DISSERTATION

VIDEOGAME-BASED LEARNING:

A COMPARISON OF DIRECT AND INDIRECT EFFECTS ACROSS OUTCOMES

Submitted by

Diana R. Sanchez

Department of Psychology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2017

Doctoral Committee:

Advisor: Kurt Kraiger
Co-Advisor: Alyssa Gibbons

Travis Maynard
Lucy Troup
Mark Prince

ABSTRACT

VIDEOGAME-BASED LEARNING:
A COMPARISON OF DIRECT AND INDIRECT EFFECTS ACROSS OUTCOMES

Recent years have shown a rise in the application of serious games used by organizations to help trainees learn and practice job related skills (Muntean, 2011). Some sources have projected a continued growth in the development and application of video games for novel purposes (Sanders, 2015). Despite the increasing use of video games for workplace training, there is limited research evidence to justify the use of video games for learning. Additionally, this research has generated mixed results on the utility of serious games (Guillen-Nieto & Aleson-Carbonell, 2012).

I review the research literature to understand why videogame-based learning research is producing inconsistent results. From this review, I identify several current challenges in the research literature that may be contributing to these inconsistencies; distinguishing videogames from similar training media, identifying game characteristics, exploring the possible mechanisms in the training experience, differentiating training outcomes, and making accurate implications for research.

I then present results from an empirical study that addresses these challenges. I tested and expanded the model from Garris et al.'s (2002) game-based learning I-P-O model to determine the extent to which one game characteristic (i.e., human interaction) influences two training outcomes (i.e., declarative knowledge and affective states), as well as the possible mechanisms through which this occurs. I found that active learning is a mechanism through which human interaction influences both declarative knowledge and affective states. Although the effect size was large for affective states, it was small for declarative knowledge. The mediating effect of active learning was greater for the relationship between human interaction and affective states than for the relationship between human interaction and declarative knowledge. I also found that perceived value of the training mediates the relationship between human interaction and affective states.

TABLE OF CONTENTS

INTRODUCTION

Recent years have shown a rise in the application of videogames for purposes other than entertainment (Muntean, 2011). Sources have projected continued growth in the development and application of videogames for novel purposes, such as being used in the classroom to teach elementary school children, in interactive displays for museum exhibits, and within organizations to train and develop employees (Sanders, 2015). Despite the increasing use of videogames for training purposes, the research evidence used to justify the investment and implementation of videogame-based learning is limited (Guillen-Nieto & Aleson-Carbonell, 2012). Research on videogame-based learning has found evidence that both supports and discredits the claim that videogames enhance learning outcomes (Coller & Scott, 2009; Wrzesien & Raya, 2010). One contribution of this study is a review of the research literature to understand why videogame-based learning research is producing inconsistent results. From this review, I generate several current challenges in the research literature that may be contributing to these inconsistencies; distinguishing videogames from similar training media, identifying game characteristics, exploring the possible mechanisms in the training experience, differentiating training outcomes, and making accurate implications for research. The primary purpose of this study is to present and test an alternative methodology for videogame-based learning that could help explain previous contradictory findings and clarify future research results. This proposed methodology is an extension of a model of videogame-based learning (Garris, Ahlers, & Driskell, 2002) that isolates game characteristics and identifies specific mechanisms through which different training outcomes are affected.

Videogames share similarities to simulations, which are a replication of reality in which the individual assumes a specific role and applies knowledge or skill in a scenario over time (Adams, Mayer, MacNamara, Koenig, & Wainess, 2012; Beal & Christ, 2004; Browning, Ryan, Scott, & Smode, 1977; Department of Defense, 1997; Gredler, 1996; Hays, 2005; Lypaczewski, Jones, & Voorhees, 1987). Institutions have a long history of using simulations for instructional purposes (Guillen-Nieto & Aleson-Carbonell, 2012; Prensky, 2001). One example is *MONOPLOGS* that the Air Force used in the 1950's to train soldiers on supply chain management (Videogame Historian, n.d.). Integrating simulations into a videogame environment began in the 1970's when videogames such as *Space Fortress* were used to train and assess flight skills (Guetzkow, 1959; Megginson, 1959). A videogame is a virtually-constructed environment with rules and goals that challenge the player while engaging the attention with a serious of decisions (Calvillo- Gámez, Cairns, & Cox, 2015; Hays, 2005; Pagulayan, Keeker, Wixon, Romero, & Fuller, 2002). Videogame environments are well-suited for teaching because they provide a high degree of control over the features and characteristics of the game. This flexibility is beneficial to organizations because game scenarios can be easily adapted to different needs (Guillen-Nieto & Aleson-Carbonell, 2012; Mayo, 2007). A further benefit has been that the burgeoning videogame industry (Lemay & Maheux-Lessard, 2015) has caused an influx of videogame designers and improved design software, making videogame development more accessible to organizations (Dieker, Rodriguez, Lignugaris, Hynes, & Hughes, 2014).

Videogame-Based Learning

The benefits and accessibility of videogames have contributed to an ongoing use of videogames for learning purposes (Adams, 2013). Despite their longstanding use, there is inconsistent research evidence that videogames are effective teaching tools (Butler, Someya, &

Fukuhara, 2014; Cannon-Bowers, 2006; Guillen-Nieto & Aleson-Carbonell, 2012; Hays, 2005). These inconsistencies have led to a contentious debate on the effectiveness of videogame-based learning methodology (Adams, 2013; Butler et al., 2014; Garris et al., 2002; Nash, 2005).

> *"I challenge anyone to show me a literature review of empirical studies about game-based learning. There are none. We are charging head-long into game-based learning without knowing if it works or not. We need studies."* (Cannon-Bowers, 2006)

> *"…advocacy research is reflected in papers and books in which game proponents extol the potential of games to transform education but do not offer adequate empirical evidence. The hallmark of advocacy research is strong claims coupled with weak evidence."* (Mayer, 2011, p. 300)

> *"There is still little consensus on the game features supporting learning effectiveness, the process by which games engage learners, and the types of learning outcomes that can be achieved through game play."* (Guillén-Nieto & Aleson-Carbonell, 2012, p. 436).

The debate on the effectiveness of videogame-based learning has been fueled by contradictory findings. Some researchers have found that videogames are effective learning tools (Blunt, 2009; Guillen-Nieto & Aleson-Carbonell, 2012; Mayo, 2007), improving the acquisition of skills such as writing and negotiating (Rising, 2009). Other researchers have found that games do not impact learning (Clark, Yates, Early, & Moulton, 2010; Kickmeier-Rust, Peirce, Conlan, Schwarz, Verpoorten, & Albert, 2007; O'Neil & Perez, 2008; O'Neil, Wainess, & Baker, 2005; Tobias & Fletcher, 2011; Wrzesien & Raya, 2010). While multiple reviewers have purported there is no clear relationship between videogame-based learning and learning outcomes (Dempsey, Rasmussen, & Lucassen, 1996; Randel, Morris, Wetzel, & Whitehall, 1992; Vogel, Vogel, Cannon-Bowers, Bowers, Muse, & Wright, 2006), others have labeled these authors as *skeptics* (Ke, 2008) citing studies that have shown games are effective for improving engagement (Malone, 1981; Rieber, 1996), collaboration (Kaptelinin & Cole, 2002), and learning (Connolly, Boyle, MacArthur, Hainey, & Boyle, 2012; Ricci, Salas, & Cannon-Bowers, 1996). Issues with the methodological design of these studies may be precipitating these contradictory results. Thus, it is important to review how videogame-based learning is being studied to better understand the potential sources of these inconsistencies.

Current research on videogame-based learning often applies the methodological framework proposed by Garris et al. (2002). This model produced a videogame-specific version of the Input-Process-Output (IPO) model first proposed by McGrath (1984) to study team interactions. Both models feature a progression of three steps in which specific *inputs* lead to *processes*, which produce *outputs*. In Garris et al.'s version of the model, the *input* step consists of *instructional content* (e.g., skills being taught) and *game characteristics* (e.g., degree of fantasy in the game design). Several studies have produced taxonomies of game characteristics to help define and distinguish these features and to provide standardized language for the field of videogame-based learning (e.g., Bedwell, Pavlas, Heyne, Lazzara, & Salas, 2012; Connolly et al., 2012; O'Neil et al., 2005). The *process* step goes by many names in the literature including *game cycle*, *game play*, and *user experience*. This step includes *user judgment* (e.g., interest), *user behaviors* (e.g., decisions made) and *system feedback* (e.g., changes produced by the user's actions). Garris et al. (2002) referred to the *process* step as the key component of the model, believing it could lead to a judgment-behavior-feedback loop, motivating users to continually return to the game and reengage with the activities. The *output* step consists of *skill-based learning* (e.g., motor or technical skills), *cognitive-based learning* (e.g., declarative knowledge), and *affective learning* (e.g., attitudes).

This model has been particularly important to the gaming research literature because it provided researchers with a framework to build additional evidence for understanding the progression of videogame-based learning from the first step of design through the final step of evaluation (e.g., Connolly et al., 2012; Ke, 2008; O'Neil et al., 2005; Wilson et al., 2009). Other researchers have spent the last several years applying this model to help understand how videogame characteristics influence training outcomes (Connolly et al., 2012; Grossman, Heyne, & Salas, 2014; Hoffman & Nadelson, 2010; Iacovides, McAndrew, Scanlon, & Aczel, 2014; Pratt & Hahn, 2015). This had led to the prolific generation of videogame-based learning research across industries (e.g., Faria, Hutchinson, Wellington, & Gold, 2009; Gee, 2005; Hsu, 1989; Jabbar & Felicia, 2015; Li & Tsai, 2013).

An unfortunate outcome of this prolific line of research has been the increasing number of contradictory results regarding the effectiveness of videogame-based learning. It is clear that a new approach is needed to evaluating videogame-based learning. Based on a review of the literature, I identify current challenges to research that may be propagating the concerns and controversies in this field. I present these challenges along with recommendations on how to address these concerns in future research methodology.

Challenge 1: Distinguishing Training Media. Training media is defined as the format that is used to deliver the training content, such as using classroom instruction or computer-based training. There is a need for researchers to define and distinguish videogames from other types of training media. It is important that researchers provide clear and accurate definitions because mislabeling results can lead to inaccurate findings in which outcomes are misattributed to the wrong type of training media (e.g., Kobes, Helsloot, de Vries, & Post, 2010). This can create challenges for researchers when attempting to interpret the aggregated results of videogame-based learning studies as is done in a meta-analysis or a review. If a researcher cannot accurately parse apart the results of a videogame-based learning tool from another form of media, their findings will likely lead to inconclusive or erroneous interpretations due to contaminated data.

Clear distinctions are imperative for preventing these issues of misattribution or contamination (Clark, 2001).

Previous researchers have addressed the need to provide clear distinctions between different forms of training media (Clark, 2001; Greenblat, 1981; Hays, 2005). To prevent issues of mislabeling, researchers have presented guidelines for defining and distinguishing similar types of training media (e.g., Crookall & Thorngate, 2009; Fletcher & Tobias, 2006; Gredler, 1996; Wilson et al., 2009). However, some researchers have still used terms like videogame and simulation interchangeably (Amer, Denehy, Cobb, Dawson, Cunningham-Ford, & Bergeron, 2011; Greenblat & Duke, 1981; Rieber, 1996; Thomas, Cahill & Santilli, 1997). For example, one paper reported a shear bonding dental simulation as a videogame although it lacked the major characteristics of a game such as presenting constraints and challenges with an objective beyond practicing a skill (Amer et al., 2011). This lack of distinction among training media can be difficult to avoid, given how blurred the lines are between different types of media. For example, in the videogame *Zero Hour*, individuals take on the role and practice the skills of an Emergency Medical Service responder (Virtual Heroes, 2007). Although this videogame meets several components in the definition of a simulation, such as taking on a role and applying skills in a replication of real-life, it also includes primary components of a game such as increasing levels of difficulty, playing with an avatar, and active decision making in an interactive virtual world (Bedwell et al., 2012; Garris et al., 2002). Some papers have presented guidelines intended to identify the distinctions between pure and mixed training media such as distinguishing pure videogames, mixed simulation games, and pure simulations (Hays, 2005; Leemkuil, de Jong, & Ootes, 2000). Acknowledging these distinctions is important because there is an assumption that game characteristics add an immersive feature to the activity that can change the game play experience for individuals (Garris et al., 2002; Bedwell et al., 2012).

The first challenge in videogame-based learning research is that poor distinctions have been made between videogames and other forms of training media. By grouping games with other training media, we misattribute results and may reach inaccurate conclusions. To address this challenge, it is recommended that researchers clearly define and distinguish videogames from other similar forms of training media using commonly established definitions (e.g., Gredler, 1996; Hays, 2005).

Challenge 2: Identifying Videogame Characteristics. There is a need to use research-based frameworks to identify and describe specific videogame characteristics. Several researchers have supported this claim by stating that different videogames influence training outcomes in different ways based on specific design characteristics of the game (Bedwell et al., 2012; Garris et al., 2002; Wilson et al., 2009). For example, the amount of control (i.e., multiple possible actions throughout the game) an individual has within the environment of the game is related to their level of motivation and desire to return to the game (Driskell & Dwyer, 1984; Habgood, 2005; Habgood, Ainsworth, & Benford, 2005; Lepper, 1985). Previous researchers have provided frameworks and taxonomies that can be used to distinguish different game characteristics (Bedwell et al., 2012; Wilson et al., 2009). Some of these frameworks have focused on different elements such as the structure of the game versus the player's experience in the game (Calvillo-Gámez et al., 2015; Ermi & Mäyrä, 2005; Hassenzahl, 2003; Preece, Rogers, & Sharp, 2002). Despite the different approaches these frameworks have taken to define game characteristics, their key importance is that they identify characteristics that can be used to distinguish different features of a game.

Despite the availability of these frameworks to define game characteristics and the previous connections that have been drawn between game characteristics and training outcomes, little empirical research has examined the impact of game characteristics (Bedwell et al., 2012; Wilson et al., 2009). This includes a lack of empirical evidence isolating and manipulating game characteristics and a lack of information describing game characteristics in enough detail so that game characteristics can be inferred when not explicitly identified. Rather than making these important distinctions between games based on their characteristics using the available taxonomies, researchers often group all videogames together, disregarding important distinctions that can be made. Aggregating games with different characteristics (in meta-analysis or qualitative reviews) likely leads to misleading or inconclusive evidence on the effectiveness of videogame-based learning.

The majority of studies in the videogame-based learning literature are media comparisons (i.e., comparing the effectiveness of one training format such as classroom instruction to another training format such as videogame-based learning; Clark, 1983; 1985; 2001; Mayer, 2011). The field has continued to produce studies about comparing media instead of studying what it is about the features within a medium that have an association with training outcomes. These studies have continued despite researchers such as Clark saying for the past 30+ years that these comparisons are misguided (Clark, 1983; 1985; 1994; 2001).

It is important that future research shifts away from media comparisons and towards closer examinations of game characteristics. The theoretical justification for making comparisons by game characteristics is consistent with Garris et al.'s (2002) approach that game characteristics impact the in-game processes that lead to learning outcomes. This is comparable to points made by Clark (1994) that training media are proxy measures for other characteristics and mechanisms within the training that are associated with learning outcomes.

The second challenge in videogame-based learning research is that studies do not provide detailed descriptions of game characteristics and do not study the impact of specific attributes on training outcomes. This lack of a central framework of game characteristics contributes to the unfortunate practice of aggregating evidence of videogame effectiveness. In other words, researchers are treating all videogames as members of a single category (even though they may have different features and attributes). To address this challenge, I recommend that the game characteristics be clearly defined using common taxonomies or frameworks (e.g., Bedwell et al., 2012) and that more attention be given to exploring the relationships between game characteristics and training outcomes.

Challenge 3: Exploring Learning Processes. There is a need to understand the learning processes of videogame-based training and to test complex relationships such as mediating relationships with videogames and training outcomes. In applications of the game-based IPO model (Garris et al., 2002), research has concentrated largely on the relationships between *inputs* (e.g., game characteristics) and *outputs* (e.g., training outcomes), largely neglecting the *process* through which the game cycle occurs. This has resulted in a weak empirical understanding of how videogames impact learning beyond theoretical arguments. Some researchers have argued that behavioral and attitudinal variables exist that mediate the relationship between game characteristics and training outcomes (Landers, 2014; Landers & Landers, 2015), yet there is little research that has examined these possible indirect effects. These gaps have led some researchers to label game-based training processes a black box (Arnab et al., 2015; Perrotta, Featherstone, Aston, & Houghton, 2013), while other researchers have merely posited that there are mechanisms that occur during the training experience without testing what they are and how effective they are.

One common assumption is the claim that game-based learning occurs because games are motivating, engaging, and immersive for trainees (e.g., Coller & Shernoff, 2009; Papastergiou, 2009; Prensky, 2001), and that these features of the game experience in turn improve learning. However, there is a lack of evidence to support these assertions (de Freitas & Jarvis, 2007). Not knowing the processes through which game characteristics impact training outcomes has been called out as an important gap that prevents clear implications and precludes the generation of sound recommendations for practitioners (Landers, 2014).[1] More research on the mechanisms mediating the relationship between videogame-based learning and training outcomes is necessary. For example, it will allow us to understand if games are engaging and learn the specific association that engagement has with different training outcomes such and declarative knowledge and affective states.

The third challenge in videogame-based learning research is the need to identify and test complex mechanisms through which learning occurs in videogame-based training. To address this challenge, it is recommended that potential mechanisms are identified and empirically testing as potential mediators extending Garris et al.'s (2002) IPO framework to include specific variables and their established relationships with different training outcomes.

Challenge 4: Differentiating Training Outcomes. There is a need to consistently define and differentiate training outcomes from one another (e.g., objective learning, subjective learning). Hays (2005) stated that deciding to use a videogame-based learning tool required a

[1]This statement was made in reference to gamification, which is adding videogame characteristics to various non-game media such as course curriculum or a computer-based training. The present study is examining a serious game, which is a videogame in which the primary goal is education not entertainment. Since both lines of research are similar and both rely on the IPO model from Garris et al., (2002) the conclusion is applicable to both lines of research.

leap of faith, because the evidence for games as effective teaching tools was lacking but there was evidence that games are engaging and create positive reactions. This is consistent with other researchers who have identified that videogames impact training outcomes differently (e.g., Shute, Ventura, & Ke, 2015). It further highlights the importance of measuring training outcomes consistently and accurately using validated methods (Clark, 2001). Currently there is wide variation in how training outcomes are defined and measured, ranging from using validated scales to a single-item indicator.

Measuring training outcomes is one of the most important aspects of training as it establishes whether the training has met its goals and objectives (Kraiger, 2002; Kraiger, Ford, & Salas, 1993). There are well-established research guidelines for defining and measuring training outcomes (e.g., Kirkpatrick & Kirkpatrick, 2006; Kraiger, 2002; Kraiger et al., 1993; Phillips & Phillips, 2016; Sackett & Mullen, 1993). Aligning training objectives with the most appropriate training outcome is a critical guideline for determining the effectiveness of the training (e.g., de Freitas & Jarvis, 2007). Establishing effective training evaluation provides input to ensure training improves over future iterations (Brown & Gerhardt, 2002; Kraiger, 2002).

Measures of learning assess the extent to which trainees demonstrate that they have acquired the knowledge and/or skills specified in the training objectives (Kraiger, 2002). Learning may be measured either objectively or subjectively. Objective learning is the process of assessing knowledge or skills acquired or developed in training using evaluative methods independent from human judgment (Beach, 1980). Subjective measures of learning refer to self or other assessments of trainee learning (Sitzmann, Ely, Brown, & Bauer, 2010). When the rater is the trainee, these include judgments of learning, which are defined as the level of confidence an individual has regarding his or her own knowledge or development in a specific skill set or within a certain domain of knowledge. These measures are valuable when the goals of the training are to understand an individual's perceptions of his or her own learning and are particularly useful in the research fields of memory and human cognition (Koriat, 1997; Koriat, Sheffer, & Ma'ayan, 2002; Rhodes & Tauber, 2011). Within videogame-based training, they can be useful during training as they may guide decisions trainees make about whether to repeat content or move on within the game. Despite considerable research providing guidance on training evaluation, there are still large variations in the clarity and consistency of how training outcomes are defined and measured within the context of videogame-based learning (as discussed by Clark, 2007).

Thus, the fourth challenge in game-based training research is the ambiguous definition and of training outcomes that at poorly defines or does not report clear explanation of how a study was measured. To address this challenge, it is recommended that future research provide clear descriptions on the definitions and measures used for training outcomes.

Challenge 5: Drawing Accurate Implications. There is a need to draw accurate implications from videogame-based learning studies that do not overstate or misrepresent the research findings. A concern is that objective measures of learning are used interchangeably with both subjective learning and affective reactions in the research literature; asking whether someone believes they have learned something or whether they enjoyed the training has been used to make claims that the training was an effective learning tool (e.g., Burguillo, 2010; Johnson, 2010; Levitan, 2010; Loon, Evans, & Kerridge; Vahed, 2008). In one study, researchers claimed that computer-based learning led to "increased performance" when only self-report survey responses were gathered from participants to make these claims (Burguillo, 2010). Similarly, another study claimed that student learning improved following a game-based simulation, but the only outcome measures were student responses to survey and interview questions along with self-ratings of their own learning on the different learning modules (Loon, Evans, & Kerridge, 2015). It is not uncommon for researchers to use language that implies that learning was achieved in a training in which only subjective, self-report measures were gathered. There are in fact an alarming number of examples in the research literature across industries that demonstrate inaccurate or overstated claims using subjective measures of evaluation (e.g., Aabakken, Adamsen, & Kruse, 2000; Chang, Lee, Ng, & Moon, 2003; Jeong & Bozkurt, 2014; Navarro, Baker, & Van der Hoek, 2004). To illustrate the magnitude of the problem, in a separate meta-analysis, I have found to date 49 studies in which subjective measures were the only criterion used to assess videogame learning effectiveness.

More importantly for present purposes, different evaluation criteria provide different types of information about training effectiveness and cannot be considered equivalent or comparable. In videogame-based training, the most common criteria are affective reactions and learning measures, with the latter consisting of both objective and subjective forms. Affective reactions measure the individual's response to the training, including their feelings and liking of the format and materials (Tan, Hall, & Boyce, 2003). Subjective measures of learning are valuable when applied in the correct context (Alliger, Tannenbaum, Bennett, Traver, & Shotland, 1997; Kraiger, 2002; Leung, 2015).

The ambiguity in measuring training outcomes becomes problematic when we measure one learning outcome and then draw inferences about another outcome from that same measure. For example, affective reactions can be useful for measuring whether trainees like the course, whether they will complete it, or whether they will recommend it. However, affective reactions cannot indicate whether an individual learned the intended content or can perform a new skill.

There is clear research evidence that any given training outcome measure cannot act as a proxy measure for other outcome measures (that is, different training criteria are at best mildly correlated). It is unreasonable to assume that subjective measures of learning can accurately represent an objective level of learning (Alliger et al., 1997; Chen, 2005; Clark, 2001; Hays, 2005; Hinds, 1975; Leung, 2015; O'Neil et al., 2005; Sitzmann et al., 2010). Previous research has shown that subjective assessments are not highly correlated with performance (Murphy, 2008). It is misleading when researchers claim they are measuring learning and imply objective learning, but they are only using subjective or affective reactions. Erroneous measures of learning can lead to false conclusions about the effectiveness of a training intervention, which may be contributing to the conflicting results regarding the usefulness of games as viable training options.

Thus, the fifth challenge in game-based training research is a lack of precision in differentiating training outcomes, which at times leads to inaccurate implications, misconstrued results, and overstated research findings. To address this challenge, I recommend that future research draw implications based solely on the training outcome that was measured.

A Different Approach to Videogame-Based Training

It is important to recognize that the debate on the effectiveness of videogame-based learning continues even as research accumulates. The prominent research question over the last several decades has been whether or not game-based learning is effective. However, there is clear evidence from the exhaustive number of studies that game-based learning can be a very effective learning tool in certain contexts (Adams, 2013). Continuing to pursue research following the same methodology that has been historically used will not resolve the current debate on videogame-based learning. Studies on media comparisons can provide little insight into the current debate and may thwart future research by perpetuating the current divide (Johnson & Mayer, 2010; Mayer, 2011; Mayer & Johnson, 2010).

What would contribute to understanding the research is a new approach to studying game-based learning that accounts for the challenges outlined above and applies a new methodology that reframes the primary research question. There is a need to better understand the nuances of videogame-based training and the game elements that lead to systematic differences in training outcomes. The following study was designed and developed to address these challenges and to develop methodology that could be used in future research studies. In this current study, one game characteristic (i.e., human interaction) will be evaluated to determine the direct effects on two types of training criteria (i.e., affective states and declarative knowledge). Additionally, I propose and test four indirect effects between human interaction and the two training outcomes, see Figure 1. A complete summary of my study follows.

Human Interaction

To explore the conditions under which game-based learning is effective, I will manipulate one game characteristic to determine its impact on the training outcomes. I have chosen to use the taxonomy of game characteristics (Bedwell et al., 2012) based on its rigorous development and consistent application in the literature (Aragón-Correa, Martín-Tapia, & Hurtado-Torres, 2013; Kozlowski, Chao, Grand, Braun, & Kuljanin, 2013; Salas, Shuffler, Thayer, Bedwell, & Lazzara, 2015). Using this taxonomy will allow me to frame my game within the broad context of videogames. This is needed since some researchers have posited that different features lead to different outcomes, which will help future researchers understand how the results from this particular game relate to the larger body of research on game-based learning.

In this study, I chose to focus on human interaction, or the degree of interaction that the individual has with someone else while playing the game (Bedwell et al., 2012). Although any or all of the characteristics in the taxonomy might affect learning outcomes, human interaction presents a simple and feasible intervention that organizations could easily manipulate within existing programs and could easily design into future training programs. Other game characteristics (e.g., control, environment) would require detailed design and programming changes to the games to be included in a training program or manipulated in a study. Given that this area of research is still developing, particularly in micro-examinations on the impact of game characteristics, it makes sense to initially address game characteristics that would be easiest for organizations to manipulate in application.

While individuals can interact with others in a variety of ways, a common way is as members of the same team. There are specific methods inherent to team work that make their process distinct from work performed by an individual (Kraut, Fussell, Lerch, & Espinosa, 2005). This is because teams coordinate their efforts in various ways to often divide up the workload catering to individual's strengths and preferences on tasks. To study the effects of human interaction, I will compare the learning performance of individuals who complete the training within a team to individuals who complete the videogame training alone. Logically, participants completing training alone will have no human interaction, while those completing training in a team will have the opportunity for at least some human interaction with teammates.

There have been few studies making direct comparisons between individuals who played with a team and to individual who played alone using the same work-related training protocol (e.g., Liang, Moreland, & Argote, 1995). These comparisons are very common in other research literatures such as training individuals with disabilities (e.g., Chadwick, Momcilovic, Rossiter, Stumbles, & Taylor, 2001). However, in the workplace training literature, many of the existing studies compared team-level performance to individual level performance for participants who were all trained as part of a team (e.g., Moreland & Myaskovsky, 2000). The present study examines whether interactions with others in a game-based learning environment impacts training outcomes.

I also chose to focus on human interaction because organizations are continually exploring the use of teams, particularly in virtual settings. With globalization and technology continually reshaping the progress and nature of work, it is important to explore future applications of training. There have been a handful of studies that directly compared individuals who played with a team to individuals who played alone in a training setting (e.g., Brodbeck & Greitemeyer, 2000b; Laughlin, Bonner, & Miner, 2002; Laughlin, Zander, Knievel, & Tan, 2003; Liang, Moreland, Argote, 1995; Sanchez, & Gibbons, n.d.). A similar stream of research has reported on the positive impact that teamwork can have on individual outcomes (e.g., Laughlin, Carey, & Kerr, 2008; Maciejovsky, Sutter, Budescu, & Bernau, 2013). A general trend in these studies has been that those training in a team have demonstrated greater learning outcomes than those who trained individually (Brodbeck & Greitemeyer, 2000b). Previous research has shown that positive team interactions lead to beneficial team processes such as knowledge sharing (e.g., Gibson & Cohen, 2003), transactive memory (e.g., Moreland & Myaskovsky, 2000), and information elaboration (e.g., Maynard, Mathieu, Gilson, & Sanchez, 2017) which can be beneficial towards learning (Kozlowski & Bell, 2008). I will explore the potential impact human interaction will have on different training outcomes, along with the likely mechanisms whereby these relationships are mediated.

Affective States

There are a variety of ways to operationalize affective responses from training. Affective outcomes relate to common outcomes such as changes in attitudes that trainees experience due to the training (Garris et al., 2002). In contract to this, affective reactions are trainees' specific perceptions and reactions to the training, such as whether or not they liked specific aspects of the training (Alliger et al., 1997). This may include reactions such as satisfaction or liking the training (Kraiger, 2002; Ponce, Franchak, Billings, & O'Reilly, 1981), utility (the usefulness of the training; Alliger et al., 1997), effectiveness (perceived success of the training; Morgan & Casper, 2000; Pannese & Carlesi, 2007), and ease (difficulty of the training; Warr & Bunce, 1995).

The affective measures in the present study fall between these two definitions. These affective states that I will be evaluating are specific to the training but are not an attitude change that is expected to persist long after the training. However, the measures go beyond asking simply about the trainees liking of the training, instead measuring personal meaning, enjoyment, emotional engagement and motivation that participants experience directly as a part of this activity. I will use the term affective states to represent the temporary experience that trainees experience as part of this particular videogame-based training activity. Other training researchers have presented similar approaches, such as Ford, Smith, Weissbein, Gully, and Salas (1998) who measured goal orientation and self-efficacy as mediating states following a training intervention. This direction reflects the transition that the team training literature has made moving away from the traditional IPO model towards a nuanced model of emergent states that teams experience during interactions (Curşeu, 2006). For the current study, I will be examining four affective states, discussed below.

Psychological Meaning. Psychological meaning is defined as an individual's positive evaluation of their experience in the videogame-based training activity based on their value system (May, Gilson, & Harter, 2004). Psychological meaning benefits training via an association with increased motivation (May et al., 2004). I anticipate that individuals who train with a team will report greater perceptions of meaningfulness from the training. This is based on previous research findings that working with others in a game can improve the meaningfulness of the content being taught (Plass et al., 2013).

Perceived Enjoyment. Perceived enjoyment is defined as satisfaction with the training from attaining personal needs or expectations in the training (Ponce et al., 1981). Previous research has shown that when grouped with a virtual team, individuals reported higher levels of enjoyment when placed with group (Guillén-Nieto & Aleson-Carbonell, 2012). Based on this I anticipate that participants who are placed within a team will report significantly higher levels of enjoyment from the training.

Emotional Engagement. Emotional engagement is the level of emotional arousal experienced from the training (Rich, Lepine, & Crawford, 2010). Previous research has shown that individuals experienced more engagement through their participation with a team (Guillén-Nieto & Aleson-Carbonell, 2012). I believe that individuals who train with a team will feel more emotional engagement in the training.

Motivation. A common claim is that games are inherently motivating and they involve trainees in the content and gameplay (Rieber, 1996). Further, other researchers have shown that interaction with others during a game was motivating to the player's experience (Plass et al., 2013). This increased motivation has been shown to relate to other positive outcomes

(Bernhaupt, 2010; Blair, 2011; Leung, 2015; Patrick, 1992). Based on these findings I believe that human interaction will lead to increased levels of motivation.

Affective States as a Latent Construct

All of the variables above describe a positive affective state that a trainee could experience, indicating a positive evaluation of their own personal experience during the videogame-based training. While some researchers have argued in favor of measuring discrete dimensions of affective measures (Tannenbaum & Yukl, 1992; Warr & Bunce, 1995), there is much evidence that affective measures are positively and significantly correlated to one another (Alliger et al., 1997; Morgan & Casper, 2000). Other researchers have grouped multiple affective measures together by providing evidence that they loaded onto the same factor (Tan et al., 2003). Based on this evidence that affective measures are correlated and load onto the same factor, I have little reason to believe that there would be large differences in examining affective states. Given that affective measures in the traditional sense tend to be positively correlated with one another, and for the sake of parsimony, I will treat the above measures of affective states as indicators of a broader latent variable, provided that they load as expected.

There is not strong theoretical justification for why perceptions about meaning, enjoyment, emotional engagement and motivation should have different relationships with particular training outcomes. Because my hypotheses predicted similar patterns of relationships for each of the reactions, I will propose a higher order factor. I propose that affective states in the training will be indicated by higher levels of psychological meaning, greater perceived enjoyment, higher emotional engagement, and higher levels of motivation in the training.

Hypothesis 1: Participants who complete the training as part of a team will report significantly higher positive affective states than participants who complete the training alone.

Declarative Knowledge

Declarative knowledge is the factual information an individual is able to retain and recall from the videogame-based training (Alexander & Judy, 1988; Anderson, 1983). Findings from previous research have been mixed on whether human interaction improves or impedes performance. While some studies have shown that individuals demonstrate superior performance when compared to individuals trained as a team (Dillon, Graham, & Aidells, 1972; Sanchez & Gibbons, 2017), other researchers have found that interactions with other players during a game create a beneficial environment in which players can guide, explain or clarify things that are misunderstood in the game (Morrison, 2010). Researchers who posit that human interactions aid learning often cite claims that human interaction aids the learning process through increased access to resources and input from others (Ma & Yuen, 2011). This is consistent with the Online Knowledge Sharing Model (OKSM), which proposes that individual learning in an online context occurs through online knowledge sharing which encompasses complex behaviors beyond meeting and discussing the content (Argote, 1999; Darr & Kurtzberg, 2000; Ko, Kirsch, & King, 2005; Ma & Yuen, 2011). This theory states that individuals develop a detailed understanding of the content and its implications through additional interactions such as observing and mimicking through frequent interactions with others (Ramos & Yudko, 2008). Understandably this would have a positive impact on the players who are receiving assistance and possibly a negative impact on those giving assistance. When further examined, there are many potential benefits that can occur when participants are placed within a team for training.

For example, teams can engage in transactive memory, in which the coordination between team members allows them to fill in knowledge or skill gaps of other team members (Kraut et al., 2005; Moreland & Myaskovsky, 2000). Other researchers have shown that when teams coordinate their efforts in an effective way, this can positively affect their efficiency, quality of work and organizational level outcomes such as profits and board evaluations (Gittell, 2002; Kraut et al., 2005). Other researchers have aggregated research evidence to demonstrate that team-based training is effective for numerous positive outcomes including cognitive outcomes, affective outcomes, teamwork processes and performance outcomes (Salas et al., 2008). In review, the majority of previous studies find that human interaction will have a positive impact on learning outcomes (de Freitas & Routledge, 2013; MacStewart, 2010). Based on these findings I predict that human interaction will promote learning and training performance during the videogame-based training simulation. I operationalize learning and training performance as declarative knowledge. Based on prior findings, I hypothesize that those trained with a team will have significantly higher level of knowledge than those trained individually. I also anticipate that affective states and declarative knowledge will be moderately correlated with one another.

Hypothesis 2: Participants who complete the training as part of a team will score significantly higher on the declarative knowledge assessment than participants who complete the training alone.

Training Processes

One purpose of the current study is to provide evidence on the processes that occur during the training experience (Adams et al., 2012). Some researchers have begun to explore possible mediating relationships evaluating performance in a game as the process through which affective reactions are experienced (Trepte & Reinecke, 2011). The current study will explore several possible indirect effects for the relationships between human interaction and two training outcomes (i.e., affective states and declarative knowledge).

Although there has been some support for the direct relationship between videogames and affective states such as motivation (Coller, Shernoff, & Strati, 2011; Derouin-Jessen, 2008; Liu, Cheng, & Huang, 2011), there is little empirical evidence to show indirect relationships between game characteristics and learning outcomes (Wilson et al., 2009). As discussed above, there is a gap in the research on the exploration of the mechanisms through which game-based learning is experienced. Other researchers have called for more research examining complex relationships between games and learning outcomes (Landers, 2014; Wilson et al., 2009).

Some researchers have described the gaming experience as unique to each individual. There is a consistent belief that a positive user experience in the game is associated with positive outcomes such as enjoyment and motivation (Calvillo-Gámez et al., 2010; Lankes, Bernhaupt, & Tscheligi, 2010; Lemay & Maheux-Lessard, 2015). Previous research has shown that interaction with others while playing a serious game can have a positive impact on different game outcomes (Bernhaupt, 2010). This is likely due in part to various mechanisms that trainees experience during the training process (Oksanen & Hämäläinen, 2014; Von Der Pütten et al., 2012). Many researchers have theorized mediating relationships but few have empirically tested these assertions. As part of this study I will test several potential processes to understand which mechanisms mediate the relationship between videogame characteristics and training outcomes.

Physical Engagement and Cognitive Engagement. Physical engagement and cognitive engagement refer to the extent to which individuals feel physically and cognitively involved in the videogame-based training activity. One of the most common perceptions about videogame-based learning is that games are engaging for individuals. Videogames are often promoted as effective learning tools through the belief that videogame-based learning generates player engagement, which leads to increased affective states (Coller & Scott, 2009; Mayo, 2007). Although there is some evidence that videogames can be engaging for players, there is little evidence to demonstrate that videogames, or the engaging experience they may create, leads to improved outcomes (Guillen-Nieto & Aleson-Carbonell, 2012), that is, there is a lack of evidence that engagement mediates the relationship between game play and training outcomes.

Evidence from previous studies demonstrates that human interaction will likely lead to an increase in both physical and cognitive engagement. I propose that this will in turn be associated with higher positive affective states. It is believed that the immersive environments of games are what leads them to be more physically and cognitively engaging (Annetta, 2008; Coller & Scott, 2009; Guillen-Nieto & Aleson-Carbonell, 2012; Mayo, 2007). I propose that the interaction with others during the game will enhance this feeling of physical and cognitive engagement because the human interaction will make the game feel more immersive, generating shared interest and experience (Jonker, van Riemsdijk, & Vermeulen, 2011). Previous research has shown that individuals in a multiplayer game were more active and involved than those in a single player game (Manninen, 2003). I further predict that these increased levels of physical and cognitive engagement will be associated with improved affective states. Other research has showed that a sense of physical presence in an activity was tied to emotional outcomes (Bernhaupt, 2010). Based on these results, I predict that human interaction will lead to increased levels of physical and cognitive engagement, which will in turn be associated with more positive affective states.

Hypothesis 3: Human interaction will lead to increased affective states through an increase in physical engagement.

Hypothesis 4: Human interaction will lead to increased affective states through an increase in cognitive engagement.

Perceived Value. Perceived value is the extent to which participants believe the training has instrumental value and would be useful in some way (Adomaityte, 2013). I believe that human interaction will lead to increased levels of perceived value, which will in turn improve the participant's affective states. The first part of this relationship, that human interaction will improve perceived learning, is supported by previous research that has shown teams were more likely to perceive information as valuable when made aware of its uses (Lingard, Whyte, Espin Ross Baker, Orser, & Doran, 2006). In general, the more individuals interact in training with others, the more they will have opportunities to learn the various uses of the information gained. I predict that these interactions with others will lead to increased perceptions on the value of the training content as those individuals will share their knowledge with their teammates during the activity in order to be successful. To support the second part of this predicted relationship, that perceived value will improve affective states, previous research has further shown that demonstrating the instrumental value of an activity to a participant will improve their affective impressions of that activity (Blair, 2011; Lepper & Gilovich, 1982; Shernoff, Csikszentmihalyi, Shneider, & Shernoff, 2003). Based on these findings I believe that human interaction will lead to increased levels of perceived value, which will in turn be associated with improved affective states.

Hypothesis 5: Human interaction will lead to increased affective states through an increase in perceived value for the activity.

Active Learning. Active learning is defined as the active attention that an individual gives to the videogame-based training activity, including effort exerted in thinking about or interacting with the components of the videogame (Bonwell & Eison, 1991; Brown, 2001; Prince, 2004). Evaluating active learning in a videogame-based learning context is valuable given that games are believed to inherently promote interest and draw attention (Adams et al., 2012; Gee, 2005), which are primary components of active learning. I argue that human interaction will promote active learning, which will in turn be associated with higher levels of declarative knowledge. Previous research has demonstrated that team interactions can lead to both active participation and improved performance (Goodman, Bradley, Paras, Williamson, & Bizzochi, 2006; Prensky, 2001; Vandercruysse, Vandewaetere & Clarebout, 2012; Vogel et al., 2006). However, I will further examine the relationships between these variables to determine if active learning is a mechanism through which human interaction improves learning.

The first part of this proposed indirect effect is that interactions within a team will lead to active learning. This has been supported by previous studies that have demonstrated that working with a team promotes active learning by holding one another accountable for paying attention to the tasks (Blair, 2011; Dillenbourg, 1999; Fandt, 1991). Team members are often motivated to encourage team participation, particularly in interdependent tasks when their own success is dependent on the actions and behaviors of their fellow teammates. The second part of this proposed indirect effect is that active learning will enhance learning. Previous research supports this assertion that active learning has a positive association with learning outcomes (Brown, 2001; Ford et al., 1998). I predict that participants who complete the activity as part of a team will demonstrate significantly higher levels of active learning, which will in turn be associated with higher levels of declarative knowledge than individual who completed the activity alone.

Hypothesis 6: Human interaction will lead to increased declarative knowledge through an increase in active learning.

Research Question – Exploration of Boundary Conditions

Some research has evaluated affective outcomes in serious games. However, the measures in these studies often vary and the findings are inconsistent. This makes it difficult to synthesize the findings because of the variability in studies (Mayer, 2011). For example, studies have found that participants in videogame conditions report higher levels of motivation than those in non-game conditions (e.g., Mautone, Spiker, & Karp, 2008). However, a meta-analysis from Wouters, Van Nimwegen, Van Oostendorp, and Van Der Spek (2013) found that serious games were not more motivating than non-game conditions. These inconsistencies highlight the need to be precise in defining training outcomes, and to understand how affective outcomes may differ from learning outcomes.

These variations in results are not the only concern. There is also some confusion between distinct types of outcomes. For example, there have been cases in which subjective reactions were described as indications of learning researchers used survey reaction questions to make implications for participant learning (e.g., Kobes et al., 2010). Based on the variations in how training outcomes are defined and measured for serious games, the likely result is outcome contamination (i.e., the behavioral measure is capturing another component not intended as part of the construct; Raykov & Marcoulides, 2011). To highlight this potential concern, it is important to evaluate and measure different outcomes within the same study to observe and better understand the differences between outcomes (e.g., cognitive outcomes versus affective outcomes).

To address this challenge, I will evaluate the extent to which predictors and mediating processes are similar for affective states and declarative knowledge. This will provide evidence regarding whether the inconsistencies in the research literature may be attributed to outcome contamination. If the paths are dissimilar, it will suggest that important variations exist between training outcomes and that these differences need to be carefully considered when measuring and making implications based on the training outcomes.

One of the challenges I identified in research is a lack of understanding the differences among training outcomes. Although I can make clear hypotheses for some of the indirect relationships, there is less theoretical justification for extending these hypotheses to other training outcomes. For example, it is clear how teams interacting together will hold each other accountable for active learning and that this active learning will likely be associated with improved learning outcomes. However, there is less theoretical justification to propose that active learning will be associated with improved affective states. Although attending to the content is likely to be associated with improvements in knowledge, it is not as practical that trainees will perceive this training to be valuable and enjoyable. Based on evidence that learning and affective states are not strongly related (Alliger et al., 1997) it is not reasonable to assume that the same relationships will exist between the mechanisms and different training outcomes.

There is little theoretical justification to predict that the indirect effects for affective states will generalize to indirect effect for declarative knowledge. I propose an exploratory research question to examine the relationships between training outcomes when using the same mechanism (i.e., how does the relationship between human interaction → active learning → declarative knowledge compare to the relationship between human interaction → active learning → affective states). The primary purpose of this research question is to explain if and how the relationships and effect sizes between mechanisms are different by training outcome.

Research Question: How will comparisons across training outcome influence the relationship of the indirect effects?

METHODS

A Priori Power Analysis

I collected a pilot sample of 513 participants as part of another study (147 individuals and 366 individuals within a team of three, with a total of 122 teams). This pilot sample included measures like those proposed in the current study. These similar measures were used in preliminary analyses of my hypotheses, using a single structure equation model in Mplus 7.4 (Muthén & Muthén, 1998-2012). I used the pilot data to conduct a power analysis for the hypotheses proposed in the current study. The results indicated that a sample of 100 participants would have over 90% power to detect the relationships in H1 & H2. With a sample of 200 participants I would have over 80% power to detect the relationships in H3-H5. However, with a sample of over 800 participants there would be just over 70% power to detect the relationship in H6. The results further indicated that with my current sample of participants (n = 385), I would have less than 50% power to detect the relationship in H6. Despite the low power for H6, I proceeded with my current sample because there was sufficient power to detect the remaining hypotheses. Further, it is important to consider that the pilot sample was from a single study with similar yet different measures of the constructs. Based on these differences the current findings may differ from what was initially found in the pilot sample.

Participants

Participants in the current study were 389 undergraduate students recruited from the psychology department and business school of a university in the western United States. Responses from four participants were removed from the data due to a technical error that prevented all necessary information from being collected from these. A total of 385 participants remained in the final sample. Participation in the study was voluntary in exchange for class credit. All participants were assigned to one of two conditions: 122 individuals completed the training alone, while 263 participants completed the training with a team.

Procedure

An overview of the experimental procedures is provided in Figure 2. Upon arrival, participants reviewed an online consent form and agreed to participate in the study. Participants then completed a self-paced computer-based training (CBT), which covered the skills and knowledge needed to play the videogame-based training. Details on the development of the CBT are provided below. Participants completed Survey 1, which was comprised of demographic questions, a measure for videogame experience, and a pretest assessing declarative knowledge retained from the CBT, see Appendix A. Survey 1 took participants on average 4 minutes (*SD* = 1 minutes) to complete. Participants were randomly assigned to a condition of either low human interaction (completing an activity in the videogame-based training alone) or high human interaction (completing that same activity with a team). All teams were composed of three individuals. When a participant was not available for the third slot, a trained research assistant acted as a confederate to ensure a full number of participants were in each team of three[2]. Although some have defined teams as groups involving two or more individuals (Kozlowski & Bell, 2013), others argue that there is a distinction between dyads and teams (Hinds & Weisband, 2003). In the pilot study, which used groups of both three and four persons, there appeared to be indications of potential differences in team performance based on team size. Thus, I chose to use only teams of exactly three to avoid having team size as a possible confounding variable in the study.

---

[2]I repeated the analyses using only data from the teams with three participants, and no confederates. These analyses did not change the outcome of the results, demonstrating that there was no significant change having used confederates for a portion of the data.

All participants spent 20 minutes in the videogame-based training to complete the activity. Their activity was recorded, capturing both audio and screen activity from the participants. All participants then completed Survey 2, see Appendix B, which took participants on average 34 minutes to complete (median = 21 minutes, *SD* = 15.5 minutes). It appeared that the mean response time was skewed by a small number of participants who needed additional time to complete the survey (e.g., students with English as their second language). Survey 2 consisted of a declarative knowledge posttest and the remaining measures described below. Participants were debriefed with a four-minute video describing the purpose of the study before being dismissed. Past studies have shown that including a debriefing promotes learning (Sawyer, Sierocka-Castaneda, Chan, Berg, Lustik, & Thompason, 2012). Sample screenshots of the debriefing video are provided in Appendix C. The value and application of using video debriefings is described in previous studies (e.g., Cheng, Eppich, Grant, Sherbino, Zendejas, & Cook, 2014). The entire experiment took 60 minutes to complete.

Computer-Based Training

The computer-based training program was developed using Captivate software. The design followed guidelines from Bell and Putman (1979). For development, subject matter experts (SMEs; i.e., those with tens of hours of experience playing the videogame-based training) met and discussed the major goals and objectives in the videogame-based training, described in further detail below. The SMEs discussed the knowledge participants would need to execute the goals of the game. This discussion produced a list of learning objectives, see Table 1. Content was generated to teach the learning objectives. The modules were designed to begin with an explanation of the learning objective and end with a summary of what was taught.

During the CBT participants interacted with the training content to progress forward in the training. Sample screenshots from the CBT are available in Appendix D.

Videogame-Based Training

The videogame-based training was a science fiction videogame named Quintet (Guida, n.d.), in which players assumed the role of crew members aboard a spaceship and carried out different missions. The game required players to learn the different roles available on the ship (i.e., Captain, Helm, Tactical, Engineer, Scientist) and to manage the tasks in each role to meet the different mission objectives (e.g., aid an ally ship in distress, escort a cargo ship transporting materials). When played as a team, the game connected team members virtually to communicate and work together. Participants scored points for meeting the mission objectives. Researchers followed a script when running the experiment to ensure consistency across iterations of the study. An example from this script is provided in Appendix E.

Declarative Knowledge Assessments

Declarative knowledge was measured during Survey 1 using a 14-item pretest and again during Survey 2 using a 26-item post-test. Both assessments were developed for this study. I followed research-based principles from Lord (1952) and Findley (1956) to construct the pretest and posttest assessments. A large pool of items was generated by a group of six SMEs. Each item was written to link to a specific learning sub-objective of the CBT (see Table 1). During this process, 123 multiple-choice items were written, each with four response options and only one correct answer. Several items were removed due to redundancy or concerns for item quality. The remaining 113 items were used in the first pilot test with a sample of 28 undergraduate student volunteers who completed the CBT and the test in exchange for class extra credit. Items were reviewed for difficulty, quality of distractors, and relationship to the overall exam score (i.e., point-biserial correlation). Based on these analyses, 48 items were retained and a second pilot test of 174 undergraduates was conducted. The same selection procedures from the first pilot test were used again in the second pilot test and a final list of 38 items was retained. The final set of items demonstrated a reasonable level of difficulty (.30-.90) and discrimination (point-biserial r >.25). The final set of items for both the pretest and the post-test had the same average difficulty (M = .71). At least two items from each learning sub-objective were included on each test. It was easier to compare results from the pretest to the posttest (i.e., score improvements) because of the similarity in content and difficulty between the two tests.

To explore possible threats to the internal validity of this design (Campbell & Riecken, 1968; Cook & Campbell, 1979; Cook, Campbell, & Perracchio, 1990), I used the Internal Referencing Strategy (IRS) proposed by Haccoun and Hamtiaux (1994). IRS is designed to improve the accuracy of estimating participant gains in knowledge. This is done by identifying the degree of change in knowledge that occurs in items that would not be expected to change due to the training intervention. These *content irrelevant items* were included on both assessments and were expected to not change, or change very little between the pretest and posttest. According to this strategy, any change in scores for *content irrelevant items* that occur between a pretest and posttest can indicate there is another explanation for score differences such as the practice effect (Haccoun & Hamtiaux, 1994).

Measures

All other measures were gathered from participants using self-report items in the two surveys. A summary of the measures with a sample item from each measure is provided in Table 2. All scales, unless otherwise specified, used a 5-point Likert type scale with 1 = Strongly

Disagree and 5 = Strongly Agree. I have not reported several additional measures included in the surveys that were gathered for future analyses and exploration and not for this study.

Demographic Measures. Participants reported their age, sex, and ethnicity.

Videogame Experience. Videogame experience is defined as the amount of experience an individual has with playing videogames. This was measured using a 14-item scale developed for a previous study (Sanchez, 2017). Four items in this scale were adapted from the Gaming Experience Measure from Singer and Knerr (2010). This measure had three subscales; *earnest video gaming*, *casual video gaming*, and *video game self-efficacy*. The data in this study from all 14 items produced a reliability of $\alpha = .94$.

Declarative Knowledge Pre-test. The declarative knowledge pre-test consisted of 14 multiple-choice questions. I calculated the average difficulty ($M = .63$).

Declarative Knowledge Post-test. The declarative knowledge post-test consisted of 26 multiple-choice questions. I calculated the average difficulty ($M = .67$). I did not calculate the alpha coefficient in either assessment for declarative knowledge because that alpha is relatively meaningless for this type of assessment. The purpose of this assessment was to create test items that sample broadly across the training content (within and across sections). Internal consistency is based on the assumption that the items come from a homogenous domain. Because the written items each tap a different objective, they are heterogeneous domains and internal consistency is not relevant (Cortina, 1993).

Psychological Meaning. Psychological meaning is defined as an individual's positive evaluation of their experience in the videogame-based training activity based on their value system (May et al., 2004). This six-item measure was taken from May et al. (2004). The reliability of psychological meaning in the current study was $\alpha = .88$.

Perceived Enjoyment. Perceived enjoyment is defined as the level of enjoyment a participant experienced from the activity in a videogame-based training activity (Adomaityte, 2013). This six-item scale was created using relevant items from enjoyment and satisfaction scales (Adomaityte, 2013; Calvillo-Gámez et al., 2015; Long, 2005; Ricci et al., 1996; Warr, Allan, & Birdi, 1999). The reliability of perceived enjoyment for the current study was $\alpha = .94$.

Motivation. Motivation is defined as the extent to which the participant felt interested and motivated by the activity such that they would voluntarily reengage in the activity. The six items used for this scale were adapted from the *attention* subscale in Keller's (2006) Instructional Materials Motivation Survey. Six items from the original scale were dropped for being not applicable to this study. The reliability of the items from the motivation scale used in this study was $\alpha = .84$.

Emotional Engagement. Emotional engagement is defined as the emotional connection the individual felt to the videogame-based training activity they were in (Rich et al., 2010). The six items from the scale in this study were taken from the *emotional engagement* subscale of the job engagement measure from Rich et al. (2010). The reliability of emotional engagement using the current study data was $\alpha = .93$.

Physical Engagement. Physical engagement is defined as the physical energies an individual felt in the videogame-based training activity used in this study (Rich et al., 2010). These six items were taken from the *physical engagement* subscale of the job engagement measure from Rich et al. (2010). The reliability of physical engagement in the current study was $\alpha = .90$.

Cognitive Engagement. Cognitive Engagement is defined as the cognitive energies an individual felt in the videogame-based training activity in this study (Rich et al., 2010). These six items were taken from the *cognitive engagement* subscale of the job engagement measure from Rich et al. (2010). The reliability of cognitive engagement in this study was α = .96.

Perceived Value. Perceived value is defined as the extent to which the participant felt the activities in the videogame-based training were useful and have instrumental value (Adomaityte, 2013). This six-item scale was adapted from the *relevance* subscale in Keller's (2006) Instructional Materials Motivation Survey. Three items from the original scale were dropped for not being applicable to the current study. The reliability of the perceived value scale in this study was α = .78.

Active Learning. Active learning is defined as the active attention and involvement given to the learning process in the videogame-based training activity (Walker & Fraser, 2005). This was measured using the *active learning* subscale from the Distance Education Learning Environments Survey (DELES) (Walker & Fraser, 2005). The reliability of active learning in this study was α = .83.

RESULTS

Data were collected from 389 participants. Four participants were not able to complete the study due to a technical error that caused the game to crash during the experiment. These participants were dismissed with credit and their data was removed from the sample because they completed neither the game nor survey 2. The final sample consisted of 385 participants with complete data. In the final sample, 51% were female, 73% were Caucasian with an average age of 19 years old (*SD* = 1.7). A summary of study means, standard deviations, and correlations is provided in Table 3.

Assumption of Independence

Participants in the team condition interacted with one another, which violates the assumption of independent observations. To assess the impact of this non-independence, I examined the intra-class correlations for each measure, see Table 4. The ICCs range from .02 to .19, indicating inconsistency in the variation of responses from within a given team. Thus, I considered clustering the data by teams. I tested the final structural equation model both with and without clustering individuals in teams and the outcomes did not change (i.e., results and fit were practically the same), see Table 5. For clarity, I report the results of the analyses without clustering.

Measurement Model

My approach to verifying the fit of my model began with building and testing the measurement model. To do this I tested a model in which each of the items loaded onto the factor it was intended to measure and then evaluated the overall model fit. I conducted my analyses in Mplus 7.4 (Muthén & Muthén, 2012) using maximum likelihood estimation to estimate the model fit indices. I applied overall model fit criteria recommended by Hu and Bentler (1998): Root Mean Square Error of Approximation (RMSEA), with a value < .08 indicating adequate fit, and Standardized Root Mean Square Residual (SRMR), with a value < .08 indicating good fit. I also considered the Comparative Fit Index (CFI) and the Tucker–Lewis Index (TLI), both indicating acceptable fit with a value > .90. I reviewed the Chi-Square test for model fit with a non-significant estimate indicating good fit. The first tested measurement model did not produce acceptable fit. A series of reviews and revisions were made retesting the model fit with each revision. Table 5 provides a summary of the fit statistics for all tested models.

Revisions to the Measurement Model. To address concerns of poor fit, I made revisions to the measurement model. First, I examined the standardized factor loadings to identify items that did not load well onto the scale factor (i.e., $\lambda < .30$). This revealed two items that did not load well onto their respective scales, see Table 6. The item PV5 did not load well onto the perceived value scale, $\lambda = .09$. I removed this item first because it had the lowest factor loading. This item may have loaded poorly onto the factor because it was the only negatively worded item in the perceived value measure (DeVellis, 2012; Raykov & Marcoulides, 2011). Previous studies have shown that negatively worded items tend to load onto a single factor different from positively worded items (Caught, Shadur, & Rodwell, 2000). Some researchers have attributed this to "careless" survey responses (Schmitt & Stuits, 1985). After removing the item, the model fit improved significantly but did not produce adequate fit, see Table 5.

I next removed item MOT4 which did not load well onto the motivation scale, $\lambda = .29$. This item was one of three negatively worded items in the measure. This item likely did not load well with the other items because it asked participants about the repetitive nature of the task whereas all remaining items in the motivation scale asked about the interest level of the activity.

The content of item MOT4 is clearly different from the content in the remaining items in the measure. After removing MOT4, the model fit improved significantly but still did not produce adequate fit, see Table 5.

Examining item discrepancies (i.e., residual correlations) can also help identify model misfit (McDonald & Ho, 2002). Items were reviewed to identify large discrepancies (i.e., > .15) with other items in the model (McDonald, 1999). These doublet items can indicate a high level of redundancy meaning they overlap substantially with other items (either on the same or in different scales). Table 7 provides a summary of the items that were removed due to high discrepancies.

I first reviewed items with the *greatest* number of large discrepancies and worked down from there, removing items and retesting the model fit. I identified seven items with more than three residual correlations beyond the .15 cutoff. I re-evaluated model fit after each of these items were removed. My model produced adequate fit after the first item was removed and showed significant improvement (via the chi-square difference test) after each of the additional items was removed. I kept the model with all seven items removed to be consistent in the criteria applied to dropping items from the model and because this model demonstrated satisfactory fit, see Table 5.

As expected, the factor scores for the four affective states all loaded onto a second-order factor; psychological meaning ($\lambda = .86$), perceived enjoyment ($\lambda = .92$), motivation ($\lambda = .95$) and emotional engagement ($\lambda = .92$). Each of these constructs similarly indicates a positive affective state regarding the participants' experience in the game-based training. However, a model collapsing all four scales into a single common factor for affective states did not fit well, indicating that (as expected) they are related but distinguishable constructs (see Table 5). As intended, the three videogame experience subscales loaded onto a higher-order factor (earnest video gaming, $\lambda = .95$; casual video gaming, $\lambda = .92$; video game self-efficacy, $\lambda = .94$). All remaining items loaded onto a single factor for each scale, see Table 6.

I added the remaining study variables (i.e., human interaction and declarative knowledge) allowing these variables to correlate with the other measures in the model. This final model demonstrated good fit ($\chi^2(1192) = 2461.44$, $p < .001$, $RMSEA = .05$ [.05, .06], $p = .07$; $CFI = .92$; $TLI = .92$; $SRMR = .05$). Factor loadings for all items in the final measurement model are reported in Table 6.

Hypothesis Testing

After demonstrating satisfactory fit with the revised measurement model, I tested Hypotheses 1-6 in one structural equation model (SEM), controlling for videogame experience. The results are shown in Figure 3. All variables were scored on a continuous scale and distributions were normal. As before, I conducted these analyses in Mplus 7.4 using maximum likelihood estimation to estimate the model fit indices.

Overall Model Fit. The first structural equation model did not have adequate fit (Table 5) therefore I considered possible sources of misfit. I reconsidered the relationships that may occur between variables in the game cycle process. It is likely that these variables relate to one another, as they are all part of an overall *training experience*. This is consistent with previous research that has referenced *gameplay* as a consistent state, generally positive or negative beliefs and attitudes that are interdependent, that the player experiences during the game (Ermi & Mäyrä, 2005; Jabbar & Felicia, 2015; Ke, 2008). Previous research has shown there can be large variations between how students experience a learning environment but consistency within an individual’s experience (Ermi & Mäyrä, 2005; Frenzel, Pekrun & Goetz, 2007). It is likely that the individual's experience in one of the measured processes (e.g., perceived value) is related with their experience in one of the other processes (e.g., cognitive engagement). Based on this conceptualization of the game cycle, it is likely that these process variables are related to one another. I modified the model to include factor-level correlations between the process variables during the game cycle (i.e., physical engagement, cognitive engagement, perceived value, and active learning). Table 8 provides a summary of the correlations between all the latent factors in the final model. This revision to the model improved the overall model fit. This revised structural equation model demonstrated good fit, see Figure 4. The overall fit indices were within acceptable ranges ($\chi^2(1200) = 2556.71$, $p < .001$, *RMSEA* = .05 [.05, .06], $p = .009$; *CFI* = .92; *TLI* = .91; *SRMR* = .11). All tested relationships in the final model are presented in Table 9.

Direct Effects. Hypothesis 1 stated that individuals who completed the training with a team would experience significantly higher affective states than individuals who played alone. Results supported this hypothesis. The total direct effect for human interaction influencing affective states was ($b = .17$, *SE* = .07, $p = .01$, *95% CIs* = .04, .30, $\beta = .11$), see Figure 5. The direct effects were also evaluated in the full model, which also included the hypothesized mediation paths. The remaining direct effect is the effect after accounting for the other indirect effects in the model. The remaining direct effect was also significant, as reported in Table 9, indicating that the effect of human interaction on affective states was only partially mediated by the mediating variables proposed in the model.

Prior to testing Hypothesis 2, I reviewed the declarative knowledge assessment given as the pretest and posttest. Using the Internal Referencing Strategy, I evaluated the differences between the pretest and posttest for both the content relevant items and the content irrelevant items (Haccoun & Hamtiaux, 1994). I found no significant difference ($F_{1, 769} = .12$, $p = .73$, $R^2 <$

.001) for the *content irrelevant items* from the pretest (*M* = .22, *SD* = .22) to the posttest (*M* = .22, *SD* = .20). Responses at both time points reflected scores comparable to an individual guessing on each item (close to 25% correct). Despite this, playing in the videogame-based training did not improve declarative knowledge outcomes beyond the level of knowledge demonstrated by participants after completing the CBT and the pretest. There was no significant difference ($F_{1, 769}$ = 1.59, *p* = .21, R^2 < .002) from the pretest scores (*M* = .63, *SD* = .21) to the posttest scores (*M* = .62, *SD* = .17) on content relevant items, meaning no significant improvement in declarative knowledge occurred between the pretest and the posttest assessments.

Although for the total sample there was no significant change in declarative knowledge, I examined my second hypothesis to see if there were significant differences in posttest scores for individuals based on human interaction. Hypothesis 2 stated that individuals who played with a team would demonstrate significantly greater levels of declarative knowledge than individuals who played alone. The results did not support this hypothesis; the direct effect between human interaction and declarative knowledge was significant but in the wrong direction (*b* = -.01, *SE* = .01, *p* = .02, *95% CIs* = -.06, .-006, β = -.03). The remaining direct effect was not significant as shown in Table 9. This is the residual direct effect after accounting for the indirect effects in the model. This is interpreted to mean that human interaction has a small but significant negative direct effect on declarative knowledge until accounting for the indirect effects in the model. Once accounting for the mediating effects of the process variables, there is no significant direct effect from human interaction to declarative knowledge. The post-test declarative knowledge scores for individuals who played with a team (*M* = .67, *SD* = .17) were not statistically different from the scores for individuals who played alone (*M* = .67, *SD* = .20). The control variable videogame experience was associated with significantly higher levels of declarative knowledge ($b = .07$, $SE = .01$, $p < .001$, *95% CIs* = .05, .09, $\beta = .43$), after accounting for the indirect effects in the model. Videogame experience has a moderate effect on declarative knowledge outcomes after accounting for other indirect effects. Individuals who reported more videogame experience had significantly higher declarative knowledge scores.

I also predicted that there would be a moderate correlation between the two outcome variables, declarative knowledge and affective states. The results did not support this, affective states and declarative knowledge were not significantly correlated, $r = -.001$, $p = .78$. Although there was theoretical justification to suggest that these outcomes were related, this was not the case in the current sample after accounting for the other relationships in the model.

Indirect Effects. Hypotheses 3 – 5 predicted indirect effects for the relationship between human interaction and affective states. I assessed the asymmetrical confidence intervals (ACIs) to ensure they did not contain zero as this is considered an indication of statistical significance for examining indirect effects. This method avoids the issue of lost statistical power when examining non-normal distributions as a product of the two regression slopes (Davey, 2009). I used a 95% bias-corrected bootstrapped confidence intervals based on 1,000 bootstrapped samples. I also examined *Pm* as an indicator of effect size based on research from Alwin and Hauser (1975) using $P_m = \frac{ab}{c}$, a ratio of indirect to total effect. The bias-corrected bootstrapped confidence intervals demonstrated that one of the predicted indirect effects was statistically significant (human interaction → perceived value → affective states, $b = .14$, $SE = .05$, *95% CIs* = .05, .23, $Pm = .37$, $p < .001$). The other proposed indirect effects were not significant (i.e., human interaction → physical engagement → affective states, and human interaction → cognitive engagement → affective states), see Table 9. The standardized estimates are provided for the structural equation model in Figure 5. These findings demonstrate that physical engagement and cognitive engagement do not mediate the relationship between the game characteristic human interaction and affective states. It is important to note that both physical engagement and cognitive engagement were significantly related to affective states. However, human interaction was not significantly related to physical or cognitive engagement, which is why no mediating relationship was found. However, human interaction does lead to improved perceptions of value, which in turn significantly improved the positive affective state of participants.

Hypothesis 7 stated that active learning would mediate the relationship between human interaction and declarative knowledge. The results did not support the hypothesized effect (human interaction → active learning → declarative knowledge, $b = -.01$, $SE = .01$, *95% CIs* = -.02, -.002, $Pm = 2.00$, $p = .72$). This effect ($Pm = 2.00$) is not significant because the direct effect is close to 0, $c = -.02$, making the valuable unreliable (see Table 11). Although there appears to be some evidence that playing the game alone will lead to higher active learning, which in turn will have a positive association with declarative knowledge, this effect appears to be negligible.

Research Question. I explored the research question using an alternative SEM model, controlling for videogame experience as before. This model included two direct effects, one between human interaction and affective states and the other between human interaction and declarative knowledge, see Figure 1. The alterations to this model included an exploration of possible indirect effects (i.e., human interaction → physical engagement → declarative knowledge, human interaction → cognitive engagement → declarative knowledge, human interaction → perceived value → declarative knowledge). Similarly, this model includes the evaluation of the indirect effect (human interaction → active learning → affective states). I compared the same indirect effects from the hypothesized model across different training outcomes using $P_m = \frac{ab}{c}$, a ratio of indirect to total effect (Alwin & Hauser, 1975). Larger values indicate a larger effect size. Due to limited studies applying the *Pm* value, standardized values indicating small, medium or large effects have not been established. The first model for the research question produced adequate fit ($\chi^2(1200) = 2808.99$, $p < .001$, $RMSEA = .06$ [.06, .06], $p < .001$; $CFI = .90$; $TLI = .89$; $SRMR = .17$), thus, no evaluations or revisions were made to the model. The bias-corrected bootstrapped confidence intervals demonstrated that one indirect effect was statistically significant (human interaction → active learning → declarative knowledge, $b = -.10$, $SE = .03$, *95% CIs* = -.16, -.04, $Pm = -.25$, $p = .01$), as shown in Table 10. The other proposed indirect effects were not significant (i.e., human interaction → physical engagement → declarative knowledge, human interaction → cognitive engagement → declarative knowledge, human interaction → perceived value → declarative knowledge). The standardized estimates are provided for the structural equation model in Figure 6. The results from this model in comparison to the previous model demonstrate that gameplay processes mediate the relationships between game characteristics and training outcomes differently. In the hypothesized model, there appears to be a significant indirect effect (i.e., human interaction → perceived value → affective states, $Pm = .37$, $p < .001$). Also, in the research question model, there appears to be a significant indirect effect (i.e., human interaction → active learning → affective states, $Pm = -.25$, $p < .01$). Together these results indicate that both perceived value and active learning mediate the relationship between human interaction and affective states, with more human interaction leading to more positive affective states through perceived value and less human interaction leading to more positive affective states through active learning.

These findings demonstrate that the processes through which game characteristics impact affective states are different from the processes through which game characteristics impacts declarative knowledge. When considering the game characteristic human interaction, affective states were impacted through the mechanisms of both perceived value and active learning. However, declarative knowledge was only impacted through active learning, which was a small effect. This provides valuable evidence for future researchers in supporting the assertion that thoughtful design needs to be considered regarding the intended outcomes of the training as the mechanism in the game processes will affect training outcomes differently.

DISCUSSION

The purpose of this paper was to explore the impact of human interaction in a videogame-based training tool on participant outcomes. I further sought to understand the possible mechanisms that mediate this relationship (i.e., between game characteristics and training outcomes). To accomplish this purpose, I compared individuals who played the training videogame alone with those who played the training videogame with a team. I gathered information about each participant's user experience (e.g., perceived value of the training) and measured the outcomes of the training (affective states and declarative knowledge). The primary findings demonstrated that after accounting for the indirect effects in the model, human interaction had a significant direct effect on affective states but not on declarative knowledge. Participants who trained as part of a team reported significantly greater levels of positive affective states (indicated by psychological meaning, perceived enjoyment, motivation and emotional engagement) but no significant difference in declarative knowledge. Further analyses showed that game-based training with a team impacted the affective states of players through the mechanisms of perceived value and active learning. Also, active learning mediated the relationship between human interaction and declarative knowledge. Below, I discuss these findings, their implications, and address limitations and recommendations for future research.

This study was developed in response to the mounting number of contradictory findings regarding the effectiveness of videogames as a learning tool. While some studies show that videogames are effective tools for promoting learning (e.g., Blunt, 2009), other studies have failed to find such effects (O'Neil & Perez, 2008). My review of the research literature revealed current challenges in how research is being conducted (e.g., vague explanations of game design features that may impact learning outcomes). These challenges may be addressed by changing the focus of the research question to understand the context in which games are effective for learning, and by increasing the scientific rigor of studies in the field.

Previous researchers have highlighted the need for quality game-based research, stating that game developers need better guidance on how to design a game in a manner that will facilitate learning outcomes (Adams et al., 2012; O'Neil & Perez, 2008). Recent research has called for studies to examine the variables that lead to effective game-based learning (Wong et al., 2007). Some researchers have stated that there is a need for specific studies that isolate game elements to determine their effects (Adams, 2013). However, researchers have been calling for these studies since the mid-1980s (e.g., Lepper, 1985), leading some researchers to criticize the lack of progress that has been made in the field over the last three decades (Butler et al., 2014). Given the amount of time and money invested in producing games for learning, it is apparent that a resolution to this contradictory research is needed.

Developing quality game-based research is further important due to recent practices in how games are being applied. Some organizations have begun to use games for high stakes purposes such as a step in a multi-hurdle selection process (Amad, 2015; Morgan, 2013). This is concerning because there is inconclusive evidence to justify the use of games for this purpose (i.e., no known published studies validating games as an effective selection assessment method for predicting job performance). Although there seems to be considerable potential in this application for game-based tools, research must catch up with application to provide evidence on the usefulness of games as high-stakes selection tools. The negative consequences of applying games without sufficient research backing were evident in the recent outcome of the lawsuit with Lumosity. Certain parties sued Lumosity for making broad reaching claims regarding the cognitive benefits of their "brain training" games (Federal Trade Commission, 2016; Span, 2016). Lumosity settled this case, with substantial financial penalty and changes to their business processes. Based on these deleterious outcomes, it is important for researchers and practitioners to be vigilant about the rigor of game research and the recommendations that are made in applying research results. Overstating claims can have critical consequences for organizations. The research to justify the use of games for these means is far from conclusive. It is apparent that organizations need to be cautious when applying videogames for purposes other than what there is clear and rigorous scientific evidence to support.

It is important for game-based research to address these challenges by conducting high quality studies that examine how game characteristics impact learning. The current study provided an expanded version of Garris et al.'s (2002) game-based learning model and addressed current challenges to the research literature. Specifically, I isolated one game characteristic (i.e., human interaction) that has been proposed to influence training outcomes. I further tested several potential variables (i.e., physical engagement, cognitive engagement, perceived value, and active learning) as possible mechanisms through which game-based learning is influencing training outcomes. Finally, I measured two training outcomes (i.e., affective states and declarative knowledge) to observe differences between these outcomes. This is a valuable contribution to the research literature given the ongoing debate regarding the effectiveness of using game-based learning as a tool to train individuals. This model followed a new approach to the research question, exploring the context in which game-based learning leads to improved training outcomes.

The current study indicated that game characteristics may influence training outcomes. Results showed that human interaction significantly impacted affective states. Those who played the game with a team experienced significantly more positive affective states than individuals who played the game alone. This is consistent with previous research that has shown that many individuals prefer to work with a group rather than alone (Gardner & Korth, 1998) and that working with a team can lead to positive affective outcomes (Taut, 2007). This means that individuals will likely experience more positive affective outcomes from the training if they are allowed to interact with other individuals as they play the game. Thus, designing or selecting a game with a multiplayer option may be beneficial in applications in which the goal is to improve a player's affective state. It is not clear whether other game characteristics can improve affective states; this warrants further exploration. I also found that human interaction did not have a direct effect on declarative knowledge after accounting for the indirect effects in the model. When examining the posttest scores there was no significant difference between groups (i.e., if they played alone or with a team) on declarative knowledge.

Implications and Future Research

It is beneficial to re-examine the primary outcomes of the current study to understand major implications for future research. These results have valuable implications for future research in a number of ways, and indicate in which additional research is needed.

Impact of Game Characteristics. The results of this study demonstrate that game characteristics can have an impact on training outcomes. Because only one game characteristic (i.e., human interaction) was explored in the current study, more research is needed to clarify the effects of other game characteristics and their impact on other training outcomes. However, it is still clear that game characteristics can impact training outcomes and should be considered from a design perspective for future game-based learning tools. Examining other game characteristics could shed light on the complex relationships between game characteristics, mediating variables, and training outcomes (e.g., Vogel et al., 2006). For example, the game characteristic *assessment* (characterized by the extent an individual receives feedback during the training) could have a greater impact on mediating factors such as active learning and cognitive engagement, which could then lead to a greater effect on learning outcomes. By only examining only one game characteristic (i.e., human interaction), I may have limited my ability to find the relationships I hypothesized. Thus, another research opportunity is for other studies to systematically explore multiple game characteristics across studies while tracking the characteristics and combinations of characteristics that produce the strongest effect. This would benefit researchers and practitioners by clarifying which characteristics they may choose to include in the design of a game and what combinations would best match the unique goals of the training.

Since game characteristics can have an impact on training outcomes, it is concerning when researcher do not disclose enough information about a game so that the game characteristics can be inferred. A primary benefit of disclosing detailed information about a game is that the effects of the game characteristics can be identified. Without considering the effects of the game characteristics, the findings from the study may be misattributed to other factors (e.g., condition assignment). Thus, future researchers should disclose information about the game in enough detail so that the game characteristics can be identified.

Another beneficial practice would be describing the game characteristics using a common framework or taxonomy. Using a standardized framework to describe game characteristics would help future researchers understand and synthesize the results of game-based learning across studies. This would allow more accurate comparisons and interpretations to be made across studies and would provide a clearer picture of the impact of game characteristics within a larger context of game-based learning. However, research synthesis and cumulative research (building on what others have found) is limited by poor descriptions or mischaracterizations of game characteristics by researchers or game manufacturers. Research would benefit from the disclosure of game characteristics using a standard framework, such as presenting a profile of game characteristics. An example profile of the game characteristics for the game used in the current study is provided in Table 12.

One primary benefit of expanding the literature on game characteristics is that research findings can lead to intentional game design decisions such as using specific game characteristics or combinations of characteristics to meet training objectives. The goal of understanding game characteristics is that research findings can inform the future design and development of videogames used for learning. As scientific findings provide clear evidence about the game characteristics that lead to specific training outcomes, games can be customized by developers or chosen by clients specifically for the intended training outcomes. An example from the present study is that a game designed to include a multiplayer option would lend itself to a training intention of improved affective states. Previous research has demonstrated that specific training design can influence training outcomes (Broad & Newstrom, 1992) and it is advantageous to consider research findings during the development phase to inform the decisions made regarding a training program.

Human Interaction. In the current study, human interaction led to decreased levels of active learning. This is potentially due to the additional attention that team members paid to interacting with others when collaborating and coordinating their efforts. In this study, participants interacted only once for 20 minutes. It is possible that initial interaction detracts from the attention paid to the learning process but that later interactions may be different, once teams have had sufficient opportunity to develop collaborative relationships. This may indicate that teams need additional time in game-based training to interact and be effective, consistent with previous research that has shown that the benefits of teams may take time and multiple interactions to emerge (Brodbeck & Greitemeyer, 2000a). Practitioners could design a training with a multiplayer component to allow extra time for team interaction, to emphasize the importance of active learning, or to allow the team ample time to prepare before the game (e.g., preparing a strategy of working together).

Other social mechanisms could contribute to the differences observed between individuals and teams, such as *diffusion of responsibility* (Darley & Latané, 1968). That is, participants playing with a team may not feel that they are responsible for learning all the components of the game if they have others who they can depend on to help them be successful during the game. Although participants were warned they would complete a test at the end of the study, they may still have experienced this diffusion when working with other team members. It would be valuable for future researchers to more closely observe the social interactions and experiences that teams are having in the game and to explore how these interactions impact learning outcomes. Awareness of the processes that occur during the game cycle and building in new features to the game may change the training process and improve training outcomes.

Mechanisms in the Game Cycle. Based on Garris et al.'s (2002) model of game-based learning, inputs are believed to impact training outcomes via various training processes in the game cycle. In previous research, these game cycle mechanisms have been considered a black box of training processes (Arnab et al., 2015). Other researchers have identified the need to examine variables that may exist in this black box of game cycle processes and that mediate or moderate the relationships between game characteristics and training outcomes (Landers, 2015).

However, there is little empirical research examining these mechanisms. Therefore, I explored several mechanisms expected to exist in this game cycle.

I found that perceived value mediated the relationship between human interaction and affective states. This is consistent with previous research, which has shown that understanding the value of a training has an important association with participants' experience and the outcomes of the training (Lingard et al., 2006). In one previous study, researchers found that perceived value of the learning process lead to increased levels of participant enjoyment (Ainley & Ainley, 2011). I also found that active learning is a mechanism through which human interaction impedes both affective states and declarative knowledge. Thus, my study was one of the first to identify specific game processes that are influenced by game characteristics.

Two of the tested process variables (perceived value and active learning) mediated the relationships between human interaction and the training outcomes, Two other process variables (physical engagement and cognitive engagement) were examined and there was no evidence that either mediated the relationship between human interaction and the two training outcomes. However, both physical engagement and cognitive engagement were significantly related to affective states. Also, cognitive engagement was significantly related to the outcome of declarative knowledge. This means that although these variables did not mediate the relationship between human interaction and the outcome measures, it is possible that they could be important mechanisms in the game cycle. It is possible that these associations indicate complex relationships between other game characteristics and training outcomes. This suggests that perhaps these variables mediate the relationship between a different game characteristic and the outcome variable. Thus, my study was one of the first to identify specific game processes that are influenced by game characteristics.

Although further research is needed to explore the mechanisms at play in the game cycle, it is clear there are complex experiences occurring for players. To understand the game cycle, other researchers have discussed this process and developed a definition for game cycle, which most researchers seem to agree is a subjective user experience designed to elicit positive reactions from players (Calvillo-Gámez et al., 2010; McCarthy & Wright 2004; Popper, 1994). One concern with this definition is that the game cycle is the individual's unique user experience, based on how he or she constructs and interprets the game (Ermi & Mäyrä, 2005). This poses a unique challenge for researchers who are attempting to make broad implications that are consistent across individuals. Although this user experience is complex, individually specific, and multi-dimensional, some researchers have developed methods for making comparisons across different user experiences (Calvillo-Gámez et al., 2010), allowing for other models of the user experience to emerge. For example, the sensory, challenge-based, and imaginative immersion model (SCI-model) is based on the immersive components of games and describes the user experience as a complex interaction of various processes that lead to the experience of game immersion. The concepts of this model are similar to those of other models. For example, the User Experience (UX) concept focuses on two parts, the interaction between the individual and the game along with the outcomes (i.e., fulfillment of goals) of the game (Hassenzahl, 2003). The player evaluates their experience during the game, working to understand how they feel in their experience and this influences their interactions during the game cycle (Calvillo-Gámez et al., 2010; Kaye, 2007; Preece et al., 2002).

Despite the complications in studying the game cycle, it is important to evaluate the user experience to understand the role it plays in affecting training outcomes (Calvillo-Gámez et al., 2010). The cycle of game elements leading to a unique user experience which influences training outcomes has been described as a feedback loop that can be designed to lead to beneficial states such as happiness, engagement, flow, and satisfaction (Calvillo-Gámez et al., 2010; Csikszentmihalyi, 1990; Dewey, 1938). Thus, it is important to understand the processes from a design perspective to help optimize the user experience and the effectiveness of the training towards the intended outcomes. It would be valuable for game designers to intentionally build in opportunities to enhance the game cycle according to the intended goals of the training.

Re-evaluating the Game-Based Learning Debate. Another purpose of this study was to provide empirical evidence that addressed the challenges previously outlined at the beginning of this paper. This study was designed to address these challenges and to lay the groundwork for future research on game-based training. Based on the results of this study, there is evidence that process variables in the game cycle are associated with training outcomes in different ways. When comparing results across training outcomes, different variables mediated the relationships between game characteristics and training outcomes (i.e., perceived value mediated the relationship with affective states but not declarative knowledge). Also, the strength of these effects differed by training outcome (i.e., active learning was a stronger mediator for affective states than declarative knowledge).

These differences are particularly concerning because there is a longstanding precedent of conflating outcome measures for learning (Clark, 2001). Previous researchers have used vague definitions for learning or overstated results implying subjective or reaction measures as objective measures of learning (Rosen & Petty, 1992). Other researchers have commented on similar concerns that researchers do not have the evidence to support the claims made in the arena of game-based learning (e.g., Mayer, 2011). Based on this it is possible that some of the mixed results on the effectiveness of game-based learning can be attributed to erroneous measures and descriptions of training outcomes. To address this concern, researchers need to be transparent about how they are defining and measuring their training outcomes. At times, the outcome measure has been clearly mismatched with the type of training being given. For example, in one study it was stated that the training game was intended to develop the player's skill. However, no outcome measure for procedural skills was given. Practice time was measured in addition to subjective responses to a survey questionnaire (Zwikael & Gonen, 2007). Conclusions about participants' skills were based on participant's responses to the survey. In instances such as this, researchers are overstating the results of their study (and not on objective skill measures) and thus making claims that are not substantiated by the research.

Researchers have misstated the results of their studies or are vague about the implications of their outcomes. This tendency makes it challenging to synthesize existing literature and make definitive statements about the effectiveness of games for learning. Erroneous and overstated research findings may be contributing to the contradictions in the research literature regarding the effectiveness of game-based learning tools. The current study begins to address this by: 1) clearly defining specific learning outcomes; 2) demonstrating connections between game characteristics and learning outcomes along with 3) clarifying that multiple training outcomes have differential relationships with the game processes (i.e., human interaction benefits affective states but has a negative effect on declarative knowledge). These results demonstrate the importance of clearly defining and distinguishing training outcomes, as they are impacted differently by game characteristics. Understanding the importance of identifying and distinguishing training outcomes in future studies, and drawing accurate and precise conclusions based on the research findings may help resolve the existing contradictions in the research literature.

Studying Business Relevant Games. A critical next step for future research on game-based learning for workplace outcomes is for researchers to study the actual games that organizations are using to develop employee knowledge and skills. Empirical studies of games used in business environments rarely examine the games used by training practitioners. There may be a number of reasons for this. Understandably, concerns about sharing proprietary information or programs may limit a company's willingness to share their training products for scientific exploration. However, it would benefit both practitioners and researchers to generate evidence on the effectiveness of game-based learning using the actual games used in organizations. At present, many researchers use games that are readily available (e.g., commercial games, entertainment-based games) as proxy measures for the training games used by organizations. There is a dearth of research studying the game-based tools that organizations are using today. In this study and other similar studies, the generalizability of the results is limited by the tools being used in the study. The present study used a game focused on building teams of individuals to coordinate their efforts on a ship in order to meet specific mission objectives. An immediate concern of this game is that participants will rarely see any real-world value in developing their skills in this game. Not only does the game lack face validity for real world application, but the skills the participants are developing (e.g., communication and knowledge sharing) are secondary skills and not the primary focus of the game. Future studies could address this concern by studying games currently being used within workplace settings. This would allow more clear connections to be made between the research and the application of the results.

The Role of Videogame Experience. Another consideration is the impact of videogame experience. For this study, I used videogame experience as a statistical control for my measures.

This is a practice that has been used repeatedly in the research literature (Bauer, Brusso, & Orvis, 2012). The belief is that previous experience with videogames can confound the results of a study, specifically that individuals with more videogame experience may learn better in videogame environments for reasons other than the game characteristic under study. The benefits of videogame experience could include preference for the instructional tool, greater enjoyment with the gameplay, or more familiarity with game controls. Despite the prevalence of measuring videogame experience as a control measure, there is little explanation in the literature as to whether this measure has an impact and the extent of that impact. It would be valuable for future research to explore a standardized definition of videogame experience, and to validate a measure of videogame experience. Developing and using a valid measure of videogame experience would enable additional research to explore the implications that videogame experience has for training outcomes.

Exploring the link between experience and performance is not new for psychologists. In the 70's and 80's, researchers explored the link between job experience and job performance. Researchers believed there was a link between experience and performance and frequently measured work experience as a control measure. However, a series of meta-analyses demonstrated that there was a small relationship, but that the definition and measure of experience was inconsistent and had a large impact on how the results should be interpreted (Quińones, Ford, & Teachout, 1995). This is not to imply that videogame experience does not have an impact on game-based learning. On the contrary, I believe there are strong theoretical arguments for why researchers should be measuring and controlling for videogame experience in game based learning. However, it is clear there needs to be a consistent method for measuring game experience and researchers need to understand and account for the impact that videogame experience has on game-based learning outcomes.

Links to Job Performance. Finally, future research should establish a clear link between game-based training and job performance. My study demonstrates that human interaction impacts affective states and declarative knowledge, but stopped short of showing impacts on job performance. Prior research has demonstrated a positive correlation between declarative knowledge and transfer of training, but not as clearly between affective reactions and transfer of training (Blume, Ford, Baldwin, & Huang, 2010). There is some theoretical support to imply that some positive affective states experienced in the training (e.g., motivation) can benefit the acquisition of declarative knowledge, which may increase the likelihood that transfer of training will occur (Cheng & Hampson, 2008). However, there are many factors that affect transfer, such as workplace support and opportunity to practice and apply the skill (Cheng & Hampson, 2008). Considering the number of factors that can influence this outcome, it is important to gather direct transfer measures rather than relying on the relationship between transfer of training and other available measures. Thus, there is little that can be said regarding the generalizability of the current results to direct work-related outcomes such as job performance. However, future researchers could design their studies to demonstrate these connections. Using a real-world sample of workers would strengthen this area of research and provide stronger evidence about the effectiveness of game-based learning for transfer of training and job performance. There are several benefits that would come with using a real-world sample. For example, student participants may be less motivated to pay attention and do their best in a lab experiment because there is little incentive for them to put forth effort in the study. In addition to this, workers have previous skills and experiences that could be applicable to the experiment and could influence the results of the study in a meaningful way. Thus, these differences between a subject pool of students and a sample of workers could impact the results and it is important to recognize that future studies using real world samples may benefit from these differences and provide a more accurate indication of game-based learning in an organizational setting.

In a design that coordinated with an organization to use a game-based training with employees and then gathered job performance data for the employees, the impact that the training had on relevant work-related outcomes such as job performance could be clearly seen. Providing evidence about the implications of game-based training for job performance can help an organization better understand the return on investment for game-based learning. A major limitation of existing research in this area and of the current study is that implications for transfer of training and job performance are being made using proxy measures of learning (e.g., declarative knowledge). Being able to draw these connections directly could aid organizations in making decisions about investing in and developing future game-based learning programs. Thus, the current platform of using a lab-based experiment limits the generalizability of these results as the measures and methodologies did not resemble the real-world context.

Limitations

There are several important factors to consider when interpreting the results of the current study. One concerns conclusions of causality. Because I manipulated human interaction, there is some confidence in arguing for a causal effect between that game characteristic and the outcome variables (i.e., affective states and declarative knowledge). However, because the mediating variables and outcome variables were measured simultaneously, causal inferences cannot be drawn for any mediating relationships. While the mediating and outcome variables were related in my study, I do not have evidence that the mediators occurred prior to the training outcomes, or that these variables were the process through which human interaction affected the training outcomes. Ideally, future studies examining gameplay processes would measure potential mediators during the flow of the game. For example, the researchers could pause the game and measure the potential mechanisms or utilize a think-aloud protocol to capture participants' thoughts during the training. Another alternative could be a repeated-measures design, in which participants complete multiple iterations of the game and report on the process variables between these interactive periods. These methods would better align the measurement of the potential mediator with the time point in which it operates and strengthen the inference that the mediators have a causal impact on the training outcomes.

Further, as noted above, although the paths from the mediating variables to the outcomes were generally significant and in the predicted direction, the paths from human interaction to the mediating variables were weak. It may be that the human interaction manipulation was not robust enough to create the expected main effects. Specifically, the differences between the experiences of those who played alone and those who played with a group may have been too similar. The experiment could have been redesigned to allow multiple iterations of the game. With several phases of team interaction, the distinctions between the individual and team group may have been clearer and had a stronger effect on the tested process variables.

Another methodological concern is that most of the measures were gathered using self-report data. This methodology has been criticized as having several drawbacks (Podsakoff & Organ, 1986). For example, self-report data relies heavily on the participant's ability to understand and accurately recall their behavior. A further concern with self-report measures is that participant responses can be easily influenced by other factors. In this case *order effect* may have been an issue, in which previous questions presented in the survey influenced how later questions were answered. This is because multiple scales were collected during time 2 in the second survey. It is possible that questions asked early in the survey influenced how participants considered and answered questions presented later in the survey. This may have inadvertently primed participants to consider, feel, or remember their experience in a particular way (Kahneman & Tversky, 1984). One further concern with the self-report measures used in the present study is that participants were assumed to have an accurate understanding of their own experience in a way that they can correctly report back that own experience. Previous research has shown that different conditions (e.g., current emotional state, triggered memories) and experiences can influence, alter, and contaminate the way individuals recall and perceives their experience (Howard, 1980; Stone, Bachrach, Jobe, Kurtzman, & Cain, 1999).

A final consideration is that many of the variables in the current study were similar in nature and highly related to one another. I found a substantial amount of overlap between the measured constructs. This made it challenging to distinguish the constructs in the measurement model, which likely contributed to the number of items that needed to be removed from the different scales (e.g., motivation). The modifications to the validated scales may impact the results. It is unclear what unintended consequences could have resulted from removing content that was too similar to content of other variables in the study. These changes may have been meaningful to the constructs or the outcomes. However, previous research has shown that game features are highly interdependent and difficult to parse apart (Bedwell et al., 2012). It is important to remember that this is a single study and replication of this methodology is needed for further confidence in the results.

Based on the current findings, it appears that there are potential benefits from continuing to explore games as a training tool. The current findings demonstrated that the game characteristic of human interaction has a large positive effect on affective states and a small but significant effect on declarative knowledge. Additionally, perceived value and active learning were important mechanisms through which human interaction impacted affective states. These emerging areas of research show there are potential avenues for continued research in the application of game-based training and education.

Table 1. Learning objectives and sub-objectives for the Computer-Based Training (CBT)

Objectives	**Sub-Objectives**
Accessing the Game	Participants will be able to locate and join the ship without assistance.
Using the Main Controls	Participants can demonstrate knowledge of the main controls and can navigate the main controls screen.
Playing the Scenarios	Participants will understand the scenarios of the game and the rules and goals for playing.
Knowing the Ship Stations	Participants know the purpose and abilities of the *Helm* station including controls and strategies.
	Participants know the purpose and abilities of the *Tactical* station including controls, strategies, and teammates to prioritize communication with.
	Participants know the purpose and abilities of the *Engineering* station including controls, strategies, and equipment this role is responsible for.
	Participants know the purpose and abilities of the *Scientist* station including controls, strategies, equipment, and primary tasks.
Knowing the Alliances	Participant knows the different alliances in the game including their abilities, strengths, and ships (style & size).

Table 2. Summary and sample items for study scales

Measure	Citation	No. of Items	Sample Item
Videogame Experience (VGE)	Sanchez, 2017; Singer & Knerr, 2010	14	*I currently play videogames several hours every week.*
Declarative Knowledge Pre-test	- -	14	*Which station does the Tactical station need to communicate with to boost their cooldown?*
Declarative Knowledge Post-test	- -	26	*Which station is responsible for communicating to the other team members the current health of the ship?*
Psychological Meaning (PM)	May et al., 2004	6	*I feel that the work I did in the activity was valuable.*
Perceived Enjoyment (PEJ)	Adomaityte, 2013; Calvillo-Gámez et al., 2015; Long, 2005; Ricci et al., 1996; Warr et al., 1999	6	*I really enjoyed this activity.*
Motivation (MOT)	Keller, 2006	6	*The variety of tasks in the activity helped keep my attention.*
Emotional Engagement (EEG)	Rich et al., 2010	6	*I was enthusiastic in this activity.*
Physical Engagement (PEG)	Rich et al., 2010	6	*I worked with intensity in the activity.*
Cognitive Engagement (CEG	Rich et al., 2010	6	*During this activity, my mind was focused on the task at hand.*
Perceived Value (PV)	Keller, 2006	6	*The content of this activity will be useful to me.*
Active Learning (AL)	Walker & Fraser, 2005	3	*I explored my own strategies in the activity.*

Note. All measures used a 5-point Likert type scale, 1 = Strongly Disagree to 5 = Strongly Agree.

Table 3. Means, standard deviations, and correlations for observed study variables

	M	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Age	18.98	1.72														
2. Sex	.49	.50	.18**													
3. Videogame Experience	2.74	.98	.12*	.66**	*.94*											
4. Human Interaction	.68	.47	-.07	-.06	.02											
5. Declarative Knowledge Pre-test	.63	.21	.06	.24**	.44**	.02										
6. Declarative Knowledge Post-test	.67	.18	.08	.30**	.50**	.01	.67**									
7. Psychological Meaning	3.12	.77	-.06	.07	.31**	.23**	.22**	.16**	*.88*							
8. Perceived Enjoyment	3.40	.89	-.02	.14**	.39**	.29**	.31**	.29**	.73**	*.94*						
9. Motivation	3.49	.78	.03	.14**	.37**	.20**	.29**	.30**	.71**	.81**	*.84*					
10. Emotional Engagement	3.32	.81	-.01	.16**	.34**	.22**	.27**	.27**	.71**	.81**	.75**	*.93*				
11. Physical Engagement	3.31	.81	.06	.05	.16**	.01	.13*	.25**	.48**	.44**	.45**	.56**	*.90*			
12. Cognitive Engagement	3.80	.75	.13**	.10	.23**	.07	.28**	.34**	.48**	.52**	.55**	.59**	.61**	*.96*		
13. Perceived Value	3.01	.78	.01	.25**	.45**	.17**	.30**	.28**	.61**	.60**	.64**	.58**	.29**	.37**	*.78*	
14. Active Learning	3.62	.81	.02	.18**	.23**	-.17**	.20**	.23**	.14**	.27**	.26**	.25**	.25**	.33**	.23**	*.83*

Note. n = 385; Sex 0 = Female, 1 = Male; Human Interaction 0 = Individual, 1 = Team; *p < .05, **p < .01, ***, p < .001. Italicized values along the diagonal are the Cronbach's alpha reliability coefficient. Scales were calculated using average scores.

Table 4. Intra-class correlations for study measures

Measure	ICC
Psychological Meaning (PM)	***range .04 - .16***
PM1	.15
PM2	.12
PM3	.12
PM4	.16
PM5	.04
Perceived Enjoyment (PEJ)	***range .11 - .18***
PEJ1	.17
PEJ2	.17
PEJ3	.16
PEJ4	.18
PEJ5	.11
Motivation (MOT)	***range .08 - .17***
MOT1	.17
MOT2	.08
MOT3	.15
MOT6	.08
MOT7	.10
Emotional Engagement (EEG)	***range .04 - .12***
EEG1	.07
EEG2	.09
EEG3	.04
EEG4	.12
EEG5	.08
EEG6	.07

Table 4. Intra-class correlations for study measures (cont.)

Measure	ICC
Physical Engagement (PEG)	***range .02 - .07***
PEG2	.07
PEG3	.05
PEG4	.02
PEG5	.02
Cognitive Engagement (CEG)	***range .02 - .19***
CEG1	.02
CEG2	.04
CEG3	.19
CEG4	.11
CEG5	.06
CEG6	.08
Perceived Value (PV)	***range .04 - .14***
PV1	.04
PV2	.09
PV4	.14
PV6	.13
Active Learning (AL)	***range .02 - .08***
AL1	.02
AL2	.04
AL3	.08

Table 5. Summary of fit statistics for progression of tested models

		χ^2	*df*	*p*	*RMSEA [95% CI]*	*p*	*CFI*	*TLI*	*SRMR*	*Delta χ^2*	*Delta df*	*p*
Measurement Model												
1	Original	3735.89	1573	<.001	.060 [.057, .062]	<.001	.884	.879	.059	- -	- -	- -
2	PV5 removed for low factor loading	3619.56	1517	<.001	.060 [.057, .063]	<.001	.887	.881	.059	116.33	56	<.001
3	MOT4 removed for low factor loading	3501.13	1462	<.001	.060 [.058, .063]	<.001	.890	.884	.059	118.44	55	<.001
4	VGE14 removed for high discrepancy	3153.06	1403	<.001	.057 [.054, .059]	<.001	.903	.898	.058	348.07	59	<.001
5	PEG1 removed for high discrepancy	3037.22	1355	<.001	.057 [.054, .059]	<.001	.905	.900	.056	115.84	48	<.001
6	VGE13 removed for high discrepancy	2850.08	1303	<.001	.056 [.053, .058]	<.001	.910	.905	.055	187.14	52	<.001
7	PEJ6 removed for high discrepancy	2697.94	1252	<.001	.055 [.052, .058]	.003	.915	.910	.054	152.13	51	<.001
8	PEG6 removed for high discrepancy	2535.20	1202	<.001	.054 [.051, .057]	.019	.920	.920	.054	162.74	50	<.001
9	VGE6 removed for high discrepancy	2449.88	1153	<.001	.054 [.051, .057]	.013	.920	.915	.053	85.32	49	<.001
10	PV3 removed for high discrepancy	2343.68	1105	<.001	.054 [.051, .057]	.016	.922	.917	.052	106.20	48	<.001
- -	*Items loaded directly onto AS (dropped)*	*3049.90*	*1109*	*<.001*	*.067 [.065, .070]*	*<.001*	*.878*	*.871*	*.055*	*706.22*	*4*	*<.001*
11	**HI and DK added into model (final)**	**2461.44**	**1192**	**<.001**	**.053 [.050, .056]**	**.073**	**.921**	**.916**	**.052**	**117.76**	**87**	**.016**
Structural Equation Model												
12	Original	2836.25	1206	<.001	.059 [.056, .062]	<.001	.899	.893	.172	374.81	14	<.001
13	**PEG, CEG, PV AL correlated (final)**	**2556.71**	**1200**	**<.001**	**.054 [.051, .057]**	**.009**	**.916**	**.911**	**.107**	**95.27**	**8**	**<.001**
- -	*Clustered by groups (dropped)*	*2334.41*	*1200*	*<.001*	*.050 [.047, .053]*	*.593*	*.919*	*.914*	*.107*	*127.03*	*8*	*<.001*
Research Question Model												
13	**Original (final)**	**2808.99**	**1200**	**<.001**	**.059 [.056, .062]**	**<.001**	**.900**	**.894**	**.170**	**347.55**	**8**	**<.001**

Note. HI = Human Interaction, DK = Declarative Knowledge, VGE = Video Game Experience, PEJ = Perceived Enjoyment, MOT = Motivation, AS = Affective States, PEG = Physical Engagement, CEG = Cognitive Engagement, PV = Perceived Value and AL = Active Learning

Table 6. Factor loadings for study scales

Factor	Initial Loading	Final Loading
Video Game Experience (VGE)		
Earnest Video Gaming	***.95***	***.95***
I currently play videogames several hours every week. VGE1	.82	.82
I have spent a lot of time reading videogame magazines or websites to find tips to improve my gaming skills. VGE3	.71	.70
Playing videogames is some of the most fun I have ever had in my life. VGE6	.77	- -
I consider myself to have a lot of videogame experience. VGE7	.93	.94
I would call myself a "serious gamer". VGE8	.81	.81
Casual Video Gaming	***.93***	***.93***
In the past I have spent several hours in one week playing videogames. VGE2	.86	.86
I would call myself a "casual gamer". VGE9	.75	.76
**I am scared of videogames. VGE10*	.47	.47
**I don't enjoy playing videogames at all. VGE11*	.69	.69
Video Gaming Self-Efficacy	***.95***	***.94***
I have a high level of confidence with playing videogames in general. VGE4	.95	.96
I have a high level of skill at playing videogames in general. VGE5	.95	.96
I have complete control over whether or not I do well in a videogame. VGE12	.58	.54
I am confident that I can perform well in a videogame. VGE13	.81	- -
I believe I have the ability to play videogames well. VGE14	.76	- -

*Reverse scored items.

Table 6. Factor loadings for study scales (cont.)

Factor	Initial Loading	Final Loading
Affective States		
Psychological Meaning (PM)	***.86***	***.86***
My time was well spent in the activity. PM1	.87	.87
Doing the activity was a meaningful experience. PM2	.78	.78
**My time doing the activity was NOT well spent. PM3*	.69	.69
Doing the activity was worthwhile. PM4	.83	.83
The activity was important to me. PM5	.73	.73
Perceived Enjoyment (PEJ)	***.92***	***.92***
I was very satisfied with this activity. PEJ1	.82	.82
I had a very positive experience in this activity. PEJ2	.86	.86
I really enjoyed this activity. PEJ3	.94	.94
This activity was very fun. PEJ4	.93	.93
I would enjoy doing this activity again. PEJ5	.86	.85
**This activity was a waste of my time. PEJ6*	.68	- -
Motivation (MOT)	***.95***	***.95***
The activity was interesting to me. MOT1	.85	.85
**The activity was dry and unappealing. MOT2*	.72	.71
This activity stimulated my curiosity. MOT3	.79	.80
**The activity was repetitive. MOT4*	.29	- -
The variety of tasks in the activity helped keep my attention. MOT5	.70	.71
**The activity was boring. MOT6*	.73	.72
Emotional Engagement (EEG)	***.92***	***.92***
I was enthusiastic in this activity. EEG1	.83	.83
I felt energetic in this activity. EEG2	.83	.83
I was interested in this activity. EEG3	.85	.85
I felt proud of my work in this activity. EEG4	.74	.74
I felt positive about this activity. EEG5	.84	.85
I felt excited about this activity. EEG6	.87	.87

*Reverse scored items.

Table 6. Factor loadings for study scales (cont.)

Factor	Initial Loading	Final Loading
Physical Engagement (PEG)		
I worked with intensity in the activity. PEG1	.70	- -
I exerted my full effort in the activity. PEG2	.84	.81
I devoted a lot of energy to the activity. PEG3	.79	.74
I tried my hardest to perform well in the activity. PEG4	.82	.87
I strove as hard as I could to complete the activity. PEG5	.84	.87
I exerted a lot of energy in this activity. PEG6	.63	- -
Cognitive Engagement (CEG)		
During this activity, my mind was focused on the task at hand. CEG1	.88	.88
During this activity, I paid a lot of attention to the task at hand. CEG2	.91	.91
During this activity, I focused a great deal of attention on the task at hand. CEG3	.90	.90
During this activity, I was absorbed by the task at hand. CEG4	.85	.85
During this activity, I concentrated on the task at hand. CEG5	.90	.91
During the activity, I devote a lot of attention to the task at hand. CEG6	.90	.89
Perceived Value (PV)		
It is clear to me how the content of this activity is related to things I already know. PV1	.64	.61
There were aspects of this activity that could be important to some people. PV2	.58	.61
The content of this activity is relevant to my interests. PV3	.82	- -
The content of this activity is worth knowing. PV4	.81	.83
**The content of this activity was not relevant to my needs. PV5*	.09	- -
The content of this activity will be useful to me. PV6	.79	.82
Active Learning (AL)		
I explored my own strategies in the activity. AL1	.67	.67
In the activity I sought my own answers. AL2	.91	.91
In the activity I solved my own problems. AL3	.80	.80

*Reverse scored items.

Table 7. Summary of items removed due to large residual correlations

Item	Discrepancy
Video Game Experience Scale (VGE) (Earnest Video Gaming) *Playing videogames is some of the most fun I have ever had in my life. VGE6*	.17 with *I was interested in this activity. EEG3* .16 with *The activity was interesting to me. MOT1* .16 with *I worked with intensity in the activity. PEG1* .19 with *There were aspects of this activity that could be important to some people. PV2*
Video Game Experience Scale (VGE) (Video Gaming Self-Efficacy) *I am confident that I can perform well in a videogame. VGE13*	.25 with *I have complete control over whether or not I do well in a videogame. VGE12* .18 with *I was interested in this activity. EEG3* .17 with *It is clear to me how the content of this activity is related to things I already know. PV1*
Perceived Enjoyment (PEJ) **This activity was a waste of my time. PEJ6*	.22 with **My time doing the activity was NOT well spent. PM3* .21 with **The activity was dry and unappealing. MOT2* .18 with **The activity was boring. MOT6*

*Reverse scored item. Note. VGE = Video Game Experience, PM = Psychological Meaning, MOT = Motivation, EEG = Emotional Engagement, PEG = Physical Engagement, PV = Perceived Value, and AL = Active Learning.

Table 7. Summary of items removed due to large residual correlations (cont.)

Item	Discrepancy
Physical Engagement (PEG) *I worked with intensity in the activity. PEG1*	.16 with *Playing videogames is some of the most fun I have ever had in my life. VGE6* .23 with *I was enthusiastic in this activity. EEG1* .23 with *I felt energetic in this activity. EEG2* .19 with *I was interested in this activity. EEG3* .15 with *I felt excited about this activity. EEG6* .15 with *The variety of tasks in the activity helped keep my attention. MOT5* .18 with *I explored my own strategies in the activity. AL1*
Physical Engagement (PEG) *I exerted a lot of energy in this activity. PEG6*	-.16 with **I am scared of videogames. VGE10* .17 with *The activity was important to me. PM5* .15 with *I felt energetic in this activity. EEG2* .17 with *I devoted a lot of energy to the activity. PEG3*
Perceived Value (PV) *The content of this activity is relevant to my interests. PV3*	.16 with *I currently play videogames several hours every week. VGE1* .19 with *Playing videogames is some of the most fun I have ever had in my life. VGE6* .17 with *I would call myself a "serious gamer". VGE8*

*Reverse scored item. Note. VGE = Video Game Experience, PM = Psychological Meaning, MOT = Motivation, EEG = Emotional Engagement, PEG = Physical Engagement, PV = Perceived Value, and AL = Active Learning.

Table 8. Correlations among latent study variables

	1	2	3	4	5
1. Video Game Experience					
2. Affective States	$.35^{***}$				
3. Physical Engagement	$.16^{**}$	$.56^{***}$			
4. Cognitive Engagement	$.21^{***}$	$.62^{***}$	$.65^{***}$		
5. Perceived Value	$.43^{***}$	$.75^{***}$	$.32^{***}$	$.41^{***}$	
6. Active Learning	$.22^{***}$	$.25^{***}$	$.26^{***}$	$.33^{***}$	$.21^{***}$

Note. n = 385. $^{**}p < .01$, ***, $p < .001$

Table 9. Direct and indirect effect results for hypotheses model

	AS	DK	PEG	CEG	PV	AL
	b(SE) *[95% CI]*	*b(SE)* *[95% CI]*	*b(SE)* *[95% CI]*	*b(SE)* *[95% CI]*	*b(SE)* *[95% CI]*	*b(SE)* *[95% CI]*
Direct Effects						
Human Interaction	.24(.05)*** [.14, .35]	.01(.02) [-.03, .04]	.01(.09) [-.17, .18]	.10(.08) [-.05, .26]	.31(.10)*** [.12, .49]	-.26(.07)*** [-.39, -.13]
Video Game Experience	.04(.02) [-.01, .09]	.07(.01)*** [.05, .09]				
Physical Engagement	.22(.05)*** [.13, .31]					
Cognitive Engagement	.24(.05)*** [.14, .34]					
Perceived Value	.47(.05)*** [.37, .56]					
Active Learning		.05(.02)** [.02, .08]				
Indirect Effects						
Via Physical Engagement	.002(.02) [-.04, .04]					
Via Cognitive Engagement	.03(.02) [-.01, .06]					
Via Perceived Value	.14(.05)** [.05, .23]					
Via Active Learning		-.01(.01)* [-.02, -.002]				

Note. n = 385; *p < .05, **p < .01, ***, p < .001. Reported values are unstandardized regression coefficients. AS = Affective States, DK = Declarative Knowledge, PEG = Physical Engagement, CEG = Cognitive Engagement, PV = Perceived Value, AL = Active Learning.

Table 10. Direct and indirect effect results for research question model

	AS	DK	PEG	CEG	PV	AL
	b(SE)* *[95% CI]	***b(SE)* *[95% CI]***	***b(SE)* *[95% CI]***	***b(SE)* *[95% CI]***	***b(SE)* *[95% CI]***	***b(SE)* *[95% CI]***
Direct Effects						
Human Interaction	.50(.08)*** [.35, .64]	-.02(.02) [-.05, .02]	.01(.09) [-.17, .18]	.10(.08) [-.05, .26]	.31(.10)*** [.12, .50]	-.26(.07)*** [-.40, -.13]
Video Game Experience	.17(.03)*** [.11, .24]	.06(.01)*** [.04, .08]				
Physical Engagement		.01(.02) [-.02, .04]				
Cognitive Engagement		.06(.02)*** [.03, .09]				
Perceived Value		.01(.01) [-.02, .04]				
Active Learning	.38(.07)*** [.24, .51]					
Indirect Effects						
Via Physical Engagement		<.001(.001) [-.002, .002]				
Via Cognitive Engagement		.01(.01) [-.004, .02]				
Via Perceived Value		.004(.01) [-.01, .01]				
Via Active Learning	-.10(.03)*** [-.16, -.04]					

Note. n = 385; $^{*}p < .05$, $^{**}p < .01$, ***, $p < .001$. Reported values are unstandardized regression coefficients. AS = Affective States, DK = Declarative Knowledge, PEG = Physical Engagement, CEG = Cognitive Engagement, PV = Perceived Value, AL = Active Learning.

Table 11. Comparison of indirect effect sizes across models

	Pm	*SE*	*p*	*95% CI*
Hypothesized Model				
human interaction → physical engagement→ affective states	.006	.08	.93	-.15, .16
human interaction → cognitive engagement→ affective states	.09	.07	.18	-.04, .23
human interaction → perceived value → affective states	.37	.10	<.001	.18, .56
human interaction → active learning → declarative knowledge	2.00	5.56	.72	-8.89, 12.89
Research Question Model				
human interaction → physical engagement→ declarative knowledge	-.01	.07	.94	-.15, .13
human interaction → cognitive engagement→ declarative knowledge	-.62	1.34	.65	-3.24, 2.01
human interaction → perceived value → declarative knowledge	-.34	.64	.59	-1.60, .91
human interaction → active learning → affective states	-.25	.10	.01	-.44, -.06

Table 12. Profile of videogame characteristics for current study using the taxonomy from Bedwell et al., 2012

Characteristic	Rating
Action Language	High
Assessment	High
Control	Moderate
Challenge/Conflict	High
Game Fiction	High
Human Interaction	Low & High
Rules/Goals	High
Environment	Low
Immersion	High

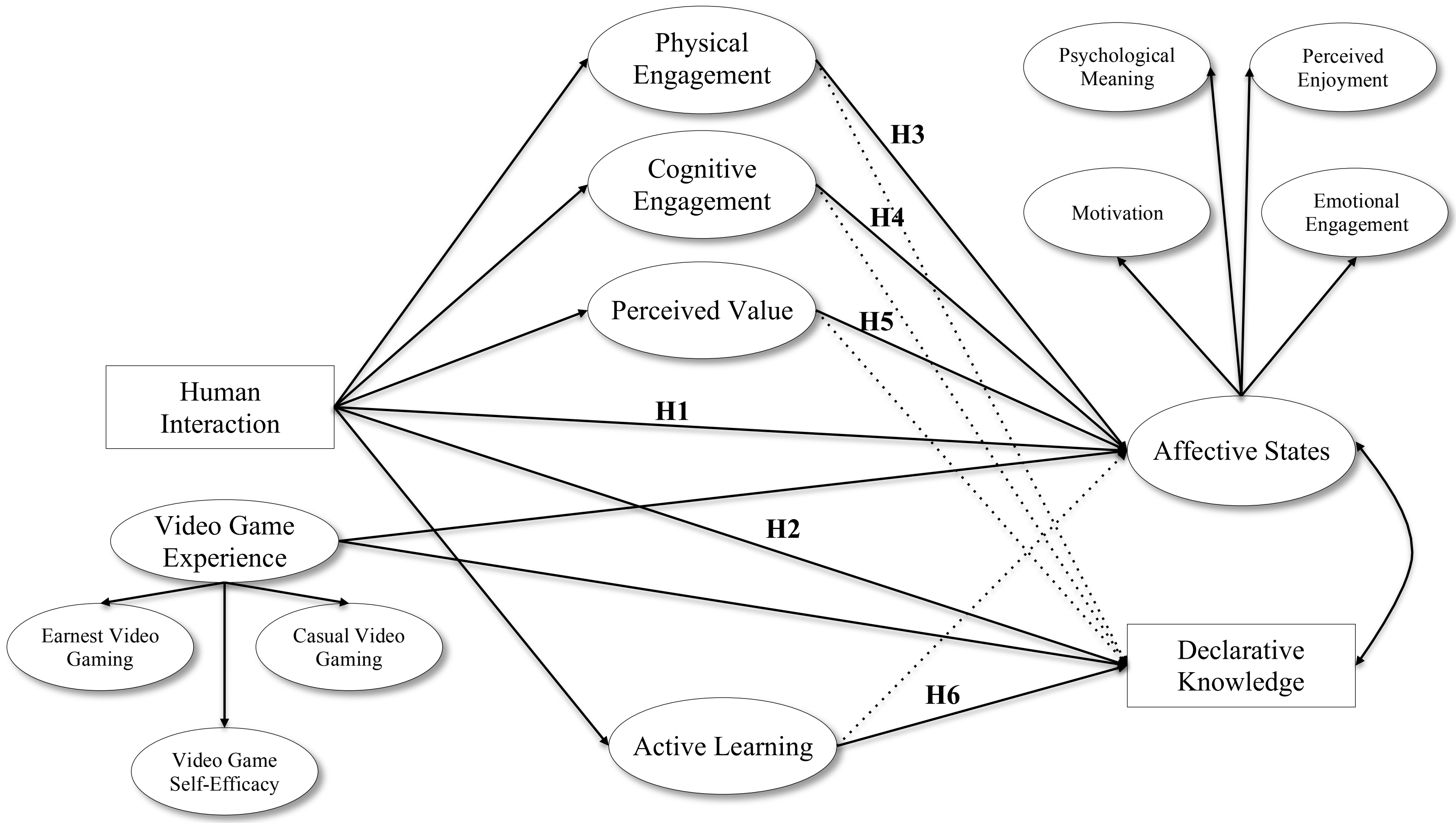


Figure 1. Hypothesized direct and indirect effects for human interaction with affective states and declarative knowledge
Note. Dotted lines denote relationships included as part of the research question.

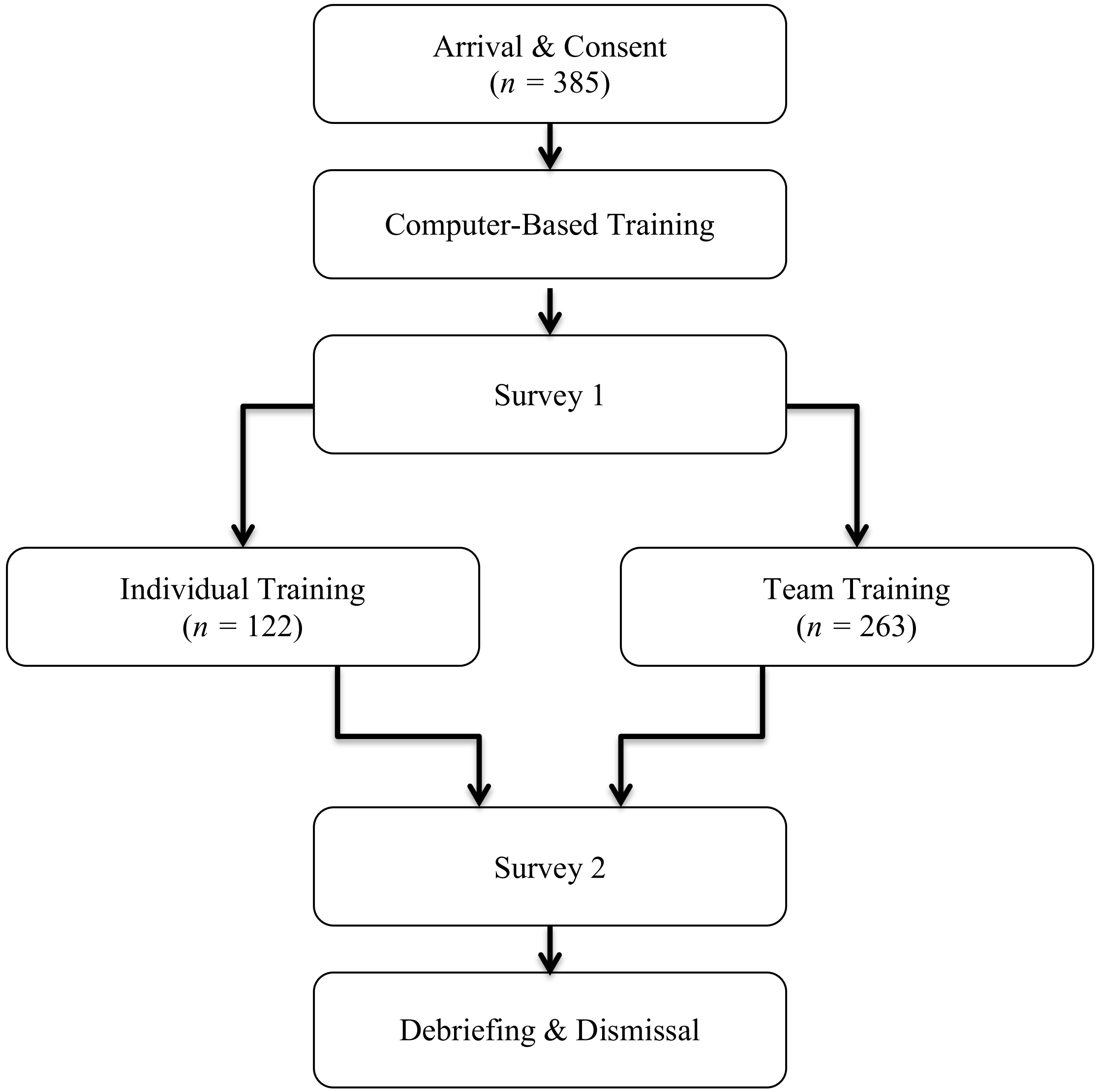


Figure 2. Procedures for experiment with final sample of participants

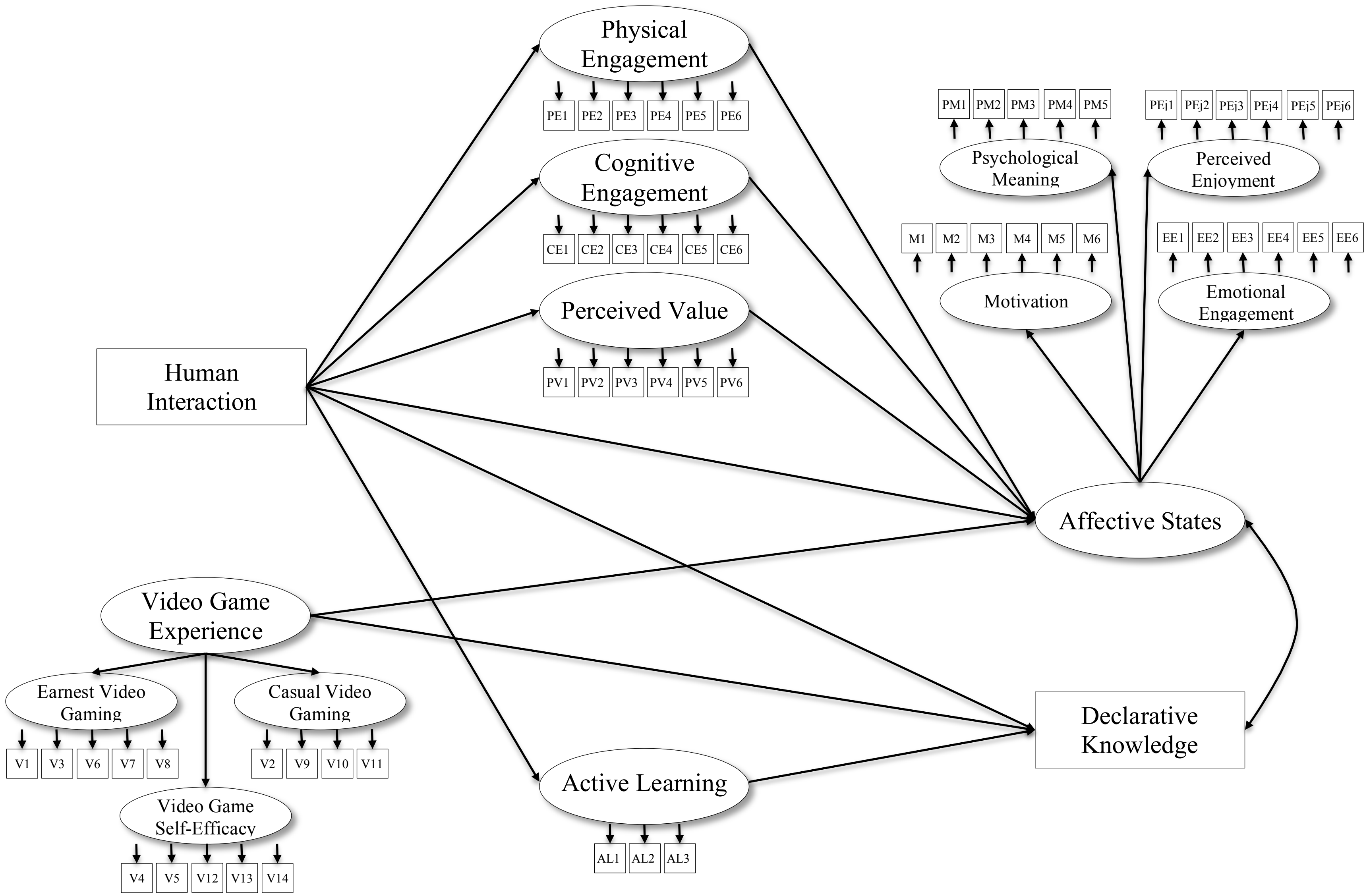


Figure 3. Proposed structural equation model

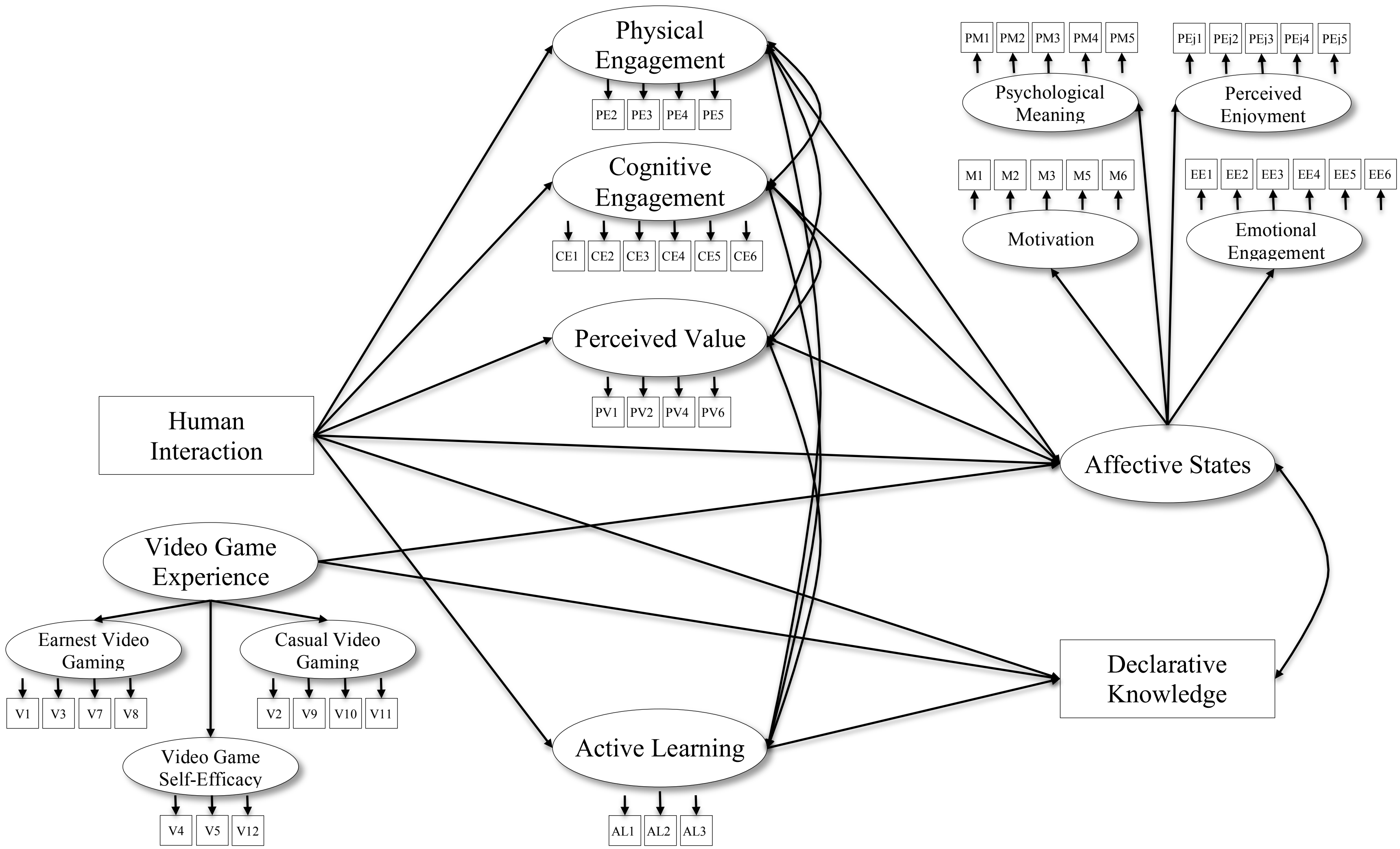


Figure 4. Final structural equation model

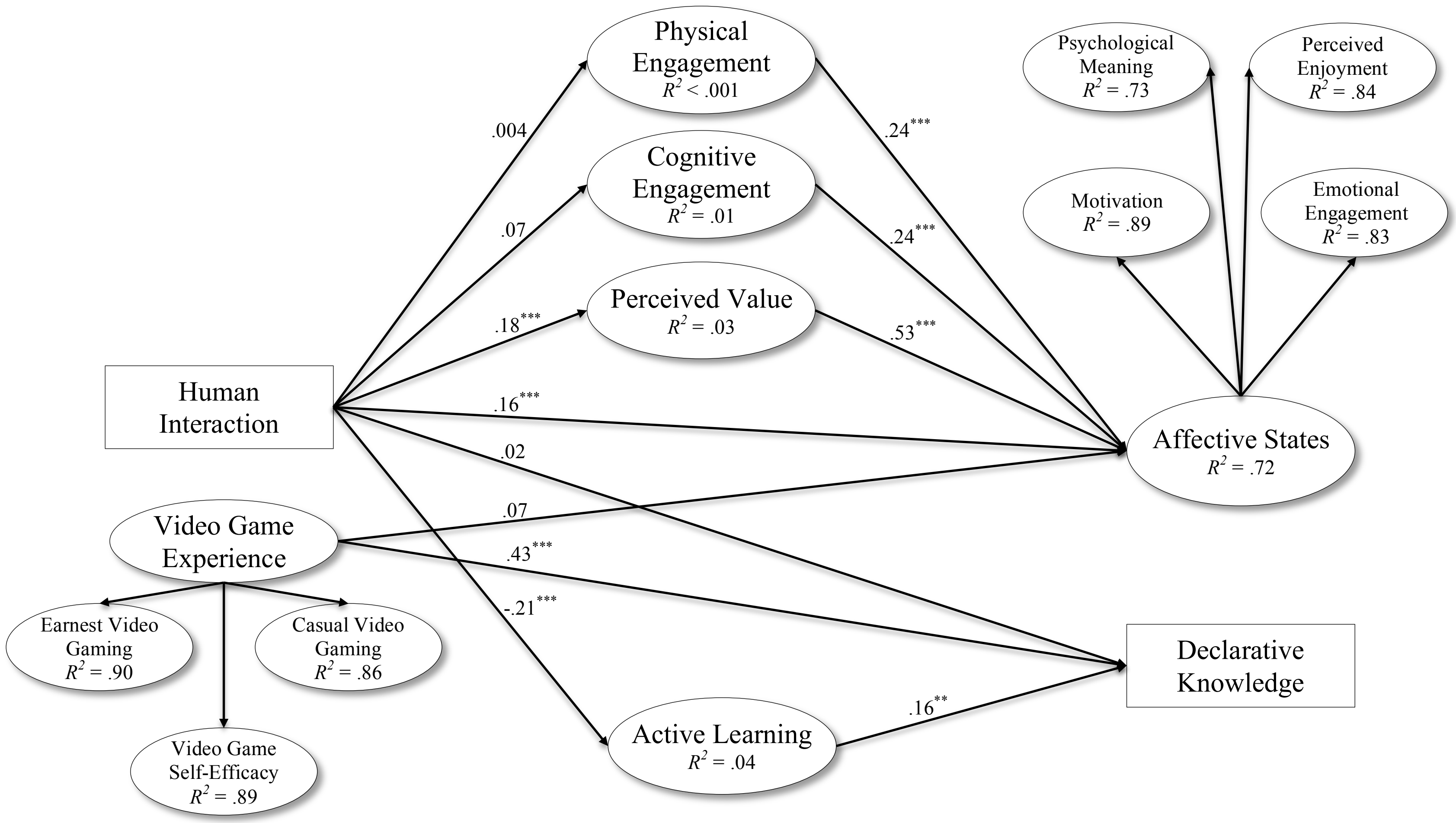


Figure 5. Standardized estimates for hypothesized direct and indirect effects

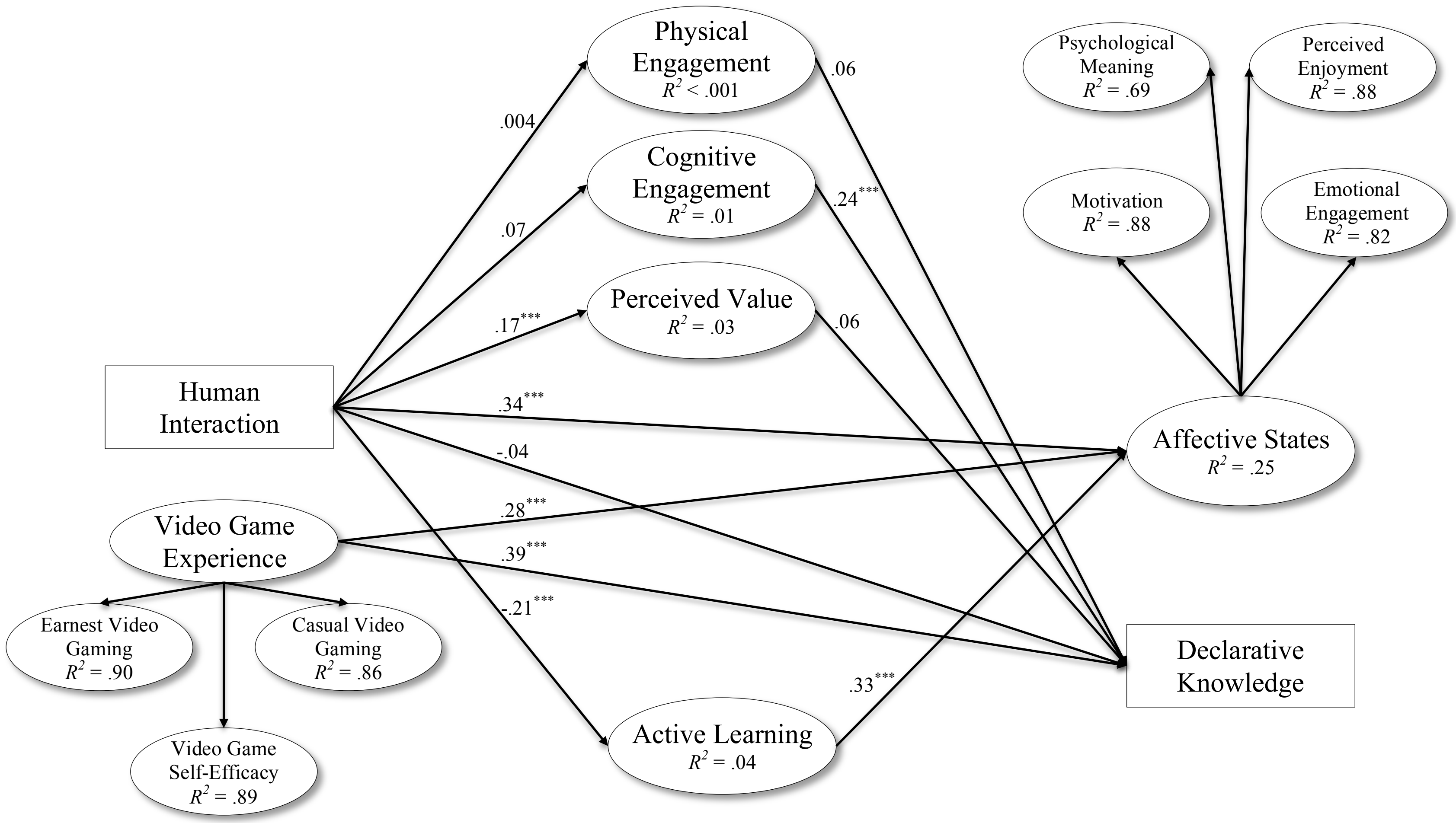


Figure 6. Standardized estimates for research question direct and indirect effects

REFERENCES

Aabakken, L., Adamsen, S., & Kruse, A. (2000). Performance of a colonoscopy simulator: Experience from a hands-on endoscopy course. *Endoscopy, 32*, 911-913.

Adams, D. M. (2013). *Exploring the effect of video game playing on static and dynamic spatial cognition skills* (Unpublished doctoral dissertation). University of California, Santa Barbara, CA.

Adams, D. M., Mayer, R. E., MacNamara, A., Koenig, A., & Wainess, R. (2012). Narrative games for learning: Testing the discovery and narrative hypotheses. *Journal of Educational Psychology, 104*, 235–249. doi:10.1037/a0025595

Adomaityte, A. (2013). *Predicting post-training reactions from pre-training attitudes* (Unpublished doctoral dissertation). University of Montreal, Quebec, Canada.

Ainley, M., & Ainley, J. (2011). Student engagement with science in early adolescence: The contribution of enjoyment to students' continuing interest in learning about science. *Contemporary Educational Psychology, 36*(1), 4-12.

Alexander, P. A., & Judy, J. E. (1988). The interaction of domain-specific and strategic knowledge in academic performance. *Review of Educational Research, 58*, 375-404.

Alliger, G. M., Tannenbaum, S. I., Bennett, W., Traver, H., & Shotland, A. (1997). A meta-analysis of the relations among training criteria. *Personnel Psychology, 50*, 341-358.

Alwin, D. F., & Hauser, R. M. (1975). The decomposition of effects in path analysis. *American Sociological Review*, *40*, 37-47.

Amad, A. (2015). The advantages of Wasabi Waiter Apps in hiring. *The Science Times*. Retrieved June 27, 2017 from http://www.sciencetimes.com/articles/7384/20150924/advantages-wasabi-waiter-apps-hiring.htm

Amer, R. S., Denehy, G. E., Cobb, D. S., Dawson, D. V., Cunningham-Ford, M. A., & Bergeron, C. (2011). Development and evaluation of an interactive dental video game to teach dentin bonding. *Journal of Dental Education, 75*, 823-831.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Annetta, L. A. (2008). Video games in education: Why they should be used and how they are being used. *Theory into Practice, 47*, 229-239.

Aragón-Correa, J. A., Martín-Tapia, I., & Hurtado-Torres, N. E. (2013). Proactive environmental strategies and employee inclusion: The positive effects of information sharing and promoting collaboration and the influence of uncertainty. *Organization & Environment, 26*, 139-161.

Argote, L. (1999). *Organizational learning: Creating, retaining and transferring knowledge*. Boston, MA: Kluwer Academic Publishers.

Arnab, S., Lim, T., Carvalho, M. B., Bellotti, F., Freitas, S., Louchart, S., ... & De Gloria, A. (2015). Mapping learning and game mechanics for serious games analysis. *British Journal of Educational Technology, 46*, 391-411.

Bauer, K. N., Brusso, R. C., & Orvis, K. A. (2012). Using adaptive difficulty to optimize videogame-based training performance: The moderating role of personality. *Military Psychology, 24*, 148-165.

Beach, D. (1980). *Personnel: The management of people at work*. Macmillan, New York.

Beal, S. A., & Christ, R. E. (2004). *Training effectiveness evaluation of the Full Spectrum Command game* (DHHS Publication No. ARI-TR-1140). Fort Benning, GA: Army Research Institute for the Behavioral and Social Sciences.

Bedwell, W. L., Pavlas, D., Heyne, K., Lazzara, E. H., & Salas, E. (2012). Toward a taxonomy linking game attributes to learning an empirical study. *Simulation & Gaming, 43*, 729-760.

Bell, C., & Putman, T. (1979). Mastering the art of training design. *Training and Development Journal, 33*, 24-27.

Bernhaupt, R. (Ed.). (2010). *Evaluating user experience in games*. London, England: Springer.

Blair, L. (2011). *The use of video game achievements to enhance player performance, self-efficacy, and motivation* (Unpublished doctoral dissertation), University of Central Florida, Orlando, FL.

Blume, B. D., Ford, J. K., Baldwin, T. T., & Huang, J. L. (2010). Transfer of training: A meta-analytic review. *Journal of Management, 36*, 1065-1105.

Blunt, R. (2009). Does game-based learning work? Results from three recent studies. *eLearn Magazine*. Retrieved from http://elearnmag.acm.org/archive.cfm?aid=1661378

Bonwell, C.C., & Eison, J. A. (1991). *Active learning: Creating excitement in the classroom*. Washington, DC: ERIC Clearinghouse on Higher Education.

Broad, M. L., & Newstrom, J. W. (1992). *Transfer of training: Action-packed strategies to ensure high payoff from training investments.* Reading, MA: Addison-Wesley.

Brodbeck, F., & Greitemeyer, T. (2000a). A dynamic model of group performance: Considering the group members' capacity to learn. *Group Processes & Intergroup Relations, 3*, 159-182.

Brodbeck, F. C., & Greitemeyer, T. (2000b). Effects of individual versus mixed individual and group experience in rule induction on group member learning and group performance. *Journal of Experimental Social Psychology*, *36*, 621-648.

Brown, K. G. (2001). Using computers to deliver training: Which employees learn and why? *Personnel Psychology, 54*, 271-296. doi:10.1111/j.1744-6570.2001.tb00093.x

Brown, K. G., & Ford, K. (2002). Using computer technology in training building an infrastructure for active learning. In K. Kraiger (Ed.) *Creating, implementing, and managing effective training and development: State-of-the-art lessons for practice* (pp. 192-225). San Francisco: CA, Jossey-Bass.

Brown, K. G., & Gerhardt, M. W. (2002). Formative evaluation: An integrative practice model and case study. *Personnel Psychology, 55*, 951-983.

Browning, R. F., Ryan, L. E., Scott, P. G., & Smode, A. F. (1977). *Training effectiveness evaluation of device 2F87F, P-3C Operational Flight Trainer* (No. ADA 035771). Orlando, FL: Training Analysis and Evaluation Group.

Burguillo, J. C. (2010). Using game theory and competition-based learning to stimulate student motivation and performance. *Computers & Education, 55*, 566-575.

Butler, Y. G., Someya, Y., & Fukuhara, E. (2014). Online games for young learners' foreign language learning. *ELT Journal, 68*, 265-275.

Calvillo-Gámez, E. H., Cairns, P., & Cox, A. L. (2015). Assessing the core elements of the gaming experience. In R. Berhnhaupt (Ed.) *Evaluating user experience in games* (pp. 37-62). London: Springer International Publishing.

Campbell, D. T., & Riecken, H. W. (1968). Quasi-experimental design*. International Encyclopedia of the Social Sciences, 5*, 259-263.

Cannon-Bowers, J. (2006). Learning activities and game genres. *Training 2006 Conference and Expo.*

Caught, K., Shadur, M. A., & Rodwell, J. J. (2000). The measurement artifact in the organizational commitment questionnaire. *Psychological Reports*, *87*, 777-788.

Chadwick, O., Momcilovic, N., Rossiter, R., Stumbles, E., & Taylor, E. (2001). A randomized trial of brief individual versus group parent training for behaviour problems in children with severe learning disabilities. *Behavioural and Cognitive Psychotherapy*, *29*, 151-167.

Chang, J., Lee, M., Ng, K. L., & Moon, K. L. (2003). Business simulation games: the Hong Kong experience. *Simulation & Gaming, 34*, 367-376.

Chen, H. H. (2005). *A formative evaluation of the training effectiveness of a computer game* (Unpublished doctoral dissertation). University of Southern California, Los Angeles, CA.

Cheng, A., Eppich, W., Grant, V., Sherbino, J., Zendejas, B., & Cook, D. A. (2014). Debriefing for technology‐enhanced simulation: A systematic review and meta‐analysis. *Medical Education, 48*, 657-666.

Cheng, E. W., & Hampson, I. (2008). Transfer of training: A review and new insights. *International Journal of Management Reviews, 10*, 327-341.

Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research, 53*, 445-459.

Clark, R. E. (1985). Confounding in educational computing research. *Journal of Educational Computing Research, 1*, 137-148.

Clark, R. E. (1994). Media will never influence learning. *Educational Technology Research and Development, 42*(2), 21-29.

Clark, R. E. (2001). *Learning from media: Arguments, analysis, and evidence*. Charlotte, NC: Information Age Publishing Inc.

Clark, R. E. (2007). Learning from serious games? Arguments, evidence, and research suggestions. *Educational Technology, 47*(3), 56-59.

Clark, R. E., Yates, K., Early, S., & Moulton, K. (2010). An analysis of the failure of electronic media and discovery-based learning: Evidence for the performance benefits of guided training methods. In K. H. Silber & W. R. Foshay (Eds.), *Handbook of improving performance in the workplace* (pp. 263–297). San Francisco, CA: Pfeiffer.

Coller, B. D., & Scott, M. J. (2009). Effectiveness of using a video game to teach a course in mechanical engineering. *Computers & Education, 53*, 900–912.

Coller, B. D., & Shernoff, D. J. (2009). Video game-based education in mechanical engineering: A look at student engagement. *International Journal of Engineering Education, 25*, 308-325.

Coller, B. D., Shernoff, D. J., & Strati, A. D. (2011). Measuring engagement as students learn dynamic systems and control with a video game. *Advances in Engineering Education, 2*(3), 1-32.

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*, 661-686.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings.* Chicago, IL: Rand McNally.

Cook, T. D., Campbell, D. T., & Perracchio, L. (1990). Quasi experimentation. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 1, pp. 491-576). Palo Alto, CA: Consulting Psychologists Press.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*, 98-104.

Crookall, D., & Thorngate, W. (2009). Acting, knowing, learning, simulating, gaming. *Simulation & Gaming, 40*, 8-26.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York, NY: Harper Perennial.

Curşeu, P. L. (2006). Emergent states in virtual teams: a complex adaptive systems perspective. *Journal of Information Technology, 21*, 249-261.

Darley, J. M., & Latané, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology, 8*, 377-383.

Darr, E. D., & Kurtzberg, T. R. (2000). An investigation of partner similarity dimensions on knowledge transfer. *Organizational Behavior and Human Decision Processes, 82*, 28–44.

Davey, A. (2009). *Statistical power analysis with missing data: A structural equation modeling approach*. New York, NY: Routledge.

de Freitas, S., & Jarvis, S. (2007). Serious games engaging training solutions: A research and development project for supporting training needs. *British Journal of Education Technology, 38*, 523–525.

De Freitas, S., & Routledge, H. (2013). Designing leadership and soft skills in educational games: The e-leadership and soft skills educational games design model (ELESS). *British Journal of Educational Technology, 44*, 951-968.

Dempsey, J. V., Rasmussen, K., & Lucassen, B. (1996). *Instructional gaming: Implications for instructional technology*. Paper presented at the annual meeting of the Association for Educational Communications and Technology. Nashville, TN.

Department of Defense. (1997). *DoD Modeling and Simulation* (DoD 5000.59-M). Alexandria, VA: Defense Modeling and Simulation Office.

Derouin-Jessen, R. E. (2008). *Game on: The impact of game features in computer-based training* (Unpublished doctoral dissertation). University of Central Florida, Orlando, FL.

DeVellis, R. F. (2012). *Scale development: Theory and applications* (3rd edition). Los Angeles: Sage. ISBN: 978-1412980449

Dewey, J. (1938). *Experience and education*. New York, NY: Kappa Delta Pi (Reprinted by Touchstone).

Dieker, L. A., Rodriguez, J. A., Lignugaris, B., Hynes, M. C., & Hughes, C. E. (2014). The potential of simulated environments in teacher education: Current and future possibilities. *Journal of the Teacher Education Division of the Council for Exceptional Children*, *37*, 21-33. doi:10.1177/0888406413512683

Dillenbourg, P. (1999). What do you mean by collaborative learning? In P. Dillenbourg (Ed.). *Collaborative-learning: Cognitive and Computational Approaches* (pp. 1-19). New York, NY: Elsevier.

Dillon, P. C., Graham, W. K., & Aidells, A. L. (1972). Brainstorming on a" hot" problem: Effects of training and practice on individual and group performance. *Journal of Applied Psychology*, *56*, 487-490.

Driskell, J. E., & Dwyer, J. D. (1984). Microcomputer videogame-based training. *Educational Technology, 24*(2), 11-16.

Ermi, L., & Mäyrä, F. (2005). Fundamental components of the gameplay experience: Analyzing immersion. *Worlds in Play: International Perspectives on Digital Games Research, 37*(2), 2-14.

Fandt, P. M. (1991). The relationship of accountability and interdependent behavior to enhancing team consequences. *Group & Organization Studies, 16*, 300-312.

Faria, A. J., Hutchinson, D., Wellington, W. J., & Gold, S. (2009). Developments in business gaming a review of the past 40 years. *Simulation & Gaming, 40*, 464-487.

Federal Trade Commission. (2016). *Lumosity to pay $2 million to settle FTC deceptive advertising charges for its "brain training" program: Company claimed program would sharpen performance in everyday life and protect against cognitive decline*. Retrieved from https://www.ftc.gov/news-events/press-releases/2016/01/lumosity-pay-2-million-settle-ftc-deceptive-advertising-charges

Findley, W. G. (1956). A rationale for evaluation of item discrimination statistics. *Educational and Psychological Measurement, 16*, 175-180.

Fletcher, J. D., & Tobias, S. (2006). *Using computer games and simulations for instruction: A research review*. Proceedings from The Society for Advanced Learning Technology Meeting, Orlando, FL.

Ford, J. K., Smith, E. M., Weissbein, D. A., Gully, S. M., & Salas, E. (1998). Relationships of goal orientation, metacognitive activity, and practice strategies with learning outcomes and transfer. *Journal of Applied Psychology, 83*, 218-233.

Frenzel, A. C., Pekrun, R., & Goetz, T. (2007). Perceived learning environment and students' emotional experiences: A multilevel analysis of mathematics classrooms. *Learning and Instruction, 17*, 478-493.

García-Carbonell, A., & Watts, F. (2009). Simulation and gaming methodology in language acquisition. In V. Guillén-Nieto, C. Marimón-Llorca, & C. Vargas-Sierra (Eds.), *Intercultural business communication and simulation and gaming methodology* (pp. 285–316). Bern, Switzerland: Peter Lang.

Gardner, B. S., & Korth, S. J. (1998). A framework for learning to work in teams. *Journal of Education for Business, 74*, 28-33.

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & Gaming, 33*, 441-467.

Gee, J. P. (2005). Learning by design: Good video games as learning machines. *E-Learning and Digital Media*, *2*, 5-16. doi:10.2304/elea.2005.2.1.5

Gibson, C. B., & Cohen, S. G. (Eds.). (2003). *Virtual teams that work: Creating conditions for virtual team effectiveness.* San Francisco, CA: Jossey-Bass.

Gittell, J. H. (2002). Coordinating mechanisms in care provider groups: Relational coordination as a mediator and input uncertainty as a moderator of performance effects. *Management Science, 48*, 1408-1426.

Goodman, D., Bradley, N. L., Paras, B., Williamson, I. J., & Bizzochi, J. (2006). Video gaming promotes concussion knowledge acquisition in youth hockey players. *Journal of Adolescence, 29*, 351-360.

Gredler, M. E. (1996). Educational games and simulations: A technology in search of a (research) paradigm. *Technology, 39*, 521-540.

Greenblat, C. S. (1981). Teaching with simulation games: A review of claims and evidence. In C. S. Greenblat & R. D. Duke (Eds.), *Principles and practices of gaming-simulation* (pp. 139-153). Beverly Hills, CA: Sage.

Greenblat, C. S., & Duke, R. D. (Eds.). (1981). *Principles and practices of gaming-simulation*. Beverly Hills, CA: Sage.

Grossman, R., Heyne, K., & Salas, E. (2014). Game- and simulation-based approaches to training. In K. Kraiger, J. Passmore, N. R. dos Santos, & S. Malvezzi (Eds.) *The Wiley Blackwell handbook of the psychology of training, development, and performance improvement* (pp. 205-224).

Guetzkow, H. (1959). A use of simulation in the study of inter-nation relations. *Systems Research and Behavioral Science, 4*, 183-191.

Guida, C. T. (n.d.). *Quintet*. Retrieved November 25, 2016 from https://www.quintet.us/

Guillen-Nieto, V., & Aleson-Carbonell, M. (2012). Serious games and learning effectiveness: The case of "It's a Deal!" *Computers & Education, 58*, 435-448.

Habgood, M. P. J. (2005, June). *Zombie division: Integration in digital learning games*. Proceedings from the Human Centered Technology Workshop, Brighton, UK.

Habgood, M. P. J., Ainsworth, S. E., & Benford, S. (2005). Endogenous fantasy and learning in digital games. *Simulation & Gaming, 36*, 483-498.

Haccoun, R. R., & Hamtiaux, T. (1994). Optimizing knowledge tests for inferring learning acquisition levels in single group training effectiveness designs: The internal referencing strategy. *Personnel Psychology, 47*, 593-604.

Hassenzahl, M. (2003). The thing and I: Understanding the relationship between user and product. In M. A. Blythe, A. F. Monk, K. Overbeeke, & P. C. Wright. (Eds.). *Funology: From usability to enjoyment.* The Netherlands: Kluwer Academic Publishers.

Hays, R. T. (2005). *The effectiveness of instructional games: A literature review and discussion* (No. NAWCTSD-TR-2005-004). Orlando, FL: Naval Air Warfare Center Training Systems Division.

Hinds, A. P. (1975). Separating the sirloin from the salami. *Training and Development Journal, 29*(3), 42-44.

Hinds, P. J., & Weisband, S. P. (2003). Knowledge sharing and shared understanding in virtual teams. In C. B. Gibson & S. G. Cohen (Eds.) *Virtual teams that work: Creating conditions for virtual team effectiveness*, (pp. 21-36). San Francisco, CA: Jossey-Bass.

Hoffman, B., & Nadelson, L. (2010). Motivational engagement and video gaming: A mixed methods study. *Educational Technology Research and Development, 58*, 245-270.

Howard, G. S. (1980). Response-shift bias: A problem in evaluating interventions with pre/post self-reports. *Evaluation Review, 4*, 5-41.

Hsu, E. (1989). Role-event gaming simulation in management education: A conceptual framework and review. *Simulation & Gaming, 20*, 409-38.

Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*, 424-453.

Iacovides, I., McAndrew, P., Scanlon, E., & Aczel, J. (2014). The gaming involvement and informal learning framework. *Simulation & Gaming, 45*, 611-626.

Jabbar, A. I. A., & Felicia, P. (2015). Gameplay engagement and learning in game-based learning a systematic review. *Review of Educational Research*, *85*, 740-779.

Jayakanthan, R. (2002). Application of computer games in the field of education. *The Electronic Library, 20*, 98–102.

Jeong, K. Y., & Bozkurt, I. (2014). Evaluating a project management simulation training exercise. *Simulation & Gaming*, *45*, 183-203.

Johnson, W. L. (2010). Serious use of a serious game for language learning. *International Journal of Artificial Intelligence in Education, 20*, 175-195.

Johnson, C. I., & Mayer, R. E. (2010). Adding the self-explanation principle to multimedia learning in a computer-based game-like environment. *Computers in Human Behavior, 26*, 1246–1252. doi:10.1016/ j.chb.2010.03.025

Jonker, C. M., van Riemsdijk, M. B., & Vermeulen, B. (2011). Shared mental models: A conceptual analysis. In M. De Vos, N. Fornara, J. V. Pitt, & G. Vouros (Eds.). Coordination, organizations, institutions, and norms in agent systems (pp. 132-151). Heidelberg, Germany: Springer. doi:10.1007/978-3-642-21268-0_8

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist, 39*, 341-350.

Kaptelinin, V., & Cole, M. (2002). Individual and collective activities in educational computer game playing. In T. Koschmann, R. P. Hall, & N. Miyake (Eds.), *Carrying forward the conversation* (pp. 297–316). Mahwah, NJ: Lawrence Erlbaum.

Kaye, J. J. (2007). Evaluating experience-focused HCI. In *Extended Abstract of CHI '07*. New York: NY, ACM Press.

Ke, F. (2008). A case study of computer gaming for math: Engaged learning from gameplay? *Computers & Education, 51*, 1609-1620.

Keller, J. M. (2006). *Development of two measures of learner motivation* (Unpublished doctoral dissertation). Florida State University, Tallahassee, FL.

Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist, 28*(2), 107-123.

Kickmeier-Rust, M. D., Peirce, N., Conlan, O., Schwarz, D., Verpoorten, D., & Albert, D. (2007). Immersive digital games: The interfaces for next-generation e-learning? In *International Conference on Universal Access in Human-Computer Interaction* (pp. 647-656). Heidelberg, Germany: Springer.

Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). *Evaluating training programs: The four levels*. San Francisco, CA: Berrett-Koehler.

Ko, D.-G., Kirsch, L. J., & King, W. R. (2005). Antecedents of knowledge transfer from consultants to clients in enterprise system implementations. *Management Information System Quarterly, 29*, 59-85.

Kobes, M., Helsloot, I., de Vries, B. & Post, J. (2010). Exit choice, (pre-)movement time and (pre-)evacuation behaviour in hotel fire evacuation – behavioural analysis and validation of the use of serious gaming in experimental research. *Procedia Engineering, 3*, 37–51.

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology, 126*, 349-362.

Koriat, A., Sheffer, L., & Ma'ayan, H. (2002). Comparing objective and subjective learning curves: Judgments of learning exhibit increased underconfidence with practice. *Journal of Experimental Psychology, 131*, 147-162.

Kozlowski, S. W., & Bell, B. S. (2013). Work groups and teams in organizations: Review update. In N. Schmitt & S. Highhouse (Eds.), *Handbook of psychology*: *Vol. 12. Industrial and organizational psychology* (2nd ed., pp. 412-469). Hoboken, NJ: Wiley.

Kozlowski, S. W., & Bell, B. S. (2008). Team learning, development, and adaptation. In V. I. Sessa & M. London (Eds.). *Work group learning: Understanding, improving & assessing how groups learn*, (pp. 15-44). New York, NY: Lawrence Erlbaum Associates.

Kozlowski, S. W., Chao, G. T., Grand, J. A., Braun, M. T., & Kuljanin, G. (2013). Advancing multilevel research design: Capturing the dynamics of emergence. *Organizational Research Methods, 16*, 581-615.

Kraiger, K. (2002). Decision-based evaluation. In K. Kraiger (Ed.) *Creating, implementing, and managing effective training and development: State-of-the-art lessons for practice* (pp. 331-375). San Francisco: CA, Jossey-Bass.

Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of Applied Psychology, 78*, 311-328.

Kraut, R., Fussell, S., Lerch, F., & Espinosa, A. (2005). Coordination in teams: Evidence from a simulated management game. *Human-Computer Interaction Institute.* Retrieved June 27, 2017 from http://repository.cmu.edu/hcii/102/

Landers, R. N. (2014). Developing a theory of gamified learning linking serious games and gamification of learning. *Simulation & Gaming, 45*, 752-768.

Landers, R. N. & Landers, A. K. (2015). An empirical test of the theory of gamified learning: The effect of leaderboards on time-on-task and academic performance. *Simulation & Gaming, 45*, 769-785. doi:10.1177/1046878114563662

Lankes, M., Bernhaupt, R., & Tscheligi, M. (2010). Evaluating user experience factors using experiments: Expressive artificial faces embedded in contexts. In R. Bernhaupt (Ed.). *Evaluating user experience in games* (pp. 165-183). London, England: Springer.

Laughlin, P. R., Bonner, B. L., & Miner, A. G. (2002). Groups perform better than the best individuals on letters-to-numbers problems. *Organizational Behavior and Human Decision Processes, 88*, 605-620.

Laughlin, P. R., Carey, H. R., & Kerr, N. L. (2008). Group-to-individual problem-solving transfer. *Group Processes & Intergroup Relations, 11*, 319-330.

Laughlin, P. R., Zander, M. L., Knievel, E. M., & Tan, T. K. (2003). Groups perform better than the best individuals on letters-to-numbers problems: Informative equations and effective strategies. *Journal of Personality and Social Psychology*, *85*, 684-694. http://dx.doi.org/10.1037/0022-3514.85.4.684

Leemkuil, H., de Jong, T., & Ootes, S. (2000). *Review of educational use of games and simulations* (No. IST-1999-13078). Enschede, The Netherlands: Knowledge Management Interactive Training System Project.

Lemay, P., & Maheux-Lessard, M. (2015). Investigating experiences and attitudes toward videogames using a semantic differential methodology. In R. Berhnhaupt (Ed.) *Evaluating user experience in games* (pp. 89-105). London, England: Springer International Publishing.

Lepper, M. R. (1985). Microcomputers in education: Motivational and social issues. *American Psychologist, 40*, 1-18.

Lepper, M. R., & Gilovich, T. (1982). Accentuating the positive: Eliciting generalized compliance from children through activity-oriented requests. *Journal of Personality and Social Psychology*, 42, 248-259. doi:10.1037/0022-3514.42.2.248

Leung, A. (2015). A conceptual model of information technology training leading to better outcomes. *International Journal of Business and Information, 1*, 74-95.

Levitan, E. P. (2010). Higher education administrators' perceptions of the use of simulation games for adult learners. *ProQuest Dissertation Publishing*, 3406153.

Li, M. C., & Tsai, C. C. (2013). Game-based learning in science education: A review of relevant research. *Journal of Science Education and Technology, 22*, 877–898.

Liang, D. W., Moreland, R., & Argote, L. (1995). Group versus individual training and group performance: The mediating role of transactive memory. *Personality and Social Psychology Bulletin*, *21*, 384-393.

Lingard, L., Whyte, S., Espin, S., Ross Baker, G., Orser, B., & Doran, D. (2006). Towards safer interprofessional communication: Constructing a model of "utility" from preoperative team briefings. *Journal of Interprofessional Care, 20*, 471-483.

Liu, C. C., Cheng, Y. B., & Huang, C. W. (2011). The effect of simulation games on the learning of computational problem solving. *Computers and Education, 57*, 1907-1918.

Long, L. K. (2005). *The role of trainee reactions in online training* (Unpublished doctoral dissertation). Kent State University, Kent, OH.

Loon, M., Evans, J., & Kerridge, C. (2015). Learning with a strategic management simulation game: A case study. *The International Journal of Management Education, 13*, 227-236.

Lord, F. M. (1952). The relation of the reliability of multiple-choice tests to the distribution of item difficulties. *Psychometrika, 17*, 181-194.

Lypaczewski, P. A., Jones, A. D., & Voorhees, J. W. (1987). Simulation of an advanced scout attack helicopter for crew station studies. Proceedings from I/ITSC '87: *The Interservice/Industry Training Systems Conference* (pp. 18-23), Salt Lake City, Utah.

Ma, W. W., & Yuen, A. H. (2011). Understanding online knowledge sharing: An interpersonal relationship perspective. *Computers & Education, 56*, 210-219. doi:10.1016/j.compedu.2010.08.004

Maciejovsky, B., Sutter, M., Budescu, D. V., & Bernau, P. (2013). Teams make you smarter: How exposure to teams improves individual decisions in probability and reasoning tasks. *Management Science, 59*, 1255-1270.

MacStewart, E. (2010). Stealth learning in online games. In S. de Freitas & P. Maharg (Eds.). *Digital games and learning* (pp. 107–128). New York, NY: Continuum Press.

Malone, T. W. (1981). What makes computer games fun? *Byte, 6*, 258–277.

Manninen, T. (2003). Interaction forms and communicative actions in multiplayer games. *The International Journal of Computer Game Research, 3*(1), 5–10.

Mautone, T., Spiker, V., & Karp, D. (2008). Using serious game technology to improve aircrew training. In proceedings of the *Interservice/Industry Training, Simulation & Education Conference (I/ITSEC)*, (paper 8183), 1-11.

May, D. R., Gilson, R. L., & Harter, L. M. (2004). The psychological conditions of meaningfulness, safety and availability and the engagement of the human spirit at work. *Journal of Occupational and Organizational Psychology, 77*, 11-37.

Mayer, R. E. (2011). Multimedia learning and games. In S. Tobias & J. D. Fletcher (Eds.), *Computer games and instruction* (pp. 281–305). Greenwich, CT: Information Age.

Mayer, R. E., & Johnson, C. I. (2010). Adding instructional features that promote learning in a game-like environment. *Journal of Educational Computing Research, 101*, 621–629.

Maynard, T., Mathieu, J., Gilson, L., & Sanchez, D. R. (2017). *Information elaboration: Unlocking the relationship between member familiarity and global virtual team effectiveness.* (Unpublished manuscript). Colorado State University.

Mayo, M. J. (2007). Games for science and engineering education. *Communications of the ACM, 50*, 31–35.

McCarthy, J., & Wright, P. (2004). *Technology as experience*. Cambridge, MA: MIT Press.

McDonald, R. P. (1999). Classical item analysis. In *Test Theory: A unified treatment* (pp. 231-247). New York, NY: Lawrence Erlbaum.

McDonald, R. P., & Ho, M. H. R. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods, 7*, 64-82.

McGrath, J. E. (1984). *Groups: Interaction and performance*. Englewood Cliffs, NJ: Prentice-Hall.

Megginson, L. C. (1959). The business game as a pedagogical technique. *The Southwestern Social Science Quarterly, 40,* 33-40.

Michael, D. & Chen, S. (2005). *Proof of learning: Assessment in serious games.* Retrieved on November 30, 2016 from http://www.gamasutra.com/view/feature/130843/proof_of_learning_assessment_in_.php

Moreland, R. L., & Myaskovsky, L. (2000). Exploring the performance benefits of group training: Transactive memory or improved communication? *Organizational Behavior and Human Decision Processes*, *82*, 117-133.

Morgan, J. (2013). Want to work here? Play this game first! *Forbes*. Retrieved from https://www.forbes.com/sites/jacobmorgan/2013/12/17/want-to-work-here-play-this-game-first/#73827d254249

Morgan, R. B., & Casper, W. J. (2000). Examining the factor structure of participant reactions to training: A multidimensional approach. *Human Resource Development Quarterly, 11*, 301-317.

Morrison, K. E. (2010). The multiplication game. *Mathematics Magazine, 83*, 100-110.

Muntean, C. I. (2011, October). Raising engagement in e-learning through gamification. *Proceedings from ICVL '11: 6th International Conference on Virtual Learning* (pp. 323-329), Cluj-Napoca, Romania.

Murphy, K. R. (2008). Explaining the weak relationship between job performance and ratings of job performance. *Industrial and Organizational Psychology, 1*(2), 148-160.

Muthén, L.K. & Muthén, B.O. (2012). *Mplus User's Guide*. Seventh Edition. Los Angeles, CA: Muthén & Muthén.

Nash, S. S. (2005). Learning objects, learning object repositories, and learning theory: Preliminary best practices for online courses. *Interdisciplinary Journal of Knowledge and Learning Objects, 1*, 217–228.

Navarro, E. O., Baker, A. & Van der Hoek, A. (2004). Teaching software engineering using simulation games. Proceedings of the ICSIE '04: *International Conference on Simulation in Education*. San Diego, CA.

O'Neil, H. F., & Perez, R. S. (Eds.). (2008). *Computer games and team and individual learning*. Amsterdam, the Netherlands: Elsevier.

O'Neil, H. F., Wainess, R., & Baker, E. L. (2005). Classification of learning outcomes: Evidence from the games literature. *Curriculum Journal, 16*, 455–474. doi:10.1080/09585170500384529

Oksanen, K., & Hämäläinen, R. (2014). Game mechanics in the design of a collaborative 3D serious game. *Simulation & Gaming, 45*, 255-278. doi:10.1177/1046878114530799

Pagulayan, R. J., Keeker, K., Wixon, D., Romero, R. L., & Fuller, T. (2002). User-centered design in games. In J. A. Jacko (Ed.). *The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications.* Boca Raton, FL: Taylor & Francis.

Pannese, L., & Carlesi, M. (2007). Games and learning come together to maximize effectiveness: The challenge of bridging the gap. *British Journal of Educational Technology, 38*, 438-454.

Papastergiou, M. (2009). Digital game-based learning in high school computer science education: Impact on educational effectiveness and student motivation. *Computers & Education, 52*, 1-12.

Patrick, J. (1992). *Training research and practice*. San Diego, CA: Academic Press.

Perrotta, C., Featherstone, G., Aston, H., & Houghton, E. (2013). *Game-based learning: Latest evidence and future directions*. Slough, England: NFER Research Programme: Innovation in Education.

Phillips, J. J., & Phillips, P. P. (2016). *Handbook of training evaluation and measurement methods*. New York, NY: Routledge.

Plass, J. L., O'Keefe, P. A., Homer, B. D., Case, J., Hayward, E. O., Stein, M., & Perlin, K. (2013). The impact of individual, competitive, and collaborative mathematics game play on learning, performance, and motivation. *Journal of Educational Psychology, 105*, 1050-1066.

Podsakoff, P. M., & Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management, 12*, 531-544.

Ponce, E. R., Franchak, S. J., Billings, R. S., & O'Reilly, P. A. (1981). *Evaluating student satisfaction: Measurement of training and job satisfaction of former vocational education students.* Columbus, OH: National Center for Research in Vocational Education.

Popper, K. R. (1994). *Knowledge and the body-mind problem: In defense of interaction*. M. A. Notturno (Ed.). London, England: Routledge.

Pratt, M. A., & Hahn, S. (2015). Effects of simulation on student satisfaction with a capstone course. *Journal of Hospitality & Tourism Education, 27*, 39-46.

Preece, J., Rogers, Y., & Sharp, H. (2002). *Interaction design: Beyond human computer Interaction.* Danvers, MA: John Wiley & Sons.

Prensky, M. (2001). Fun, play and games: What makes games engaging. *Digital Game-Based Learning, 5*, 1-05.

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education, 93*, 223-231.

Quińones, M. A., Ford, J. K., & Teachout, M. S. (1995). The relationship between work experience and job performance: A conceptual and meta‐analytic review. *Personnel Psychology, 48*, 887-910.

Ramos, C., & Yudko, E. (2008). “Hits” (not “discussion posts”) predict student success in online courses: A double cross-validation study. *Computers & Education, 50*, 1174–1182.

Randel, J., Morris, B., Wetzel, C. D., & Whitehall, B. (1992). The effectiveness of games for educational purposes: A review of recent research. *Simulation & Gaming, 23*, 261–276.

Raykov, T., & Marcoulides, G. A. (2011). *Introduction to psychometric theory*. New York: Taylor & Francis. ISBN: 978-0415878227

Rhodes, M. G., & Tauber, S. K. (2011). The influence of delaying judgments of learning on metacognitive accuracy: A meta-analytic review. *Psychological Bulletin, 137*, 131-148.

Ricci, K., Salas, E., & Cannon-Bowers, J. A. (1996). Do computer-based games facilitate knowledge acquisition and retention? *Military Psychology, 8*, 295–307.

Rich, B. L., Lepine, J. A., & Crawford, E. R. (2010). Job engagement: Antecedents and effects on job performance. *Academy of Management Journal, 53*, 617-635.

Rieber, L. P. (1996). Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games. *Educational Technology, Research, and Development, 44*, 43–58.

Rising, B. (2009). Business simulations as a vehicle for language acquisition. In V. Guillén-Nieto, C. Marimón-Llorca, & C. Vargas-Sierra (Eds.), *Intercultural business communication and simulation and gaming methodology* (pp. 317–353). Bern, Switzerland: Peter Lang.

Rosen, E. F., & Petty, L. C. (1992). Computer-aided instruction in a physiological psychology course. *Behavior Research Methods, Instruments & Computers, 24*, 169–171.

Sackett, P. R., & Mullen, E. J. (1993). Beyond formal experimental design: Towards an expanded view of the training evaluation process. *Personnel Psychology, 46*, 613-627.

Salas, E., DiazGranados, D., Klein, C., Burke, C. S., Stagl, K. C., Goodwin, G. F., & Halpin, S. M. (2008). Does team training improve team performance? A meta-analysis. *Human Factors*, *50*, 903-933.

Salas, E., Shuffler, M. L., Thayer, A. L., Bedwell, W. L., & Lazzara, E. H. (2015). Understanding and improving teamwork in organizations: A scientifically based practical guide. *Human Resource Management, 54*, 599-622.

Sanchez, D. R. 2017). *Videogame experience scale development: Designing and validating a scale with implications for game-based learning.* (Unpublished document). Colorado State University.

Sanchez, D. R. & Gibbons, A. (2017). *Videogame-based training simulations: The impact and interaction of videogame characteristics on learning outcomes.* (Unpublished document). Colorado State University.

Sanders, J. (2015). By the numbers: 10 stats on the growth of gamification. Retrieved on April 12, 2017 from http://www.gamesandlearning.org/2015/04/27/by-the-numbers-10-stats-on-the-growth-of-gamification/

Sawyer, T., Sierocka-Castaneda, A., Chan, D., Berg, B., Lustik, M., & Thompson, M. (2012). The effectiveness of video-assisted debriefing versus oral debriefing alone at improving neonatal resuscitation performance: a randomized trial. *Simulation in Healthcare, 7*, 213-221. doi:10.1097/SIH.0b013e3182578eae

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3*, 207-217.

Schmitt, N., & Stuits, D. M. (1985). Factors defined by negatively keyed items: The result of careless respondents? *Applied Psychological Measurement, 9*, 367-373.

Shernoff, D. J., Csikszentmihalyi, M., Shneider, B., & Shernoff, E. (2003). Student engagement in high school classrooms from the perspective of flow theory. *School Psychology Quarterly, 18*, 158-176. doi:10.1521/scpq.18.2.158.21860

Shute, V. J., Ventura, M., & Ke, F. (2015). The power of play: The effects of Portal 2 and Lumosity on cognitive and noncognitive skills. *Computers & Education, 80*, 58-67.

Singer, M. J., & Knerr, B. W. (2010). *Evaluation of a game-based simulation during distributed exercises.* Orlando, FL: Army Research Institute for the Behavioral and Social Sciences.

Sitzmann, T., Ely, K., Brown, K. G., & Bauer, K. N. (2010). Self-assessment of knowledge: A cognitive learning or affective measure? *Academy of Management Learning & Education, 9*, 169-191.

Span, P. (2016). After the F.T.C.'s failing grade for Lumosity. *New York Times*, January 19, 2016, (page D5). Retrieved from https://www.nytimes.com/2016/01/19/health/ftcs-lumosity-penalty-doesnt-end-brain-training-debate.html?_r=0

Stone, A. A., Bachrach, C. A., Jobe, J. B., Kurtzman, H. S., & Cain, V. S. (Eds.). (1999). The science of self-report: Implications for research and practice. Mahwah, NJ: Lawrence Erlbaum.

Tan, J. A., Hall, R. J., & Boyce, C. (2003). The role of employee reactions in predicting training effectiveness. *Human Resource Development Quarterly, 14*, 397-411.

Tannenbaum, S. I., & Yukl, G. (1992). Training and development in work organizations. *Annual Review of Psychology, 43*, 399-441.

Taut, S. (2007). Studying self-evaluation capacity building in a large international development organization. *American Journal of Evaluation, 28*, 45-59.

Thomas, R., Cahill, J., & Santilli, L. (1997). Using an interactive computer game to increase skill and self-efficacy regarding safer sex. *Health Education and Behavior, 24*, 71-86.

Tobias, S., & Fletcher, J. D. (Eds.). (2011). *Computer games and instruction*. Greenwich, CT: Information Age.

Trepte, S., & Reinecke, L. (2011). The pleasures of success: Game-related efficacy experiences as a mediator between player performance and game enjoyment. *Cyberpsychology, Behavior, and Social Networking, 14*, 555-557.

Vahed, A. (2008). The tooth morphology board game: an innovative strategy in tutoring dental technology learners in combating rote learning. Proceedings from ECGBL '08: *The European Conference on Games-Based Learning*, Barcelona, Spain.

Vandercruysse, S., Vandewaetere, M., & Clarebout, G. (2012). Game based learning: A review on the effectiveness of educational games. In M. M. Cruz-Cunha (Ed.), *Handbook of research on serious games as educational, business, and research tools* (pp. 628-647). Hershey, PA: IGI Global.

Videogame Historian. (n.d.) Retrieved from https://videogamehistorian.wordpress.com/2014/01/22/the-priesthood-at-play-computer-games-in-the-1950s/

Virtual Heroes. (2007). *Zero Hour: America's medic.* Retrieved from www.virtualheroes.biz/VHFree/Content/ReferenceMaterials/RefMaterialZeroHour.zip

Vogel, J. F., Vogel, D. S., Cannon-Bowers, J., Bowers, C. A., Muse, K., & Wright, M. (2006). Computer gaming and interactive simulations for learning: A meta-analysis. *Journal of Educational Computing Research, 34*, 229–243.

Von Der Pütten, A. M., Klatt, J., Ten Broeke, S., McCall, R., Krämer, N. C., Wetzel, R., ... & Klatt, J. (2012). Subjective and behavioral presence measurement and interactivity in the collaborative augmented reality game TimeWarp. *Interacting with Computers, 24*, 317-325. http://dx.doi.org/10.1016/j.intcom.2012.03.004

Walker, S. L., & Fraser, B. J. (2005). Development and validation of an instrument for assessing distance education learning environments in higher education: The Distance Education Learning Environments Survey (DELES). *Learning Environments Research, 8*, 289-308.

Warr, P., Allan, C., & Birdi, K. (1999). Predicting three levels of training outcome. *Journal of Occupational and Organizational Psychology, 72*, 351-375.

Warr, P., & Bunce, D. (1995). Trainee characteristics and the outcomes of open learning. *Personnel Psychology, 48*, 347-375.

Wilson, K. A., Bedwell, W. L., Lazzara, E. H., Salas, E., Burke, C. S., Estock, J. L., . . . & Conkey, C. (2009). Relationships between game attributes and learning outcomes: Review and research proposals. *Simulation & Gaming, 40*, 217-266.

Wong, W. L., Shen, C., Nocera, L., Carriazo, E., Tang, F., Bugga, Sh., . . . Ritterfeld, U. (2007). Serious video game effectiveness. Proceedings from *ACE'07: The International Conference on Advances in Computer Entertainment Technology*, Salzburg, Austria.

Wouters, P., Van Nimwegen, C., Van Oostendorp, H., & Van Der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology, 105*, 249-265. doi:10.1037/a0031311

Wrzesien, M., & Raya, M. A. (2010). Learning in serious virtual worlds: Evaluation of learning effectiveness and appeal to students in the E-Junior project. *Computers & Education, 55*, 178–187.

Zwikael, O., & Gonen, A. (2007). Project Execution Game (PEG): Training towards managing unexpected events. *Journal of European Industrial Training, 31*, 495-512.

APPENDICES

Appendix A - Survey 1

Your responses to this survey are anonymous and cannot be linked to your identity in any way.

Participant # ___________________________

Section 1: Knowledge Test

The following questions relate to the computer-based training you just completed. This section is scored. It is important for our study that you do your best to provide the correct answers for each question.

(Declarative Knowledge – Pretest)
Which of the following is a station that is available for you to select when joining/rejoining the ship? (1a7)

- ☐ Science
- ☐ Captain
- ☐ Guide
- ☐ Mechanic

During the game, when can the Engineer switch to a new station? (2a3)

- ☐ Anytime a station is unoccupied
- ☐ Anytime the Captain "Allows" the switch
- ☐ Anytime during the game as long as a repair isn't in progress
- ☐ Anytime the ship is in a neutral "uncontested" zone

Which of the enemy ships has the worst defenses and is easiest to destroy? (3a3)

- ☐ Fighter
- ☐ Cruiser
- ☐ Drone
- ☐ Station

Which station is responsible for aiming the front of the ship towards the enemy ship being targeted? (4a2)

- ☐ Helm
- ☐ Tactical
- ☐ Engineering
- ☐ Captain

Which of the following keys would roll the ship right? (4a9)

- ☐ E
- ☐ R
- ☐ X
- ☐ >

Which station is responsible for communicating to the other team members, which ships are attacking? (4b2)

☐ Tactical
☐ Helm
☐ Captain
☐ Engineering

Which of the following has the longest cool down time? (4b9)
☐ Torpedoes
☐ Beam Lasers
☐ Boosters
☐ Thrusters

In the image below clicking 'range' will take the boost off which of the following? (4c11NEW)


☐ thrust
☐ turn
☐ damage
☐ cooldown

Which station does the Tactical station need to communicate with to boost their cooldown? (4c14)
☐ Engineering
☐ Helm
☐ Science
☐ Captain

Which is the only station that can hack an enemy ship to temporarily disable their Weapons, Shields, or Engines? (4d5)
☐ Science
☐ Engineering
☐ Tactical
☐ Captain

Which station can use a tractor beam to push or pull another ship? (4d7)
☐ Science
☐ Engineering

☐ Tactical
☐ Helm

Which race has the slowest ship? (5a5)
☐ Davul
☐ Zurna
☐ Santoor
☐ Hive

Which race rivals the Santoor? (5a12)
☐ Udu
☐ Davul
☐ Hive
☐ Zurna

Which race has the capabilities of using a phase shift? (5a18)
☐ Udu
☐ Davul
☐ Zurna
☐ Santoor

Section 2: Demographic Information

What is your current age in years (e.g., 18)? _____

How do you identify?
☐ Male
☐ Female
☐ Other ___________________________

Which best describes your ethnicity? (select one)
☐ Caucasian
☐ Black
☐ Hispanic
☐ Asian
☐ Native American
☐ Hawaiian / Pacific Islander
☐ Other / Mixed Race (please explain) ___________________________

(Game Experience)
SCALE: Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, Strongly Agree

The following questions relate to your experience **playing videogames**.
_____ I currently play videogames several hours every week.
_____ In the past I have spent several hours in one week playing videogames.

_____ I have spent a lot of time reading videogame magazines or websites to find tips to improve my gaming skills
_____ I have a high level of confidence with playing videogames in general.
_____ I have a high level of skill at playing videogames in general.
_____ Playing videogames is some of the most fun I have ever had in my life.
_____ I consider myself to have a lot of videogame experience.
_____ I would call myself a "serious gamer".
_____ I would call myself a "casual gamer".
_____ I am scared of videogames. (R)
_____ I don't enjoy playing videogames at all. (R)
_____ I have complete control over whether or not I do well in a videogame.
_____ I am confident that I can perform well in a videogame.
_____ I believe I have the ability to play videogames well.

Thank you for completing the survey.

Please let the researcher know you have finished the survey. They will provide you with instructions for what to do next.

Appendix B - Survey 2

Your responses to this survey are anonymous and cannot be linked to your identity in any way.

Participant # ____________________________

Section 1: Knowledge Test

The following questions relate to the computer-based training you completed earlier. This section is scored. It is important for our study that you do your best to provide the correct answers for each question.

(Declarative Knowledge – Posttest)
Which of the following is a station that is available for you to select when joining/rejoining the ship? (1a6)

☐ Engineering
☐ Captain
☐ Pilot
☐ Viewscreen

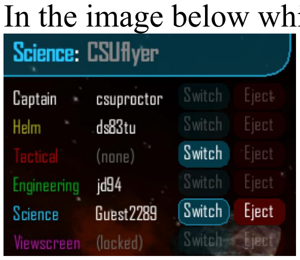
Which of the following is a station that is available for you to select when joining/rejoining the ship? (1a7)

☐ Science
☐ Captain
☐ Guide
☐ Mechanic

During the game, when can the Engineer switch to a new station? (2a3)

☐ Anytime a station is unoccupied
☐ Anytime the Captain "Allows" the switch
☐ Anytime during the game as long as a repair isn't in progress
☐ Anytime the ship is in a neutral "uncontested" zone

In the image below which station are you currently **occupying**? (2a8)



☐ Science
☐ Helm
☐ Engineering
☐ Tactical

Which of the enemy ships has the worst defenses and is easiest to destroy? (3a3)

☐ Fighter
☐ Cruiser
☐ Drone
☐ Station

When a ship is cloaked you can do which of the following actions? (3a5)
☐ Disrupt Cloak
☐ Scan
☐ View on screen
☐ Drain Shields

Which station is responsible for aiming the front of the ship towards the enemy ship being targeted? (4a2)
☐ Helm
☐ Tactical
☐ Engineering
☐ Captain

What action is taken to increase the speed of the ship? (4a3)
☐ Scroll up
☐ Zoom out
☐ Press the up arrow
☐ Press the space bar

Which station is responsible for communicating to the other team members which ships are attacking? (4b2)
☐ Tactical
☐ Helm
☐ Captain
☐ Engineering

Which of the following has the longest cool down time? (4b9)
☐ Torpedoes
☐ Beam Lasers
☐ Boosters
☐ Thrusters

Which station is responsible for communicating to the other team members the current health of the ship? (4c3)
☐ Engineering
☐ Helm
☐ Captain
☐ Science

Which station can teleport the ship across large distances? (4c6)
☐ Engineering

☐ Helm
☐ Captain
☐ Tactical

Which station is responsible for communicating to the other stations the health of an enemy ship? (4d3)
☐ Science
☐ Engineering
☐ Tactical
☐ Captain

Which is the only station that can hack an enemy ship to temporarily disable their Weapons, Shields, or Engines? (4d5)
☐ Science
☐ Engineering
☐ Tactical
☐ Captain

Which is the only station that can bring the ship to an immediate halt? (4a5)
☐ Helm
☐ Captain
☐ Engineering
☐ Science

Which of the following keys would roll the ship right? (4a9)
☐ E
☐ R
☐ X
☐ >

Which station <u>needs</u> to know the frequency of an attacking ship? (4b12)
☐ Tactical
☐ Helm
☐ Engineering
☐ Science

What do the red letters in the top right corner represent? (4b15)


☐ Hull
☐ Helm
☐ Hit Points
☐ Healing

In the image below clicking 'range' will take the boost off which of the following? (4c11NEW)

- ☐ thrust
- ☐ turn
- ☐ damage
- ☐ cooldown

Which station does the Tactical station need to communicate with to boost their cooldown? (4c14)

- ☐ Engineering
- ☐ Helm
- ☐ Science
- ☐ Captain

Which station can use a tractor beam to push or pull another ship? (4d7)

- ☐ Science
- ☐ Engineering
- ☐ Tactical
- ☐ Helm

Which of the following must happen before a Scan action can be done? (4d9)

- ☐ The enemy must be within range
- ☐ The scan action must be boosted
- ☐ The enemy ship must be in view
- ☐ The ship must not be under attack

Which of the following races represents their fleet with ships colored yellow? (5a4)

- ☐ Zurna
- ☐ Udu
- ☐ Santoor
- ☐ Davul

Which race has blue ships? (5a11)

- ☐ Udu
- ☐ Davul
- ☐ Zurna
- ☐ Hive

Which race rivals the Santoor? (5a12)

- ☐ Udu

□ Davul
□ Hive
□ Zurna

Which race has the capabilities of using a phase shift? (5a18)
□ Udu
□ Davul
□ Zurna
□ Santoor

Section 2. Training Debriefing

In this section please explain your understanding of the activity you participated in.

(Debriefing Emotion)
What type of emotional responses did you have to the activity? ____________________________

(Debriefing Group/Individual)
How do you think working **with a group/or individually** impacted your experience in the activity? ____________________________

All of the following measures use a 5-point scale
Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, Strongly Agree

Section 3: Activity Response

The following questions relate to your **reactions and experiences in the activity**.

(Psychological Meaning)
_____ My time was well spent in the activity.
_____ Doing the activity was a meaningful experience.
_____ My time doing the activity was **NOT** well spent. (R)
_____ Doing the activity was worthwhile.
_____ The activity was important to me.

(Perceived Enjoyment)
_____ I was very satisfied with this activity.
_____ I had a very positive experience in this activity.
_____ I really enjoyed this activity.
_____ This activity was very fun.
_____ I would enjoy doing this activity again.
_____ This activity was a waste of my time. (R)

(Motivation)
_____ The activity was interesting to me.

______ The activity was dry and unappealing. (R)
______ This activity stimulated my curiosity.
______ The activity was repetitive. (R)
______ The variety of tasks in the activity helped keep my attention.
______ The activity was boring. (R)

(Emotional Engagement)
______ I was enthusiastic in this activity. (Emotion)
______ I felt energetic in this activity. (Emotion)
______ I was interested in this activity. (Emotion)
______ I felt proud of my work in this activity. (Emotion)
______ I felt positive about this activity. (Emotion)
______ I felt excited about this activity. (Emotion)

(Physical Engagement)
______ I worked with intensity in the activity. (Physical)
______ I exerted my full effort in the activity. (Physical)
______ I devoted a lot of energy to the activity. (Physical)
______ I tried my hardest to perform well in the activity. (Physical)
______ I strove as hard as I could to complete the activity. (Physical)
______ I exerted a lot of energy in this activity. (Physical)

(Cognitive Engagement)
______ During this activity, my mind was focused on the task at hand. (Cognitive)
______ During this activity, I paid a lot of attention to the task at hand. (Cognitive)
______ During this activity, I focused a great deal of attention on the task at hand. (Cognitive)
______ During this activity, I was absorbed by the task at hand. (Cognitive)
______ During this activity, I concentrated on the task at hand. (Cognitive)
______ During the activity, I devote a lot of attention to the task at hand. (Cognitive)

(Perceived Value)
______ It is clear to me how the content of this activity is related to things I already know.
______ There were aspects of this activity that could be important to some people.
______ The content of this activity is relevant to my interests.
______ The content of this activity is worth knowing.
______ The content of this activity was not relevant to my needs. (R)
______ The content of this activity will be useful to me.

(Active Learning)
______ I explored my own strategies in the activity.
______ In the activity I sought my own answers.
______ In the activity I solved my own problems.

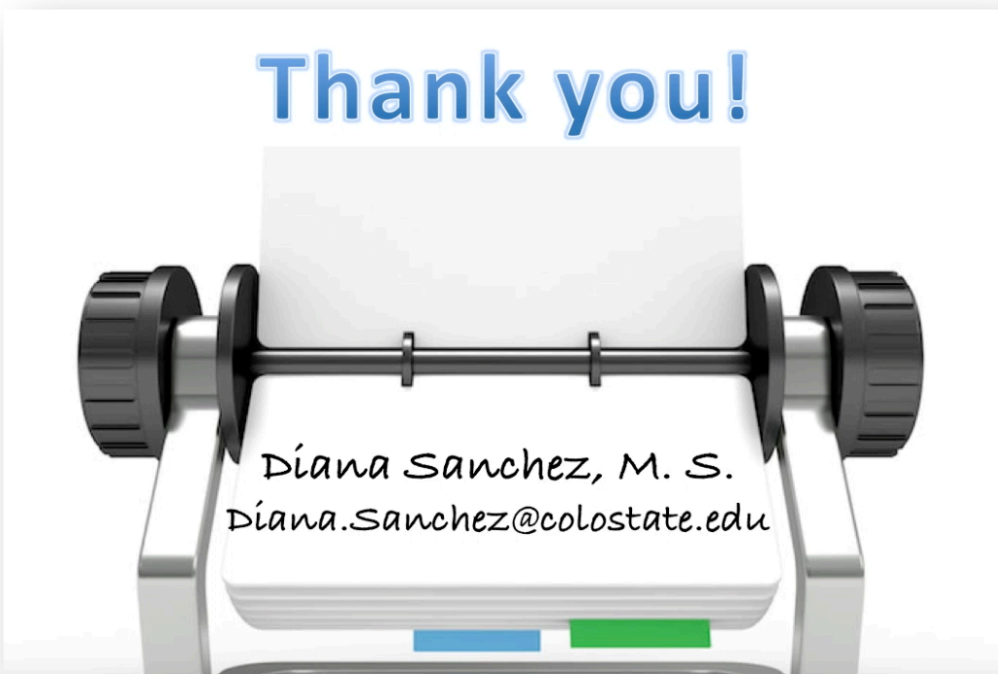
Thank you for completing this survey.

Click the link below. You will be redirected to a short video describing the purpose of our study. When the video is done please let the researcher know and they will provide you with instructions for what to do next.

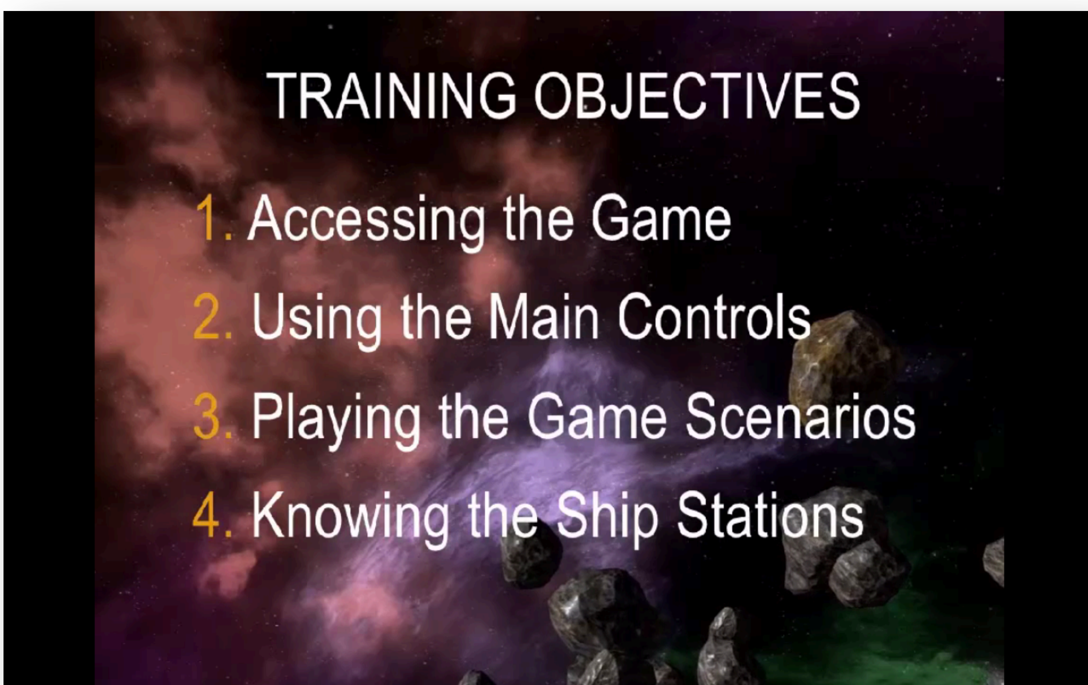
https://youtu.be/jEbg8hdDPog

Appendix C - Study Debriefing Screenshots

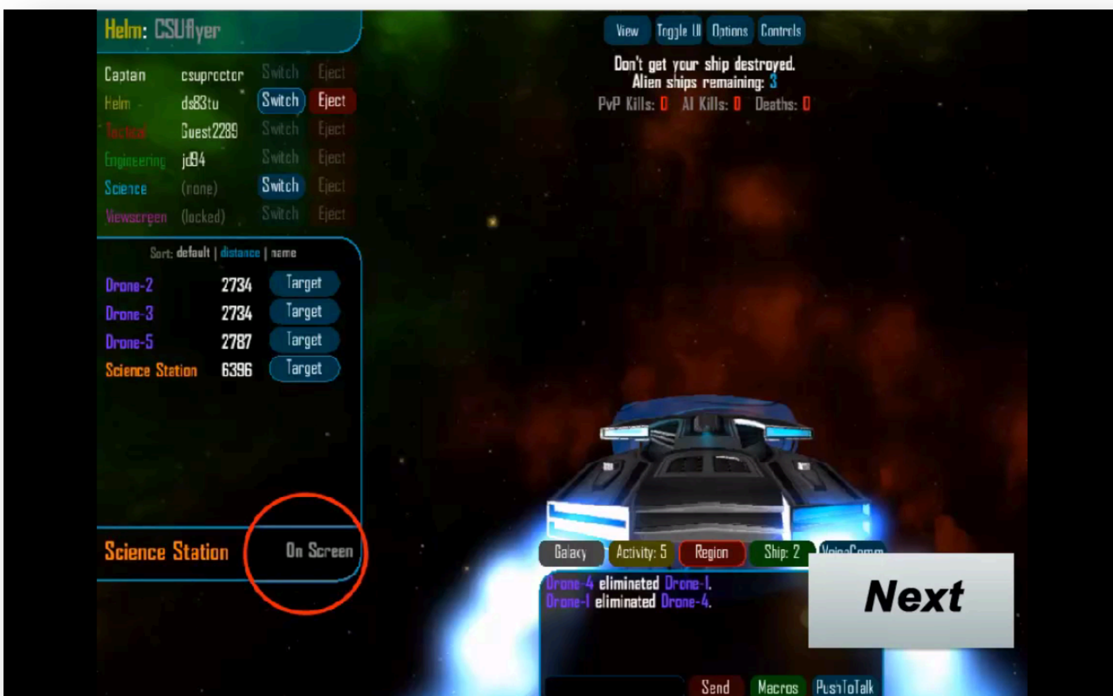

CONFIDENTIAL
CONTRACT
John Hancock


Thank you!
Diana Sanchez, M. S.
Diana.Sanchez@colostate.edu

Appendix D - Computer-Based Training Screenshots




Network Setup
Options
Controls
Feedback
Create a Ship
What is the name of your ship?
Such as: USS Second Prize
Next >
Password to join your ship?
Keeps strangers off your ship (optional)
< Back
Join a Ship
console: all
friends: all
Ship
Scenario
Captain
Helm
Tactical
Engineering
Science
ViewScreen
{P} CSUFlyer
Swarm of Nine
ds83tu
Join
Join
Join
Join
(locked)
Galaxy
Activity
Region
Ship
VoiceComm
News
BumFluff launched the SS BadLuck (Human Quartet).
BumFluff launched the SS BadLuck (Santoor Duet).
Nabrotar launched the USSR Hammer (Zurna Quartet).


Helm: CSUflyer
View
Toggle UI
Options
Controls
Captain csuprector Switch Eject
Helm ds83tu Switch Eject
Tactical Guest2289 Switch Eject
Engineering jd94 Switch Eject
Science (none) Switch Eject
Viewscreen (locked) Switch Eject
Don't get your ship destroyed.
Alien ships remaining: 3
PvP Kills: 0 AI Kills: 0 Deaths: 0
Sort: default | distance | name
Drone-2 2734 Target
Drone-3 2734 Target
Drone-5 2787 Target
Science Station 6396 Target
Science Station
On Screen
Galaxy
Activity: 5
Region
Ship: 2
Drone-4 eliminated Drone-1.
Drone-1 eliminated Drone-4.
Next
Send
Macros
PushToTalk

Appendix E - Game Script

- Start 20 minute timer
  - *You have 20 minutes to complete this scenario. Good luck!*

- Begin Each Scenario
  - Launch Ship
  - *You will see the ship FLYER. Select a station, click JOIN and enter the password.*
  - Read scenario description
  - EXAMPLE: *In this scenario locate the transport ship and escort him to the rendezvous point within the given time limit.*
  - When scenario is complete
  - *Great job. Click MENU and wait while I record your score.*

- After 20 minutes
  - The activity is now over. I'll end the game now. Please wait by your computer for me to come by and record your score.