SSQ Post-Game Survey	SSQ - Kennedy et al., 1993
SSQ_Unusual	SSQ Post-Game Survey	SSQ - Kennedy et al., 1993
SSQ_UnusualDesc	SSQ Post-Game Survey	SSQ - Kennedy et al., 1993
Excel Calculated Variables		
CIS_percentile	Creative Imagination percentile based on raw total	CIS - Wilson & Barber, 1978
CIS_t	Creative Imagination *t*-score based on percentile	CIS - Wilson & Barber, 1978
SSQ_Pre_N	SSQ Nausea subscale (pre-game)	SSQ - Kennedy et al., 1993
SSQ_Pre_O	SSQ Occulomotor subscale (pre-game)	SSQ - Kennedy et al., 1993
SSQ_Pre_D	SSQ Disorientation subscale (pre-game)	SSQ - Kennedy et al., 1993
SSQ_Pre_TS	SSQ Total Severity subscale (pre-game)	SSQ - Kennedy et al., 1993
SSQ_Post_N	SSQ Nausea subscale (post-game)	SSQ - Kennedy et al., 1993
SSQ_Post_O	SSQ Occulomotor subscale (post-game)	SSQ - Kennedy et al., 1993
SSQ_Post_D	SSQ Disorientation subscale (post-game)	SSQ - Kennedy et al., 1993
SSQ_Post_TS	SSQ Total Severity subscale (post-game)	SSQ - Kennedy et al., 1993
Involvment_Level	Involvement level	Revised PII - Zaichkowsky, 1994 + new items

Presence_Count	Number of times rated a presence item 6 or 7	Presence Survey - Slater & Steed, 2000
Presence_Class	Classified as present if spent >50% "present" (self-report)	Presence Survey - Slater & Steed, 2000
SUS_Percentile	Usability percentile based on SUS score	SUS - Sauro, 2011
SPSS Calculated Variables		
Attention	Attention - Total	PQ ver 3 - Witmer, et al. (2005) - reconfigured (PreReq)
Involvement	Involvement - Total	Revised PII - Zaichkowsky, 1994 + novel items
Immersion	Immersion - Total	PQ ver 3 - Witmer, et al. (2005) - reconfigured (Imm)
Flow_CORE	Core flow - mean	C FSS - Martin & Jackson, 2008
Flow_CSBalance	Flow subscale mean: Challenge/Skill Balance	FSS-2 - Jackson & Eklund, 2002
Flow_Merging	Flow subscale mean: Merging Action and Awareness	FSS-2 - Jackson & Eklund, 2002
Flow_Goals	Flow subscale mean: Clear Goals	FSS-2 - Jackson & Eklund, 2002
Flow_Feedback	Flow subscale mean: Feedback	FSS-2 - Jackson & Eklund, 2002
Flow_Concentration	Flow subscale mean: Concentration	FSS-2 - Jackson & Eklund, 2002
Flow_Control	Flow subscale mean: Control	FSS-2 - Jackson & Eklund, 2002
Flow_LossofSC	Flow subscale mean: Loss of Self-Consciousness	FSS-2 - Jackson & Eklund, 2002
Flow_Time	Flow subscale mean: Time Distortion	FSS-2 - Jackson & Eklund, 2002
Flow_Autotelic	Flow subscale mean: Autotelic Experience	FSS-2 - Jackson & Eklund, 2002
Usability	SUS Score for overall usability	SUS - Brooke, 1986, reprinted 1996
CognitiveAbsorption	Cognitive Absorption mean overalls core	CAS - Agarwal & Karahanna, 2000
CAS_Time	CAS subscale mean: Time Distortion	CAS - Agarwal & Karahanna, 2000
CAS_Immersion	CAS subscale mean: Immersion	CAS - Agarwal & Karahanna, 2000
CAS_Enjoyment	CAS subscale mean: Heightened Enjoyment	CAS - Agarwal & Karahanna, 2000
CAS_Control	CAS subscale mean: Control	CAS - Agarwal & Karahanna, 2000
CAS_Curiosity	CAS subscale mean: Curiosity	CAS - Agarwal & Karahanna, 2000
VGSE	Video Game Self-Efficacy - Total	VGSE - Pavlas, 2010
ImmInvolTend	Immersion/Involvement Tendencies - Total	ITQ - Witmer & Singer, 1998
Disp_Flow	Dispositional Flow - Mean	S-DFS-2 - Jackson et al., 2008
Willing_General	General Willingness to Suspend Disbelief - Mean	Willingness to Suspend Disbelief - Sas & O'Hare, 2004 + New
CreativeImagination_Total	Creative Imagination Score - Total	CIS - Wilson & Barber, 1978
TraitAbsorption	Trait Absorption - Total	TAS - Tellegen & Atkinson, 1974
Willing	Post-game Willingness to Suspend Disbelief (with all 4 items)	Willingness to Suspend Disbelief - Sas & O'Hare, 2004 + New
Involvement_Recomp	The total involvement score with a bad item (Invol_14) dropped	Revised PII - Zaichkowsky, 1994 + novel items
Flow_FSSMean	Flow subscale mean: Overall for all of FSS-2	FSS-2 - Jackson & Eklund, 2002
LowGE_Total	A total score of the Immersion and	Presence Questionnaire v3 - Witmer

	Involvement scales	et al., 2005
Immersion_Mean	Immersion Mean Score	Presence Questionnaire v3 - Witmer et al., 2005
Involvement_Mean	Involvement Mean Score	Presence Questionnaire v3 - Witmer et al., 2005
LowGE_Mean	A mean score based on the total of the Imm and Involv scales	Revised PII - Zaichkowsky, 1994 + novel items
Willing_3	Post-game Willingness to Suspend Disbelief (with all 3 items)	Willingness to Suspend Disbelief - Sas & O'Hare, 2004 + New
Involvement_RPIITot	The total involvement score from the Revised PII	Revised PII - Zaichkowsky, 1994
Transformed Variables		
SUS_Percentile_RelLog	Reflect & Logarithm transformation of SUS_Percentile	SUS - Sauro, 2011
Flow_Merging_RelLog	Reflect & Logarithm transformation of Flow_Merging	FSS-2 - Jackson & Eklund, 2002
Flow_Goals_RelLog	Reflect & Logarithm transformation of Flow_Goals	FSS-2 - Jackson & Eklund, 2002
Flow_LossofSC_RelLog	Reflect & Logarithm transformation of Flow_LossofSC	FSS-2 - Jackson & Eklund, 2002
Flow_Time_RelLog	Reflect & Logarithm transformation of Flow_Time	FSS-2 - Jackson & Eklund, 2002
CAS_Control_RelLog	Reflect & Logarithm transformation of CAS_Control	CAS - Agarwal & Karahanna, 2000
CAS_Curiosity_RelLog	Reflect & Logarithm transformation of CAS_Curiosity	CAS - Agarwal & Karahanna, 2000
SSQ_Post_TS_Log	Logarithm transformation of SSQ_Post_TS	SSQ - Kennedy et al., 1993

Outliers, Skew, and Kurtosis

The following variables were examined for outliers, skew, and kurtosis:

- CIS_percentile
- CIS_t
- SSQ_Post_TS
- Presence_Count
- SUS_Percentile
- Attention
- Involvement
- Immersion
- Flow_CORE
- Flow_CSBalance
- Flow_Merging
- Flow_Goals
- Flow_Feedback
- Flow_Concentration

- Flow_Control
- Flow_LossofSC
- Flow_Time
- Flow_Autotelic
- Usability
- CognitiveAbsorption
- CAS_Time
- CAS_Immersion
- CAS_Enjoyment
- CAS_Control
- CAS_Curiosity
- VGSE
- ImmInvolTend
- Disp_Flow
- Willing_General
- TraitAbsorption
- Willing
- Involvement_Recomp
- Flow_FSSMean
- LowGE_Total
- Immersion_Mean
- Involvement_Mean
- LowGE_Mean

Hypothesis 8 Stepwise Regression Variables

For Hypothesis 8 – Revisited, the following are the variables tested:

Individual Differences
- Age
- HoursMinecraft
- MCRating_Survival
- CIS_t
- Attention
- Usability
- VGSE
- ImmInvolTend
- Disp_Flow
- TraitAbsorption
- Willing

Conditions

- Imm_Level
- Diff_Level

Other States

- Involvement
- Immersion
- Flow_CORE
- Presence_Count

APPENDIX G: ASSUMPTION TESTING

Assumptions of Correlations

Affected hypotheses: 0.1; 0.3; 0.4a; 0.5; 4; 8, 9a; 9b; 9c; 9d

1. Variables have a linear relationship.
   - ➢ Tested by generating scatterplots and examining them for a roughly linear relationship.

2. Variables must be on an interval or ratio scale.

3. There must be independent of observations.
   - ➢ Tested by creating a regression line between variables and examining the reported Durbin-Watson variable. Ideally, the value should be around 2.0

Hypothesis	Predictor	Outcome	Durbin-Watson	Evaluation
H0.1	Willing_02	Willing_01	2.021	Acceptable
H0.1	Willing_03	Willing_01	2.202	Acceptable
H0.1	Willing_04	Willing_01	2.178	Acceptable
H0.3	Involv_11	Involvement_RPIITot	2.079	Acceptable
H0.3	Involv_12	Involvement_RPIITot	2.144	Acceptable
H0.3	Involv_13	Involvement_RPIITot	2.147	Acceptable
H0.3	Involv_14	Involvement_RPIITot	2.116	Acceptable
H0.3	Involv_15	Involvement_RPIITot	2.117	Acceptable
H0.4a	Presence_Count	Pres_Percent	2.082	Acceptable
H0.5	Flow-FSSMean	Flow_CORE	1.704	Acceptable
H4	Involvement	Immersion	2.126	Acceptable
H8	Flow_CORE	Presence_Count	2.035	Acceptable
H9a	CognitiveAbsorption	Flow_CORE	1.818	Acceptable
H9b	CAS_Time	Flow_Time_RelLog	2.261	Acceptable
H9c	Flow_Control	CAS_Control_RelLog	2.124	Acceptable
H9d	Flow_Autotelic	CAS_Enjoyment	0.888	Not Acceptable

4. Each variable should have a normal distribution.
   - ➢ Assess for the presence of outliers, skew, and kurtosis.

5. There should be homogeneity of variance, meaning that the variability for Variable X is similar across all values of Variable Y.
   - ➢ Tested by generating scatterplots and visually checking that the data are clustered in roughly the same way across an imagined line.

Assumptions of Logistic Regression

Affected hypotheses: 0.4b

1. The dependent variable must be dichotomous.

2. Categories must be mutually-exclusive.

3. Sample size must include 50 cases per predictor.

Assumptions of Exploratory Factor Analysis

Affected hypotheses: 0.2

1. Adequate sample size
   - A ratio of 20 participants per item produces the most accurate results (see Costello & Osborne, 2005).

For Hypothesis 0.2, there are 27 predictors. This means that the sample size should be at least $N = 540$.

2. Factorability
   - Examine the correlations between variables. If there are correlations above r = .30, this means that the correlation matrices are factorable.

3. Linearity of the variables
   - Tested by generating scatterplots and examining them for a roughly linear relationship.

4. Lack of outliers
   - Assess for the presence of outliers, skew, and kurtosis.

5. Absence of multicolinearity
   - Examine the correlations between variables. Variables that are correlated above $r = .70$ indicate multicolinearity, while correlations above $r = .90$ indicate singularity.

Assumptions of ANOVA

Affected hypotheses: 1; 2

1. The dependent variable needs to be continuous.

2. The independent variable(s) need to be independent, categorical groups

3. The variables are normally distributed and free of outliers
   - Assess for the presence of outliers, skew, and kurtosis.

4. The observations should be independent (e.g., a participant isn't in more than one group, there isn't a relationship between groups, etc.)

5. The group sizes should be roughly equal.

6. The variances should be homogenous.
   - Levene's test significance should be greater than .05 (it's a problem if significant)

Hypothesis	Levene's Test *F*	*df*1	*df*2	*p*-level	Evaluation
H1	2.569	3	80	.060	Acceptable
H2	0.899	3	80	.446	Acceptable

Assumptions of Multiple Regression

Affected hypotheses: 3a; 3b; 5a; 5b; 6; 7

1. The sample size should be large enough.
   - Some suggest that you need at least 15 per predictor. Tabachnick and Fidell (2007) suggest the sample size should be larger than (50 + 8M), where M is the number of predictors.

Hypothesis	# of Predictors	Min. Sample Size	Evaluation
3a	3	74	Acceptable
3b	2	66	Acceptable
5a	4	82	Acceptable
5a revisited	4	82	Acceptable
5b	4	82	Acceptable
5b revisited	4	82	Acceptable
6	7	106	Not Acceptable
H6 revisited	4	82	Acceptable
7	14	162	Not Acceptable
H8 revisited	16	178	Not Acceptable

2. There should not be any range restriction with any of the variables.

3. There must be a linear relationship between the independent variables and the dependent variable.
   - Tested by generating scatterplots and examining them for a roughly linear relationship.

4. Each variable should have a normal distribution.
   - Assess for the presence of outliers, skew, and kurtosis.

5. There should be homogeneity of variance, meaning that the variability for Variable X is similar across all values of Variable Y.
    a. Tested by generating scatterplots and visually checking that the data are clustered in roughly the same way across an imagined line

6. There should not be any multicolinearity or singularity between the predictors.
    - Correlations should be calculated between predictors. Correlations above .70 are indicative of multicolinearity and correlations above .90 are indicative of singularity.

Hypothesis	Variable A	Variable B	Correlation	Evaluation
H3b	Attention	Usability	*r*(82) = .403, *p* < .001	Acceptable
H5a / H5b / H6	Involvement	Immersion	*r*(82) = 426, *p* <.001	Acceptable
H5a / H5b / H6 / H7	Involvement	Flow_CSBalance	*r*(82) = .545, *p* <.001	Acceptable
H5a / H5b	Involvement	Attention	*r*(82) = .079, *p* = .238	Acceptable
H5a / H5b / H7	Immersion	Flow_CSBalance	*r*(82) = .523, *p* < .001	Acceptable
H5a / H5b	Immersion	Attention	*r*(82) = .237, *p* = .015	Acceptable
H5a / H5b	Attention	Flow_CSBalance	*r*(82) = - .005, *p* = .483	Acceptable
H5a & H5b revisited	Usability	Involvement	*r*(82) = .310, *p* = .002	Acceptable
H5a & H5b revisited	Usability	Immersion	*r*(82) = .401, *p* < .001	Acceptable
H5a & H5b revisited	Usability	Flow_CSBalance	*r*(82) = .215, *p* = .025	Acceptable
H7	Willing	Involvement	*r*(82) = .205, *p* = .031	Acceptable
H7	Willing	Immersion	*r*(82) = .323, *p* = .001	Acceptable
H7	Willing	Flow_CSBalance	*r*(82) = .400, *p* < .001	Acceptable
H7	CIS_t	Involvement	*r*(82) = .199, *p* = .035	Acceptable
H7	CIS_t	Immersion	*r*(82) = .330, *p* = .001	Acceptable
H7	CIS_t	Flow_CSBalance	*r*(82) = .265, *p* = .007	Acceptable
H7	CIS_t	Willing	*r*(82) = .092, *p* = .204	Acceptable
H7	CIS_t	ImmInvolvTend	*r*(82) = .424, *p* < .001	Acceptable
H7	ImmInvolvTend	Involvement	*r*(82) = .425, *p* < .001	Acceptable
H7	ImmInvolvTend	Immersion	*r*(82) = .253, *p* = .010	Acceptable
H7	ImmInvolvTend	Flow_CSBalance	*r*(82) = .475, *p* < .001	Acceptable
H7	ImmInvolvTend	Willing	*r*(82) = .290, *p* = .004	Acceptable

Note: These tests are all one-tailed.

Please note that for Hypothesis 3a, there were only two predictor variables, one of which was a dummy-coded variable. In this instance, a correlation is not possible. This is why H3a was omitted from the above table.

Since so many additional variables were used in Hypothesis 7 and Hypothesis 8 revisited, the full list of correlations is omitted. None of the correlations exceeded *r* = +/- .60, meaning that this assumption was satisfactorily met

REFERENCES

Admiraal, W., Huizenga, J., Akkerman, S., & ten Dam, G. (2011). The concept of flow in collaborative game-based learning. *Computers in Human Behavior, 27*(3), 1185–1194. doi:10.1016/j.chb.2010.12.013

Adobe Photoshop (CS6) [Computer software]. San Jose, California: Adobe.

Agarwal, R., & Karahanna, E. (2000). Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage. *MIS Quarterly, 24*(4), 665–594. doi:10.2307/3250951

Anderson, C. A., Shibuya, A., Ihori, N., Swing, E. L., Bushman, B. J., Sakamoto, A., ... & Saleem, M. (2010). Violent video game effects on aggression, empathy, and prosocial behavior in eastern and western countries: a meta-analytic review. *Psychological Bulletin, 136*(2), 151-173. doi:10.1037/a0018251.supp

Bagozzi, R. P. (2007). The legacy of the Technology Acceptance Model and a proposal for a paradigm shift. *Journal of the Association for Information Systems*, *8*(4), 244–254.

Balk, S. A., Bertola, M. A., & Inman, V. W. (2013). Simulator sickness questionnaire: Twenty years later. *Proceedings of the Seventh International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design* (pp. 257-263). Iowa City, IA: University of Iowa.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*(2), 122–147. doi:10.1037/0003-066X.37.2.122

Bangor, A., Joseph, K., Sweeney-Dillon, W., Stettler, G., & Pratt, J. (2013). Using the SUS to help demonstrate usability's value to business goals. *Proceedings of the Human Factors* *and Ergonomics Society Annual Meeting, 57*(1), 202-205. doi:10.1177/1541931213571045

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction, 24*(6), 574–594. doi:10.1080/10447310802205776

Barber, T. X., & Wilson, S. C. (1977). Hypnosis, suggestions, and altered states of consciousness: Experimental evaluation of the new cognitive-behavioral theory and the traditional trance-state theory of "hypnosis." *Annals of the New York Academy of Sciences, 296*(1), 34–47. doi:10.1111/j.1749-6632.1977.tb38159.x

Barber, T. X., & Wilson, S. C. (1978). The Barber Suggestibility Scale and the Creative Imagination Scale: Experimental and clinical applications. *The American Journal of Clinical Hypnosis, 21*(2-3), 84–108. doi:10.1080/00029157.1978.10403966

Boot, W. R., Blakely, D. P., & Simons, D. J. (2011). Do action video games improve perception and cognition? *Frontiers in Psychology, 2*, article no. 226. doi:10.3389/fpsyg.2011.00226

Bowman, D. A., & McMahan, R. P. (2007). Virtual reality: How much immersion is enough? *Computer, 40*(7), 36–43. doi:10.1109/MC.2007.257

Boyle, E. A., Connolly, T. M., Hainey, T., & Boyle, J. M. (2012). Engagement in digital entertainment games: A systematic review. *Computers in Human Behavior, 28*(3), 771–780. doi:10.1016/j.chb.2011.11.020

Brockmyer, J. H., Fox, C. M., Curtiss, K. A., McBroom, E., Burkhart, K. M., & Pidruzny, J. N. (2009). The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology, 45*(4), 624-634. doi:10.1016/j.jesp.2009.02.016

Brooke, J. (1996). SUS: A “quick and dirty” usability scale. In P. W. Jordan, B. Thomas, B. W. Weerdmeester, & I. L. McClelland (Eds.), *Usability evaluation in industry* (pp. 189-194). London, England: Taylor & Francis.

Brooke, J. (2013). SUS: A retrospective. *Journal of Usability Studies, 8*(2), 29-40.

Brown, E., & Cairns, P. (2004). A grounded investigation of game immersion. *Proceedings of CHI ‘04 Extended Abstracts on Human Factors in Computing Systems* (pp. 1297–1300). New York, NY: ACM. doi:10.1145/985921.986048

Brown, J. G. (2007). Teaching about genocide in a new millennium. *Social Education, 71*, 21-23.

Canossa, A. (2012). Give me a reason to dig: Qualitative associations between player behavior in Minecraft and life motives. *Proceedings of the International Conference on the Foundations of Digital Games* (pp. 282–283). New York, NY: ACM. doi:10.1145/2282338.2282400

Caveman. (2012, November 24). [SOLVED] - Annoying Crash - Exception in server tick loop [Msg 5]. Message posted to http://terrafirmacraft.com/f/topic/2652-solved-annoying-crash-exception-in-server-tick-loop/

Chen, H., Wigand, R. T., & Nilan, M. S. (1999). Optimal experience of web activities. *Computers in Human Behavior, 15*(5), 585–608. doi:10.1016/S0747-5632(99)00038-2

Chen, J. (2007). Flow in games (and everything else). *Communications of the ACM, 50*(4), 31-34. doi:10.1145/1232743.1232769

“Chroma Hills.” (n.d.). *Chroma Hills*. Retrieved May 3, 2014, from http://www.chromahills.com/forum/downloads.php

Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education, 59*(2), 661–686. doi:10.1016/j.compedu.2012.03.004

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.

Costello, A. B., & Osborne, J. W. (2005). Best practices in Exploratory Factor Analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation, 10*(7), 1-9.

Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., Kenyon, R. V., & Hart, J. C. (1992). The CAVE: Audio visual experience automatic virtual environment. *Communications of the ACM, 35*(6), 64-72. doi:10.1145/129888.129892

Csikszentmihalyi, M. (1975a). *Beyond boredom and anxiety*. San Francisco, CA: Jossey-Bass.

Csikszentmihalyi, M. (1975b). Play and intrinsic rewards. *Journal of Humanistic Psychology, 15*(3), 41–63. doi:10.1177/002216787501500306

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York, NY: Harper Perennial.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319-340. doi:10.2307/249008

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science, 35*(8), 982–1003. doi:10.1287/mnsc.35.8.982

Delle Fave, A., Massimini, F., & Bassi, M. (2011). *Psychological selection and optimal experience across cultures: Cross-cultural advancements in positive psychology* (Vol. 2). Dordrecht, Netherlands: Springer.

Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P., & Kaiser, S. (2012). Guidelines for choosing between multi-item and single-item scales for construct measurement: A predictive validity perspective. *Journal of the Academy of Marketing Science, 40*(3), 424-449. doi:10.1007/s11747-011-0300-3

Drachen, A., Yannakakis, G., Nacke, L. E., & Pedersen, A. L. (2010). Correlation between heart rate, electrodermal activity and player experience in first-person shooter games. *Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games* (pp. 49–54). New York, NY: ACM. doi:10.1145/1836135.1836143

Draper, J. V., Kaber, D. B., & Usher, J. M. (1998). Telepresence. *Human Factors, 40*(3), 354–375. doi:10.1518/001872098779591386

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*(1), 32-64. doi:10.1518/001872095779049543

ESA (2012). *Essential factors about the computer and video game industry*. Retrieved from http://www.theesa.com/facts/pdfs/esa_ef_2012.pdf

Evans, A. W., III. (2010). *Learning for the next generation: Predicting the usage of synthetic learning environments* (Doctoral dissertation). Retrieved from http://etd.fcla.edu/

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149-1160. doi:10.3758/BRM.41.4.1149

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*(2), 175-191. doi:10.3758/BF03193146

Ferguson, C. J. (2013). Violent video games and the Supreme Court: Lessons for the scientific community in the wake of Brown v. Entertainment Merchants Association. *American Psychologist*, *68*(2), 57-74. doi:10.1037/a0030597

Ferguson, C. J., Garza, A., Jerabeck, J., Ramos, R., & Galindo, M. (2013). Not worth the fuss after all? Cross-sectional and prospective data on violent video game influences on aggression, visuospatial cognition and mathematics ability in a sample of youth. *Journal of Youth and Adolescence, 41*(1), 109-122. doi:10.1007/s10964-012-9803-6

"Frame Rate". (n.d.). Frame rate. Wikipedia. Retrieved May 3, 2014 from http://en.wikipedia.org/wiki/Frame_rate

Fu, F.-L., Su, R.-C., & Yu, S.-C. (2009). EGameFlow: A scale to measure learners' enjoyment of e-learning games. *Computers & Education, 52*(1), 101–112. doi:10.1016/j.compedu.2008.07.004

G*Power (Version 3.1) [Computer software]. Dusseldorf, Germany: Heinrich-Heine-University of Dusseldorf. Retrieved from http://www.gpower.hhu.de/

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & Gaming, 33*(4), 441–467. doi:10.1177/1046878102238607

Gaudiosi, J. (2012, July). New reports forecast global video game industry will reach $82 billion by 2017. *Forbes*. Retrieved from http://www.forbes.com/sites/johngaudiosi/2012/07/18/new-reports-forecasts-global-video-game-industry-will-reach-82-billion-by-2017/

Gibson, E. J., Gibson, J. J., Smith, O. W., & Flock, H. (1959). Motion parallax as a determinant of perceived depth. *Journal of Experimental Psychology, 58*(1), 40-51. doi:10.1068/p080125

Gibson, J. J. (1950). *The perception of the visual world*. Oxford, England: Houghton Mifflin.

Gilbert, B. (2014, March). ‘Project Morpheus’ is Sony’s virtual reality headset for the PlayStation 4. *Engadget*. Retrieved from http://www.engadget.com/2014/03/18/playstation-virtual-reality/

Girard, C., Ecalle, J., & Magnan, A. (2012). Serious games as new educational tools: How effective are they? A meta-analysis of recent studies. *Journal of Computer Assisted Learning*, *29*(3), 207–219. doi:10.1111/j.1365-2729.2012.00489.x

Glisky, M. L., Tataryn, D. J., Tobias, B. A., Kihlstrom, J. F., & McConkey, K. M. (1991). Absorption, openness to experience, and hypnotizability. *Journal of Personality and Social Psychology*, *60*(2), 263-272. doi:10.1037/0022-3514.60.2.263

GOzzi. (2014, April 16). How to change the darkness in Sonic Ether shaders preview 2 [Video file]. Retrieved from http://youtu.be/TGZej3gZJPA

Hamari, J., & Koivisto, J. (2014). Measuring flow in gamification: Dispositional Flow Scale-2. *Computers in Human Behavior, 40*, 133–143. doi:10.1016/j.chb.2014.07.048

Held, R. M., & Durlach, N. I. (1992). Telepresence. *Presence, 1*(1), 109-112.

Hendrix, C., & Barfield, W. (1995). Presence in virtual environments as a function of visual and auditory cues. *Proceedings of the ’95 Virtual Reality Annual International Symposium ’95* (pp. 74–82). New York, NY: IEEE. doi:10.1109/VRAIS.1995.512482

Huynh-Thu, Q., & Ghanbari, M. (2008). Temporal aspect of perceived quality in mobile video broadcasting. *IEEE Transactions on Broadcasting*, *45*(3), 641-651. doi:10.1109/TBC.2008.2001246

IJsselsteijn, W., de Kort, Y., Poels, K., Jurgelionis, A., & Bellotti, F. (2007). Characterising and measuring user experiences in digital games. *Proceedings of the '07 ACE Conference*. New York, NY: ACM.

Jackson, S. A., & Eklund, R. C. (2002). Assessing flow in physical activity: The Flow State Scale-2 and Dispositional Flow Scale-2. *Journal of Sport & Exercise Psychology, 24*(2), 133–150.

Jackson, S. A., & Eklund, R. C. (2004). *The flow scales manual*. Morgantown, WV: Fitness Information Technology.

Jackson, S. A., & Marsh, H. (1996). Development and validation of a scale to measure optimal experience: The Flow State Scale. *Journal of Sport & Exercise Psychology*, *18*(1), 17-35.

Jackson, S. A., Eklund, B., & Martin, A. (2010a). *The FLOW manual: The manual for the Flow Scales*. Menlo Park, CA: Mind Garden.

Jackson, S. A., Eklund, R. C., & Martin, A. J. (2010b). *The Flow Scales instrument and scoring guide*. Menlo Park, CA: Mind Garden.

Jackson, S. A., Kimiecik, J. C., Ford, S. K., & Marsh, H. W. (1998). Psychological correlates of flow in sport. *Journal of Sport & Exercise Psychology, 20*(4), 358-378.

Jackson, S. A., Martin, A. J., & Eklund, R. C. (2008). Long and short measures of flow: The construct validity of the FSS-2, DFS-2, and new brief counterparts. *Journal of Sport and Exercise Psychology*, *30*(5), 561–587.

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies, 66*(9), 641–661. doi:10.1016/j.ijhcs.2008.04.004

Jo, N. Y., Lee, K. C., & Park, B.-W. (2011). Exploring the optimal path to online game loyalty: Bayesian networks versus theory-based approaches. In T.-h. Kim, H. Adeli, R. J. Robles, & M. Balitanas (Eds.), *Communications in Computer and Information Science: Vol. 151. Ubiquitous Computing and Multimedia Applications* (pp. 428-437). Heidelberg, Germany: Springer. doi:10.1007/978-3-642-20998-7_51

Karyonix. (2014, May 1). [1.7.2/1.7.4/1.7.5]SHADERS MOD V2.3.12 (UPDATED BY KARYONIX) [Msg 1]. Message posted to http://www.minecraftforum.net/topic/1544257-172174175shaders-mod-v2312-updated-by-karyonix/

Kato, P. M., Cole, S. W., Bradlyn, A. S., & Pollock, B. H. (2008). A video game improves behavioral outcomes in adolescents and young adults with cancer: A randomized trial. *Pediatrics, 122*(2), e305-e317. doi:10.1542/peds.2007-3134

Kennedy, R. S., Berbaum, K. S., & Lilienthal, M. G. (1997). Disorientation and postural ataxia following flight simulation. *Aviation, Space, and Environmental Medicine, 68*(1), 13-17.

Kennedy, R. S., Drexler, J. M., Compton, D. E., Stanney, K. M., Lanham, Susan, D., & Harm, D. L. (2001). *Configural scoring of simulator sickness, cybersickness and space adaptation syndrome: Similarities and differences?* (Report No. 20100033371). Retrieved from NASA Technical Reports Server: http://ntrs.nasa.gov/search.jsp?R=20100033371

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology, 3*(3), 203-220. doi:10.1027/2192-0923/a000022

Kihlstrom, J. F. (2011). *Tellegen Absorption Scale*. Faculty home page, Department of Psychology, University of California, Berkeley. Retrieved from http://socrates.berkeley.edu/~kihlstrm/TAS.htm

Kremen, A. M., & Block, J. (2001). Absorption: Construct explication by Q-sort assessments of personality. *Journal of Research in Personality, 36*(3), 252-259. doi:10.1006/jrpe.2001.2344

Krijin, M., Emmelkamp, P. M. G., Olafsson, R. P., & Biemond, R. (2004). Virtual reality exposure therapy of anxiety disorders: A review. *Clinical Psychology Review, 24*(3), 259-281. doi:10.1016/j.cpr.2004.04.001

Lee, K. M. (2004). Presence, explicated. *Communication Theory, 14*(1), 27–50. doi:10.1111/j.1468-2885.2004.tb00302.x

Levillain, F., Orero, J. O., Rifqi, M., & Bouchon-Meunier, B. (2010). Characterizing player's experience from physiological signals using fuzzy decision trees. *2010 IEEE Symposium on Computational Intelligence and Games* (pp. 75-82). New York, NY: IEEE. doi:10.1109/ITW.2010.5593370

Lewis, J. R., & Sauro, J. (2009). The factor structure of the System Usability Scale. In M. Kurosu (Ed.), *Lecture Notes in Computer Science: Vol. 5619. Human Centered Design* (pp. 94-103). Heidelberg, Germany: Springer. doi:10.1007/978-3-642-02806-9_12

Ling, Y., Brinkman, W.-P., Nefs, H. T., & Heynderickx, I. (2012). Effects of stereoscopic viewing on presence, anxiety, and cybersickness in a virtual reality environment for public speaking. *Presence, 21*(3), 254–267.

Locke, E. A., Shaw, K. N., Saari, L. M., & Latham, G. P. (1981). Goal setting and task performance: 1969-1980. *Psychological Bulletin, 90*(1), 125-152. doi:10.1037/0033-2909.90.1.125

Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication, 3*(2). doi:10.1111/j.1083-6101.1997.tb00072.x

Martin, A. J., & Jackson, S. A. (2008). Brief approaches to assessing task absorption and enhanced subjective experience: Examining 'short' and 'core' flow in diverse performance domains. *Motivation and Emotion, 32*(3), 141-157. doi:10.1007/s11031-008-9094-0

"Minecraft Forge Downloads". (n.d.). Minecraft Forge. Retrieved May 3, 2014 from http://files.minecraftforge.net/

Minsky, M. (1980, June). Telepresence. *Omni Magazine, 2*(9), 45–51. Retrieved from http://web.media.mit.edu/~minsky/papers/Telepresence.html Mizobata, R., Silpasuwanchai, C., & Ren, X. (2014). Only for Casual Players? Investigating player differences in full-body game interaction. *Proceedings of the Second International Symposium of Chinese CHI '14* (pp. 57-65). New York, NY: ACM. doi:10.1145/2592235.2592244

"MPQ Overview." (2011). *University of Minnesota Press*. Retrieved from http://www.upress.umn.edu/test-division/mpq

"MPQ Standard." (2011). *University of Minnesota Press*. Retrieved from http://www.upress.umn.edu/test-division/mpq/copy_of_mpq_BF-overview

Nacke, L., Stellmach, S., Sasse, D., & Lindley, C. A. (2009). Gameplay experience in a gaze interaction game. *The 5th Conference on Communication by Gaze Interaction* (pp. 49-54).

Nolan, C. (Producer/Director)., & Thomas, E. (Producer). (2010). *Inception* [Motion picture]. United States: Warner Bros. Pictures.

"Normal Map." (n.d.). *Polycount Wiki*. Retrieved May 3, 2014 from http://wiki.polycount.com/NormalMap/

"OptiFine Downloads". (n.d.). OptiFine. Retrieved May 3, 2014 from http://www.optifine.net/downloads.php

Parsons, T. D., & Rizzo, A. A. (2008). Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis. *Journal of Behavior Therapy and Experimental Psychiatry, 39*(3), 250-261. doi:10.1016/j.jbtep.2007.07.007

Pavlas, D. (2010). *A model of flow and play in game-based learning: The impact of game characteristics, player traits, and player states* (Doctoral dissertation). University of Central Florida, Orlando, FL.

Pavlas, D., Heyne, K., Bedwell, W., Lazzara, E., & Salas, E. (2010). Game-based learning: The impact of flow state and videogame self-efficacy. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 54*(28), 2398–2402. doi:10.1177/154193121005402808

Pavlas, D., Jentsch, F., Salas, E., Fiore, S. M., & Sims, V. (2012). The Play Experience Scale: Development and validation of a measure of play. *Human Factors, 54*(2), 214-225. doi:10.1177/0018720811434513

Peres, S. C., Pham, T., & Phillips, R. (2013). Validation of the System Usability Scale (SUS): SUS in the wild. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 57*(1), 192-196. doi:10.1177/1541931213571043

Persson, M. (2009). *Minecraft* [PC game]. Stockholm, Sweden: Mojang.

Petty, R. E., & Cacioppo, J. T. (1984). The effects of involvement on responses to argument quantity and quality: Central and peripheral routes to persuasion. *Journal of Personality and Social Psychology*, *46*(1), 69–81. doi:10.1037/0022-3514.46.1.69

Plante, C. (2015, January). Minecraft — more than anything else — could make HoloLens a hit, and here's why: Minecraft could be the FreeCell of HoloLens. *The Verge*. Retrieved from http://www.theverge.com/2015/1/21/7868363/minecraft-hololens-microsoft-freecell

Popa, J. (2014, August). Oculus Rift: DK1 vs DK2. *in2gpu*. Retrieved from http://in2gpu.com/2014/08/10/oculus-rift-dk1-vs-dk2/

Prensky, M. (2001). “Simulations”: Are they games? *Digital game-based learning* (2007 reprint, pp. 210–220). St. Paul, MN: Paragon House.

Procci, K., & Bowers, C. (2011). An examination of flow and immersion in games. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *55*(1), 2183-2187. doi:10.1177/1071181311551455

Procci, K., & Bowers, C. (2014). Instructional and gaming elements: What are they and where do they intersect? In T. S. Hussain & S. L. Coleman (Eds.), *Design and development of training games: Practical guidelines from a multidisciplinary perspective* (pp. 55-92). New York, NY: Cambridge University Press.

Procci, K., Bohnsack, J., & Bowers, C. (2011). Patterns of gaming preferences and serious game effectiveness. In R. Shumaker (Ed.), *Lecture Notes in Computer Science: Vol. 6774. Virtual and Mixed Reality - Systems and Applications, part II* (pp. 37-43). Heidelberg, Germany: Springer. doi:10.1007/978-3-642-22024-1_5

Procci, K., James, N., & Bowers, C. (2013). The effects of gender, age, and experience on game engagement. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 57*(1), 2132-2136. doi:10.1177/1541931213571475

Procci, K., Singer, A. R., Levy, K. R., & Bowers, C. (2012). Measuring the flow experience of gamers: An evaluation of the DFS-2. *Computers in Human Behavior, 28*(6), 2306-2312. doi:10.1016/j.chb.2012.06.039

Przybylski, A. K., Deci, E. L., Rigby, C. S., & Ryan, R. M. (2014). Competence-impeding electronic games and players' aggressive feelings, thoughts, and behaviors. *Personality Process and Individual Differences, 106*(3), 441–457. doi:10.1037/a0034820

Przybylski, A. K., Rigby, C. S., & Ryan, R. M. (2010). A motivational model of video game engagement. *Review of General Psychology, 14*(2), 154–166. doi:10.1037/a0019440

Przybylski, A. K., Ryan, R. M., & Rigby, C. S. (2009). The motivating role of violence in video games. *Personality and Social Psychology Bulletin, 35*(2), 243–259. doi:10.1177/0146167208327216

Qin, H., Rau, P.-L. P., & Salvendy, G. (2007). Player immersion in the computer game narrative. In L. Ma, M. Rauterberg, & R. Nakatsu (Eds.), *Lecture Notes in Computer Science: Vol 4740. Entertainment Computing - ICEC 2007* (pp. 458-461). Heidelberg, Germany: Springer. doi:10.1007/978-3-540-74873-1_60

Rau, P.-L. P., Peng, S.-Y., & Yang, C.-C. (2006). Time distortion for expert and novice online game players. *CyberPsychology & Behavior, 9*(4), 396–403. doi:10.1089/cpb.2006.9.396

Recchia, G. (2013, Jan.). Intrinsic/extrinsic motivation: A false dichotomy? *Gamastura*. Retrieved from http://www.gamasutra.com/blogs/GabrielRecchia/20130114/184715/Intrinsicextrinsic_motivation_A_false_dichotomy.php

Ross, S. R., & Keiser, H. N. (2014). Autotelic personality through a five-factor lens: Individual differences in flow-propensity. *Personality and Individual Differences, 59*, 3-8. doi:10.1016/j.paid.2013.09.029

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68-78. doi: 10.1037/0003-066X.55.1.68

Salen, K., & Zimmerman, E. (2004). *Rules of play: Game design fundamentals*. Cambridge, MA: MIT Press.

Sas, C., & O'Hare, G. M. P. (2004). Presence equation: An investigation into cognitive factors underlying presence. *Presence, 12*(5), 523–537. doi:10.1162/105474603322761315

Sauro, J. (2011). *A practical guide to the System Usability Scale: Background, benchmarks & best practices*. Denver, CO: Measuring Usability LLC.

Schwarzer, R., & Jerusalem, M. (1995). Generalized Self-Efficacy scale. In J. Weinman, S. Wright, & M. Johnston (Eds.), *Measures in health psychology: A user's portfolio. Causal and control beliefs* (pp. 35-37). Windsor, UK: NFER-Nelson.

"SEUS Modifications" (2014, January 17). SEUS v10.1 Modifications and fixes FPS Improved Clouds fixed. *Pastebin*. Retrieved May 3, 2014 from http://pastebin.com/XqCb2n39

Sharek, D., & Wiebe, E. (2012). Embedding secondary tasks in video games to measure real-time cognitive load: An approach to developing adaptive video games. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 56*(1), 2147–2151. doi:10.1177/1071181312561455

Sheridan, T. B. (1992). Musings on telepresence and virtual presence. *Presence, 1*(1), 120–125.

Sheridan, T. B. (1994). Further musings on the psychophysics of presence. *Proceedings of IEEE International Conference on Systems, Man and Cybernetics* (pp. 1073–1077). New York, NY: IEEE. doi:10.1109/ICSMC.1994.399986

Shernoff, D. J., Csikszentmihalyi, M., Schneider, B., & Shernoff, E. S. (2003). Student engagement in high school classrooms from the perspective of flow theory. *School Psychology Quarterly, 18*(2), 158–176. doi:10.1521/scpq.18.2.158.21860

Sims, V. K., Moshell, J. M., Hughes, C. E., Cotton, J. E., & Xiao, J. (2001). Salient characteristics of virtual trees. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 45*(27), 1935-1938. doi:10.1177/154193120104502717

Singer, M. J., & Witmer, B. G. (1999). On selecting the right yardstick. *Presence, 8*(5), 566–573. doi:10.1162/105474699566486

Sitzmann, T. (2011). A meta-analytic examination of the instructional effectiveness of computer-based simulation games. *Personnel Psychology, 64*(2), 489–528. doi:10.1111/j.1744-6570.2011.01190.x

Slater, M. (1999). Measuring presence: A response to the Witmer and Singer Presence Questionnaire. *Presence, 8*(5), 560–565. doi:10.1162/105474699566477

Slater, M. (n.d.). *Notes on presence*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.100.3517&rep=rep1&type=pdf

Slater, M., & Steed, A. (2000). A virtual presence counter. *Presence*, *9*(5), 413–434. doi:10.1162/105474600566925

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence, 6*(6), 603–616.

Slater, M., Usoh, M., & Steed, A. (1994). Depth of presence in virtual environments. *Presence, 3*(2), 130–144.

Sonic Ether. (2014, March 12). v10.1 Preview 2 is ready! [Msg 1]. Message posted to http://sonicether.com/forum/viewtopic.php?f=5&t=168&sid=8deaca988c2e66f12900f89f179d66b8#p643

SPSS Statistics (Version 20) [Computer software]. Armonk, NY: IBM.

Stanney, K. M., & Kennedy, R. S. (1997). The psychometrics of cybersickness. *Communications of the ACM, 40*(8), 67-68. doi:10.1145/257874.257889

SycloneSJS. (2014). It's been a while but latest update of Chroma Hills resource pack is finally out! Check it out.....[Msg] Message posted to http://www.reddit.com/r/Minecraft/comments/1y0xov/its_been_a_while_but_latest_update_of_chroma/cfgmdqp

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson.

Tellegen, A., & Atkinson, G. (1974). Openness to absorption and self-altering experiences (“absorption”), a trait related to hypnotic susceptibility. *Journal of Abnormal Psychology, 83*(3), 268–277. doi:10.1037/h0036681

Treisman, A. M. (1969). Strategies and models of selective attention. *Psychological Review, 76*(3), 282-299. doi:10.1037/h0027242

Usoh, M., Arthur, K., Whitton, M. C., Bastos, R., Steed, A., Slater, M., & Brooks, F. P. (1999). Walking > walking-in-place > flying, in virtual environments. *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques at SIGGRAPH '99* (pp. 359–364). New York, NY: ACM. doi:10.1145/311535.311589

Venkatesh, V. (2000). Determinants of perceived ease of use: Integrating control, intrinsic motivation, and emotion into the Technology Acceptance Model. *Information Systems Research, 11*(4), 342–365. doi:10.1287/isre.11.4.342.11872

Venkatesh, V., & Bala, H. (2008). Technology Acceptance Model 3 and a research agenda on interventions. *Decision Sciences*, *39*(2), 273–315. doi:10.1111/j.1540-5915.2008.00192.x

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. *Management Science, 46*(2), 186–204. doi:10.1287/mnsc.46.2.186.11926

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425–478.

Wang, C. K. J., Liu, W. C., & Khoo, A. (2009). The psychometric properties of dispositional Flow Scale-2 in internet gaming. *Current Psychology*, *28*(3), 194–201. doi:10.1007/s12144-009-9058-x

Wiebe, E. N., Lamb, A., Hardy, M., & Sharek, D. (2014). Measuring engagement in video game-based environments: Investigation of the User Engagement Scale. *Computers in Human Behavior, 32*, 123-132. doi:10.1016/j.chb.2013.12.001

Wild, C. T., Kuiken, D., & Schopflocher, D. (1995). The role of absorption in experiential involvement. *Journal of Personality and Social Psychology, 69*(3), 569–579. doi:10.1037/0022-3514.69.3.569

Wilson, S. C., & Barber, T. X. (1978). The Creative Imagination Scale as a measure of hypnotic responsiveness: Applications to experimental and clinical hypnosis. *The American Journal of Clinical Hypnosis, 20*(4), 235–249. doi:10.1080/00029157.1978.10403940

Wirth, W., Hartmann, T., Böcking, S., Vorderer, P., Klimmt, C., Schramm, H., … Jäncke, P. (2007). A process model of the formation of spatial presence experiences. *Media Psychology, 9*(3), 493–525. doi:10.1080/15213260701283079

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence, 7*(3), 225–240. doi:10.1162/105474698565686

Witmer, B. G., Jerome, C. J., & Singer, M. J. (2005). The factor structure of the Presence Questionnaire. *Presence, 14*(3), 298–313. doi:10.1162/105474605323384654

Yi, M. Y., & Hwang, Y. (2003). Predicting the use of web-based information systems: Self-efficacy, enjoyment, learning goal orientation, and the Technology Acceptance Model. *International Journal of Human-Computer Studies, 59*(4), 431–449. doi:10.1016/S1071-5819(03)00114-9

Zaichkowsky, J. L. (1985). Measuring the involvement construct. *Journal of Consumer Research, 12*(3), 341–352. doi:10.1086/208520

Zaichkowsky, J. L. (1994). The Personal Involvement Inventory: Reduction, revision, and application to advertising. *Journal of Advertising, 23*(4), 59–70. doi:10.1080/00913367.1943.10673459