PERSONNEL PSYCHOLOGY
THE STUDY OF PEOPLE AT WORK

PERSONNEL PSYCHOLOGY
2011, 64, 489–528

A META-ANALYTIC EXAMINATION OF THE INSTRUCTIONAL EFFECTIVENESS OF COMPUTER-BASED SIMULATION GAMES

TRACI SITZMANN
University of Colorado Denver

Interactive cognitive complexity theory suggests that simulation games are more effective than other instructional methods because they simultaneously engage trainees' affective and cognitive processes (Tennyson & Jorczak, 2008). Meta-analytic techniques were used to examine the instructional effectiveness of computer-based simulation games relative to a comparison group ($k = 65$, $N = 6{,}476$). Consistent with theory, post-training self-efficacy was 20% higher, declarative knowledge was 11% higher, procedural knowledge was 14% higher, and retention was 9% higher for trainees taught with simulation games, relative to a comparison group. However, the results provide strong evidence of publication bias in simulation games research. Characteristics of simulation games and the instructional context also moderated the effectiveness of simulation games. Trainees learned more, relative to a comparison group, when simulation games conveyed course material actively rather than passively, trainees could access the simulation game as many times as desired, and the simulation game was a supplement to other instructional methods rather than stand-alone instruction. However, trainees learned less from simulation games than comparison instructional methods when the instruction the comparison group received as a substitute for the simulation game actively engaged them in the learning experience.

Video games such as Pac-Man (developed by Namco, 1980) and Pong (manufactured by Atari Corporation, 1972) are icons of popular culture from the late 20th century. These video games utilized simplistic graphics and entertained millions of players. As video games increased in popularity, game developers realized the potential of capitalizing on the entertainment value of games and teaching instructional content during game play in order to advance into the education arena. As a result, popular computer games such as *Where in the World is Carmen Sandiego?* (produced by Brøderbund Software in 1985) and *Oregon Trail* (produced by MECC in 1974) were developed to teach geography and life on the

I would like to thank Katherine Ely for her assistance coding reports included in the meta-analysis and commenting on an earlier version of the manuscript.

Correspondence and requests for reprints should be addressed to Traci Sitzmann, University of Colorado Denver, School of Business, PO Box 173364, Campus Box 165, Denver, CO 80217-3364; Traci.Sitzmann@ucdenver.edu.

American frontier (CNET Networks Entertainment, 2004). Recently, the term “serious games” was coined to refer to simulation games designed to address more complicated and thought-provoking issues such as genocide, military combat, civil court procedures, and training first responders (Schollmeyer, 2006).

Simulation games refer to instruction delivered via personal computer that immerses trainees in a decision-making exercise in an artificial environment in order to learn the consequences of their decisions. Simulation games are intrinsically motivating (Malone, 1981), and people report experiencing a loss of time when playing their favorite games (Wood, Griffiths, & Parke, 2007). Furthermore, they are widely popular in the United States—approximately 40% of adults play video games (Slagle, 2006). The ultimate goal for training professionals is to harness the motivational capacity of video games to enhance employees’ work-related knowledge and skills (Garris, Ahlers, & Driskell, 2002).

The goal of this study is to statistically summarize the literature on the instructional effectiveness of computer-based simulation games for teaching work-related knowledge and skills. This will provide insight as to whether game play is effective for improving adults’ work competencies. Furthermore, it will clarify the characteristics of simulation games that are essential for maximizing learning outcomes.

Empirical and anecdotal evidence suggest that computer-based simulation games are effective for enhancing employees’ skill sets. Cold Stone Creamery developed a simulation game to teach customer service and portion control in a virtual Cold Stone store (Jana, 2006). Players race against the clock to serve customers in a timely fashion while maximizing the company’s profit by avoiding wasting too much ice cream. The first week the simulation game was available more than 8,000 employees—representing 30% of the workforce—voluntarily downloaded the simulation game. Corporate trainers believe that the entertainment value will incline employees to continuously play the simulation game while simultaneously teaching them retail sales, technical, and managerial skills. Canon Inc. utilizes a simulation game to teach copy machine repairs. Players must drag and drop parts into the right spot on a copier and, similar to the board game *Operation*, a light flashes and a buzzer sounds when a mistake is made. Employees who played the game obtained training assessment scores 5% to 8% higher than those trained with older techniques, such as manuals (Jana, 2006).

In 2003, the corporate simulation-based training industry was spending between $623 and $712 million globally (Summers, 2004). However, there is dissention among researchers and practitioners on the effectiveness of simulation games (O’Neil, Wainess, & Baker, 2005; Prensky, 2001; Randel, Morris, Wetzel, & Whitehill, 1992; Tobias & Fletcher, 2007;

Vogel et al., 2006). Others have noted that there are not clear guidelines on the features of simulation games that enhance learning (Bell, Kanar, & Kozlowski, 2008; Federation of American Scientists, 2006; Garris et al., 2002; Tennyson & Jorczak, 2008).

The goal of this study is to address the debate regarding whether simulation games enhance work-related knowledge and skills as well as to determine which simulation game design features enhance learning. I conducted a meta-analysis of 65 independent samples and data from more than 6,000 trainees in order to examine the instructional effectiveness of computer-based simulation games. Trainees taught with simulation games were evaluated against a comparison group on key affective and cognitive training outcomes. The comparison group differed across studies and ranged from a no-training control condition to trainees who received alternative instructional methods as a substitute for the simulation game. I also examined trends across studies to determine if the effectiveness of simulation games, relative to the comparison group, differed based on features of simulation games and the comparison group, characteristics of the instructional context, and methodological factors.

There have been three previous quantitative reviews of the effectiveness of simulation games (Lee, 1999; Randel et al., 1992; Vogel et al., 2006). This meta-analysis expands on these reviews in several ways. First, they all utilized a combination of child and adult samples, precluding a precise estimate of the effectiveness of this instructional method for teaching adults work-related skills. Second, Randel et al. focused strictly on cognitive learning outcomes, whereas Lee and Vogel et al. examined both cognitive learning and attitudes toward training. This meta-analysis examined several training outcomes (i.e., self-efficacy, declarative knowledge, procedural knowledge, and retention) and expanded on the moderators that have been examined—including both theoretical and methodological moderators. Specifically, five theoretical moderators (entertainment value, simulation game instruction was active or passive, unlimited access to the simulation game, whether the simulation game was the sole instructional method, and whether the instructional methods used to teach the comparison group were active or passive) and four methodological moderators (random assignment to experimental conditions, rigor of the study design, publication status, and year of the publication, dissertation, or presentation) were examined. Third, the most recent studies in previous reviews were published in 2003. This review includes studies from as far back as 1976 to as recent as 2009 in order to examine whether the instructional effectiveness of simulation games has evolved over time. Fourth, Vogel et al. and Lee included one and two unpublished studies in their reviews, respectively. Sixteen of the studies in this review were unpublished, permitting an examination of whether there is publication bias in this line of research. Thus, the goal of this investigation was to provide a considerably more comprehensive, quantitative review of the effects of simulation games on training outcomes. Accurate effect size estimates are imperative for both researchers and practitioners in that they provide a basis for comparing the effects of alternative instructional methods, for conducting power analyses in future research, and for estimating training utility. In the following section, I review definitions of games and simulations in order to alleviate some of the confusion that the interchangeable use of these terms has generated. I then review several motivation and learning theories in order to provide an integrated theoretical framework for understanding the instructional effectiveness of simulation games.

Definition of Simulation Games

The simulation game literature is plagued by an abundance of definitions and little consensus on the defining features of instructional simulations and games (Garris et al., 2002; Hays, 2005; O'Neil et al., 2005). Several popular definitions of games agree that they are entertaining, interactive, rule-governed, goal-focused, competitive, and they stimulate the imagination of players (Driskell & Dwyer, 1984; Gredler, 1996; Tobias & Fletcher, 2007; Vogel et al., 2006). The distinguishing feature of simulations is that they are reality based, but they can also incorporate common game features such as rules and competition (Bell et al., 2008; Hays, 2005; Tobias & Fletcher, 2007).

Consistent with Tennyson and Jorczak (2008), I propose that there are no longer clear boundaries between these two instructional methods. For example, Ricci, Salas, and Cannon-Bowers (1996) researched the effectiveness of a game for teaching chemical, biological, and radiological defense. This game involved answering a series of multiple-choice questions, but it provided very little entertainment value, which is a key feature of games in most definitions. In addition, simulations are proposed to be based on reality, but the literature is rich with examples of computer-based simulations that do not faithfully recreate work-related experiences (e.g., North, Sessum, & Zakalev, 2003; Shute & Glaser, 1990; Taylor & Chi, 2006). Given the blurred boundaries, it is not valuable to categorize these educational tools as either simulations or games; instead, I use the broad term *simulation games*. Consistent with previous definitions (e.g., Siemer & Angelides, 1995; Tennyson & Jorczak, 2008), I define computer-based simulation games as instruction delivered via personal computer that immerses trainees in a decision-making exercise in an artificial environment in order to learn the consequences of their decisions. Thus, learning must be the primary goal of the simulation game in order to be included in this research. Furthermore, online training is distinct from simulation games—online training is always delivered via the Internet, whereas simulation games may be online or hosted on a single workstation computer; online training also utilizes a breadth of instructional methods (e.g., lecture, assignments, discussion), one of which may be simulation games.

Theoretical Framework for Simulation Games

Several theories have been proposed for understanding the role of simulation games in enhancing the motivation and learning of trainees. The earliest theories focused exclusively on the motivational potential of simulation games, ignoring their potential for enhancing work-related skills. Consistent with cognitive-oriented learning theories (Bruner, 1962; Piaget, 1951), Malone's (1981) theory emphasized the importance of intrinsically motivating, play-type activities for promoting deep learning. When trainees are intrinsically motivated, they exert more effort to learn the material, enjoy learning more, and are more likely to apply the material outside of the game environment.

Garris et al.'s (2002) input–process–outcome model also focused on the motivational capacity of simulation games. They proposed that instructional content and game characteristics serve as input to the game cycle (the process), which ultimately influences learning outcomes. The game cycle is the motivational aspect of the model. It represents a cyclical relationship among user judgments (i.e., enjoyment, task involvement, and self-efficacy), user behavior (i.e., effort expended and the decision to continue playing), and system feedback (i.e., information about one's performance). Simulation games should be designed to be engaging and engrossing to perpetuate the game cycle. This creates a flow state that represents an optimal performance situation in which trainees are so involved in the simulation game that nothing else seems to matter (Csikszentmihalyi, 1990). After one or more game cycles, trainees should participate in a debriefing session in which the simulation game as well as its applicability to the real world is discussed. Debriefing enhances the transfer of what trainees have learned in the simulation game to the job. The benefits of this model are that it captures the process by which simulation games motivate trainees and engage them in game play and demonstrates the essential role of debriefing in enhancing transfer from game play to the job.

Tennyson and Jorczak (2008) pushed simulation game theory a step further by focusing on the cognitive systems of trainees that affect learning. Interactive cognitive complexity is an integrative information processing model that proposes learning is the result of an interaction between variables internal and external to the cognitive systems of trainees. Trainees' affective (e.g., motivation and attitudes) and cognitive (e.g., memory, knowledge base, and executive control) structures interact with each other and with sensory information from the simulation game in order to enhance trainees' knowledge base. The process is iterative as sensory information continuously interacts with trainees' cognitive system and new information is stored. According to the theory, both affective and cognitive structures are essential components of the cognitive system, and simulation games should be more effective than other instructional methods because they target both of these structures.

Together these theories suggest that, ideally, simulation games utilize a combination of entertainment and active learning principles to immerse trainees in learning the course material. Entertainment will ensure trainees repeatedly engage in the learning experience, enhancing learner motivation. Active learning principles provide trainees with responsibility for making important learning decisions and rely on inductive learning in which trainees must explore the task in order to infer the rules for effective performance (Bell & Kozlowski, 2008; Frese, Brodbeck, Heinbokel, Mooser, Schleiffenbaum, & Thiemann, 1991). Learning and adaptive transfer are enhanced via self-regulatory processes when active learning principles are incorporated in the training design (Bell & Kozlowski, 2008). However, simulation games do not always utilize entertainment and active learning principles to enhance the learning experience.

In the following sections, I rely on these theoretical frameworks for hypothesizing the effects of simulation games on affective and cognitive training outcomes. I then rely on theory and frameworks of simulation game characteristics (e.g., Garris et al., 2002; Tennyson & Jorczak, 2008; Wilson et al., 2009) to develop hypotheses regarding the characteristics of simulation games and the instructional context that enhance learning.

Effect of Simulation Games on Training Outcomes

My goal was to examine the effect of simulation games on a comprehensive set of training outcomes. Kraiger, Ford, and Salas (1993) proposed that learning is multidimensional and may be observed by changes in affective, cognitive, or skill capabilities. Furthermore, simulation game theories emphasize that affective, behavioral, and cognitive processes are all critical indicators of training effectiveness (Garris et al., 2002; Malone, 1981; Tennyson & Jorczak, 2008). Thus, I reviewed the literature in order to determine the effect of simulation games, relative to a comparison group, on three affective (i.e., motivation, trainee reactions, and self-efficacy), one behavioral (i.e., effort), two cognitive (i.e., declarative knowledge and retention), and two skill-based (i.e., procedural knowledge and transfer) training outcomes. Although other evaluation criteria are valuable for clarifying the effectiveness of simulation games, insufficient primary studies were available to examine other indicators of training effectiveness. Below I review the literature on the effect of simulation games on training outcomes and present hypotheses for the four outcomes—self-efficacy, declarative knowledge, procedural knowledge, and retention—where sufficient research has been conducted to calculate the meta-analytic effect size.

Motivation, effort, and trainee reactions. Malone (1981) and Garris et al. (2002) both suggested that simulation games are effective because they target affective processes. There is a cyclical relationship among trainees' enjoyment of game play, intrinsic motivation, and the decision to continue playing (Garris et al., 2002). Thus, both of these theories suggest that motivation, effort, and trainee reactions are key training outcomes that should be affected by simulation games. Motivation is a psychological training outcome and refers to the extent to which trainees strive to learn the content of a training program, whereas effort is a behavioral outcome and refers to the amount of time and energy devoted to training (Fisher & Ford, 1998; Noe, 1986; Sitzmann & Ely, 2010). Reactions refer to trainees' satisfaction with the instructional experience (Sitzmann, Brown, Casper, Ely, & Zimmerman, 2008).

Despite these theoretical assertions, this review revealed that a limited range of studies has compared a simulation game group to a comparison group on these three outcomes—only one study has compared posttraining motivation (DeRouin-Jessen, 2008), two studies have compared effort levels (DeRouin-Jessen, 2008; Sukhai, 2005), and three have compared trainee reactions (DeRouin-Jessen, 2008; Parchman, Ellis, Christinaz, & Vogel, 2000; Ricci et al., 1996). Thus, the scarcity of research in this area precludes an empirical test of the effect of simulation games on posttraining motivation, effort, and trainee reactions.

Self-efficacy. Posttraining self-efficacy refers to trainees' confidence that they have learned the information taught in training and can perform training-related tasks (Bandura, 1997). In contrast to the aforementioned training outcomes, sufficient empirical research has been conducted to compare posttraining self-efficacy for trainees taught with simulation games, relative to a comparison group. A critical precursor to high self-efficacy is experience with work-related tasks (Bandura, 1991). Simulation games are interactive and tend to be more engaging than other instructional methods (Ricci et al., 1996; Vogel et al., 2006). High interactivity and the opportunity to make choices while participating in simulation games may result in trainees feeling empowered, ultimately enhancing trainees' self-efficacy (Bandura, 1993; Tennyson & Jorczak, 2008). Simulation games should also promote mastery of the material via letting the trainee attempt to apply the knowledge and skills, enhance metacognitive activity due to actively engaging with the material, and promote positive emotional arousal, all of which have positive effects on self-efficacy (Bandura, 1977; Bell & Kozlowski, 2008; Brown & Ford, 2002; Garris et al., 2002; Kozlowski & Bell, 2006; Malone, 1981). This is consistent with empirical research by Randell, Hall, Bizo, and Remington (2007) that found simulation games resulted in higher posttraining self-efficacy than a comparison group when learning how to treat children with autism.

Hypothesis 1: Posttraining self-efficacy will be higher for trainees in the simulation game group than the comparison group.

Learning. I also compared simulation games to a comparison group in terms of their effect on four learning outcomes: declarative knowledge, procedural knowledge, retention, and training transfer. Declarative knowledge refers to trainees' memory of the facts and principles taught in training and the relationship among knowledge elements (Kraiger et al., 1993). Procedural knowledge refers to information about how to perform a task or action. Retention is a delayed assessment of declarative knowledge and refers to trainees' memory of the factual information taught in training several weeks or months after leaving the training environment. Finally, training transfer refers to the successful application of the skills gained in a training context to the job (Baldwin & Ford, 1988).

Interactive cognitive complexity theory proposes that simulation games maximize learning because they simultaneously engage the affective and cognitive processes of trainees (Tennyson & Breuer, 1997; Tennyson & Jorczak, 2008). Simulation games tend to be more interactive than other instructional methods, and interactivity is a critical component of effective instruction (Jonassen, 2002; Northup, 2002; Sitzmann, Kraiger, Stewart, & Wisher, 2006). Several previous reviews have examined the effect of simulation games on cognitive learning outcomes. Randel et al. (1992) reviewed the effectiveness of simulation games, primarily for teaching children, and found 27 out of 68 studies favored the use of simulation games over classroom instruction, whereas 3 favored classroom instruction. Moreover, 14 studies examined retention, and 10 of these studies found retention was greater for trainees taught with simulation games than classroom instruction. Similarly, Vogel et al. (2006) conducted a meta-analysis of the instructional effectiveness of simulation games for teaching children and adult learners. When averaging across these trainee populations and learning dimensions, they found learning gains were greater for trainees taught with simulation games than traditional teaching methods ($z = 6.05$, $N = 8{,}549$).

Only one study has compared simulation games to other instructional methods for enhancing training transfer. Meyers, Strang, and Hall (1989) found trainees who used a simulation game to practice counseling preschoolers significantly outperformed trainees who learned by coding audiotapes of children's disfluency on six out of eight transfer measures (d ranged from $-.21$ to 2.39; $N = 20$). Despite the paucity of research examining transfer, there are a sufficient number of studies to conduct a meta-analysis for the other three learning outcomes (declarative knowledge, procedural knowledge, and retention), and based on interactive cognitive complexity theory and previous research, I hypothesize a learning advantage for trainees taught with simulation games.

Hypotheses 2–4: Posttraining declarative knowledge (Hypothesis 2), posttraining procedural knowledge (Hypothesis 3), and retention of the training material (Hypothesis 4) will be higher for trainees in the simulation game group than the comparison group.

Moderators of the Effectiveness of Simulation Games

A second objective of the study is to examine moderators of the effectiveness of simulation games relative to the comparison group. The moderator variables were chosen based on the features of simulation game design and the instructional situation discussed in previous reviews of this literature (e.g., Bell et al., 2008; Garris, et al. 2002; Hays, 2005; Malone, 1981; Wilson et al., 2009) as well as course design features that theory and previous meta-analyses indicate influence learning (e.g., Bell & Kozlowski, 2008; Brown & Ford, 2002; Keith & Frese, 2008; Sitzmann et al., 2006). In the following sections, I hypothesize the moderating effect of two simulation game characteristics (i.e., entertainment value and whether the majority of instruction in the simulation game was active or passive), two instructional context characteristics (i.e., whether trainees had unlimited access to the simulation game and whether the simulation game was the sole instructional method for the treatment group), and one characteristic of the comparison group (i.e., whether the instructional methods used to teach the comparison group as a substitute for the simulation game were active or passive) on learning from a simulation game relative to the comparison group. Finally, four methodological moderators were examined in order to ensure that observed differences in effect sizes are driven by the hypothesized moderators rather than other factors.

Entertainment value. Simulation games that were high in entertainment value contained at least one feature common to either board games or video games including rolling virtual dice and moving pegs around a board, striving to make the list of top scorers, playing the role of a character in a fantasy world, and shooting foreign objects. For example, Moshirnia (2008) had players assume the role of either George Washington or

King George III as they fought other characters while learning about the American Revolutionary War. This simulation game had several entertaining features including playing the role of a character in a fantasy world and fighting other characters. The simulation game utilized by Boyd and Murphrey (2002) was much less entertaining. Players assumed the role of a human resource director with personal knowledge of the background of a job applicant. The objective of the simulation game was to improve leadership skills as players decided whether to reveal what they knew about the candidate to the search committee and discovered the consequences of their actions.

Malone (1981) theorized that the features that make a simulation game intrinsically motivating are challenge, curiosity, and fantasy. Intrinsically motivating features of simulation games increase self-determination because trainees choose to engage in game play as they find it interesting and enjoyable (Deci & Ryan, 1985). Moreover, researchers have demonstrated that instruction embedded in a fantasy context increases both interest and learning (Cordova & Lepper, 1996; Parker & Lepper, 1992). Trainees become immersed in entertaining and fantasy-based simulation games more than other instructional methods, thereby increasing learning (Cordova & Lepper, 1996; Garris et al. 2002; Wilson et al., 2009).

Hypothesis 5: The entertainment value of the simulation game will moderate learning from simulation games; relative to the comparison group, trainees will learn more from simulation games that are high rather than low in entertainment value.

Activity level of the simulation game group. Consistent with the definition advanced for simulation games, all are interactive. However, some utilize interaction to keep trainees engaged (e.g., fighting other characters), but the interaction does not contribute to learning the course material. Indeed, some of the simulation games included in the review presented the majority of the learning content in a passive manner via text or audio explanations. For example, Parchman et al. (2000) embedded presentations in the simulation game so trainees could review course topics or participate in linear computer-based instruction.

Some theorists propose that learning requires active engagement with the material (Brown & Ford, 2002; Jonassen, 2002). Active learning enhances metacognition; that is, trainees who actively learn the material exert more cognitive effort to evaluate information and integrate it with their existing knowledge base (Bell & Kozlowski, 2008; Brown & Ford, 2002). Practicing the key components of a task during training should help trainees develop an understanding of the deeper, structural features of the task (Newell, Rosenbloom, & Laird, 1989; Sitzmann, Ely, & Wisher, 2008). Teaching core training material in a passive manner in simulation games is contrary to theory suggesting that one of the instructional advantages of simulation games is that they engage trainees in the learning experience (Chen & O'Neil, 2008; Garris et al., 2002). By relying on passive instruction in a simulation game, it dilutes the instructional advantage of the simulation game.

Hypothesis 6: The activity level of the instruction in the simulation game will moderate learning from simulation games; relative to the comparison group, trainees will learn more from simulation games that actively engage trainees in learning rather than passively conveying the instructional material.

Unlimited access to the simulation game. Garris et al.'s (2002) model theorizes that one of the advantages of simulation games is that they are intrinsically motivating. There is a cyclical relationship among users' enjoyment of the simulation game, the decision to continue playing, and feedback on one's performance. Learning benefits occur when trainees choose to repeatedly engage in game play, mastering the skills that are taught. Thus, the full learning potential of simulation games is only realized if trainees can access the simulation game as many times as desired. Consistent with this argument, cognitive learning theorists (Bruner, 1962; Piaget, 1951) argued that intrinsically motivating play-type activities are crucial for deep learning (Malone, 1981). When trainees participate in traditional learning activities, they rarely display the level of effort and motivation that is typical of simulation games, thereby limiting the learning potential (Tennyson & Jorczak, 2008).

Hypothesis 7: Whether trainees have unlimited access to the simulation game will moderate learning from simulation games; relative to the comparison group, trainees will learn more from simulation games when they have unlimited access to the simulation game than when access to the simulation game is limited.

Simulation game as sole instructional method. Courses differed in terms of whether the simulation game was the only instruction the treatment group received (e.g., Kim, Kim, Min, Yang, & Nam, 2002; North et al., 2003) or the simulation game was used as a supplement to other instructional methods (e.g., Ebner & Holzinger, 2007; Ortiz, 1994). Garris et al.'s (2002) model proposes that a debriefing session, to review and analyze what happened during game play, mediates the relationship between the game cycle and learning. Lee's (1999) review of the literature revealed that the instructional effectiveness of simulations, relative to a comparison group, is enhanced when trainees have the opportunity to review information before practicing in the simulation. Moreover, Hays (2005) proposed that simulation games should be embedded in instructional programs that elaborate on how the information conveyed in the game is pertinent to trainees' jobs. Rarely are simulation games so well designed that trainees can learn an instructional topic through game play alone (Tennyson & Jorczak, 2008). When training utilizes a breadth of instructional methods, trainees who are having difficulty learning the material can continue to review the material with multiple instructional methods to increase their mastery of the course content (Sitzmann et al., 2006; Sitzmann, Ely, et al., 2008).

Hypothesis 8: Whether simulation games are embedded in a program of instruction will moderate learning from simulation games; relative to the comparison group, trainees will learn more from simulation games that are embedded in a program of instruction than when they are the sole instructional method.

Activity level of the comparison group. Trainees in the comparison group were often taught via a different instructional method as a substitute for utilizing the simulation game. However, studies differed in terms of whether the comparison group learned by means of active (e.g., Hughes, 2001; Mitchell, 2004; Willis, 1989) or passive (e.g., Bayrak, 2008; Frear & Hirschbuhl, 1999; Shute & Glaser, 1990) instructional methods. Trainees are active when they are reviewing with a computerized tutorial, participating in a discussion, and completing assignments. Trainees are passive when they are listening to a lecture, reading a textbook, or watching a video.

One of the advantages of simulation games is that they typically require trainees to be active while learning the course material (Ricci et al., 1996; Vogel et al., 2006). Actively engaging with the course material enhances learning (Newell et al., 1989; Sitzmann et al., 2006; Webster & Hackley, 1997), regardless of whether trainees are participating in a simulation game or learning from another instructional method. Active learning assists trainees in developing both a refined mental model of the training topic and the adaptive expertise necessary to apply trained skills under changing circumstances (Bell & Kozlowski, 2008; Keith & Frese, 2005, 2008). Thus, the difference in learning between the simulation game and comparison groups should be less when the comparison group is active while learning the course material.

Hypothesis 9: The activity level of the comparison group will moderate learning; relative to trainees taught with simulation games, the comparison group will learn more when they are taught with active rather than passive instructional methods.

Methodological moderators. One of the advantages of meta-analysis is it allows for a comparison of studies that differ in experimental rigor and other methodological factors (Lipsey, 2003). It is only by controlling for methodological artifacts that one can be certain that observed differences in effect sizes are driven by the hypothesized moderators rather than factors that are spuriously correlated with the outcome variable. Thus, this meta-analysis examined whether the effect of simulation games, relative to a comparison group, on learning was related to four methodological moderators: random assignment to experimental conditions, rigor of the study design (pretest–posttest vs. posttest only), publication status (published vs. unpublished), and year of the publication, dissertation, or presentation. Media comparison studies often confound instructional media with instructional quality, student motivation, and other factors (Clark, 1983, 1994; Sitzmann et al., 2006). Randomly assigning trainees to experimental conditions and utilizing a rigorous study design would allow researchers to rule out alternative explanations for differences in learning between simulation game and comparison groups. The publication status moderator analysis will reveal whether there is evidence of a file drawer problem in simulation games research (Begg, 1994). That is, do published studies tend to report effect sizes that are larger in magnitude than unpublished studies? Finally, the year moderator results will clarify whether the effectiveness of simulation games, relative to a comparison group, has increased with improvements in training technology in recent years.

Method

Literature Search

Computer-based literature searches of PsycInfo, ERIC, and Digital Dissertations were used to locate relevant studies. To be included in the initial review, each abstract had to contain terms relevant to games or simulations and training or education. Initial searches resulted in 4,545 possible studies. A review of abstracts limited the list to 264 potentially relevant reports, of which 40 met the inclusion criteria. In addition, I manually searched reference lists from recently published reports focusing on the effectiveness of simulation games (e.g., Lee, 1999; Schenker, 2007; Tobias & Fletcher, 2007; Vogel et al., 2006; Wilson et al., 2009). These searches identified an additional 13 reports.

A search for additional published and unpublished studies was also conducted. First, I manually searched the Academy of Management and Society for Industrial and Organizational Psychology conference programs. Second, practitioners and researchers with expertise in training were asked to provide leads on published and unpublished work. In all, I contacted 117 individuals. These efforts identified an additional two studies for a total of 55 reports, yielding 65 independent samples.

Inclusion Criteria

Due to the upward bias in effect sizes from gain score research (Lipsey & Wilson, 2001), this report focuses exclusively on studies that compared posttraining outcomes for simulation game and comparison groups. Trainees in the simulation game group received all or some of their learning content via a simulation game. The comparison group differed across studies and ranged from a no-training control condition to trainees who received alternative instructional methods as a substitute for the simulation game.

To be included in this review, studies had to meet four additional criteria: (a) The article reported results that allowed for the calculation of a d statistic—group means and standard deviations, a correlation, t-test, or univariate F-test; (b) data had to be collected at the individual-level of analysis; data collected at the group-level were excluded (e.g., Brannick, Prince, & Salas, 2005; Rapp & Mathieu, 2007) because there was insufficient research to conduct a meta-analysis for team-based training outcomes; (c) participants were nondisabled adults ages 18 or older; (d) the training facilitated potentially job-relevant knowledge or skills (i.e., not coping with physical or mental health challenges). The last two criteria support generalization to work-related adult training programs.

Coding and Interrater Agreement

In addition to recording all relevant effect sizes and sample sizes, the following information was coded from each study: (a) self-efficacy, (b) learning outcomes, (c) entertainment value, (d) majority of simulation game instruction was active or passive, (e) whether trainees could access the simulation game as many times as desired, (f) simulation game as standalone instruction or supplement to other instructional methods, (g) activity level of the instruction the comparison group received as a substitute for the simulation game, (h) random assignment to experimental conditions, (i) rigor of the study design, (j) publication status, and (k) year of the publication, dissertation, or presentation. Scales for each moderator were drafted prior to coding and modified following initial attempts to code articles. The coding rules are described below.

Self-efficacy. Posttraining self-efficacy refers to trainees' confidence that they have learned the information taught in training and can perform training-related tasks (Bandura, 1997). For example, Ricci et al. (1996) measured self-efficacy by asking trainees to rate their confidence that they would remember what they had learned that day.

Learning outcomes. Declarative and procedural knowledge were coded based on Kraiger et al.'s (1993) multidimensional framework of learning. Declarative outcomes are knowledge assessments designed to measure if trainees remembered concepts presented during training; they were always assessed with a written test. Procedural outcomes were defined as the ability to perform the skills taught in training. They were assessed by participating in an activity (e.g., simulation or role-play) or with a written test that required trainees to demonstrate memory of the steps required to complete the skills taught in training. Retention was coded as delayed measures of declarative knowledge. The majority of studies assessed retention between 1 and 4 weeks after the end of training, but one study—Lawson, Shepherd, and Gardner (1991)—assessed retention between 88 and 167 days after training ended.

Entertainment value. Simulation games were coded as having either a high or low level of entertainment value. Simulation games had a high level of entertainment value when they contained at least one feature common to either board games or video games including rolling virtual dice and moving pegs around a board, striving to make the list of top scorers, playing the role of a character in a fantasy world, and shooting foreign objects. For example, DeRouin-Jessen (2008) examined the effectiveness of an entertaining simulation game in which trainees maneuvered through a virtual town while learning about equal employment opportunity laws. The simulation game included elements of fantasy, such as a "mentor character" who appeared as a floating ball of light and spoke to participants in an "ethereal voice" (p. 67). The McGill negotiation simulator was low in entertainment value (Ross, Pollman, Perry, Welty, & Jones, 2001); trainees spent up to an hour negotiating the sale of an airplane.

Simulation game instruction was passive or active. Studies were coded as to whether the majority of the instructional content in the simulation game was passively or actively conveyed to trainees. Simulation games that taught the majority of the course content via text or audio explanations were coded as providing trainees with passive instruction; simulation games that taught the majority of the course content via activities, practice, and engaging with the simulation game were coded as providing trainees with active instruction. For example, DeRouin-Jessen (2008) examined a simulation game that taught with passive instruction.

Trainees navigated a virtual environment (during which time they were not learning course content) and then stopped to read factual information on equal employment opportunity laws from books in the simulation game. In contrast, Gopher, Weil, and Bareket's (1994) simulation game relied on active instruction as trainees honed flight-relevant skills by controlling a spaceship while defending it from mines and trying to destroy an enemy's fortress.

Unlimited access to the simulation game. Studies were coded as to whether trainees were allowed to utilize the simulation game as many times as desired (e.g., Cataloglu, 2006; Sterling & Gray, 1991) or a limit was imposed on trainees' interaction with the simulation game (e.g., Hughes, 2001; Mitchell, 2004).

Simulation game as standalone instruction. Studies were coded as to whether the simulation game was the sole instructional method used to teach the treatment group (e.g., Desrochers, Clemmons, Grady, & Justice, 2000; Faryniarz & Lockwood, 1992) or the simulation game was a supplement to other instructional methods (e.g., Sukhai, 2005; Willis, 1989).

Activity level of comparison group. I also examined whether the comparison group received instruction as a substitute for the instruction received in the simulation game and, if so, whether the instructional methods used to teach the comparison group were active or passive. Passive instructional methods include listening to lectures and reading case studies or textbooks (e.g., Ivers & Barron, 1994; Taylor & Chi, 2006). Active instructional methods include computerized tutorials, completing assignments, and conducting laboratory experiments (e.g., Hughes, 2001; Zacharia, 2007). However, several studies did not provide instruction for the comparison group as a substitute for utilizing the simulation game (e.g., Cameron & Dwyer, 2005; Ortiz, 1994).

Methodological moderators. Four methodological moderators were coded: random assignment to experimental conditions, rigor of the study design (pretest–posttest vs. posttest only), publication status (published vs. unpublished), and year of the publication, dissertation, or presentation.

Coding and Interrater Agreement

All articles were coded independently by two trained raters. Interrater agreement (Cohen's kappa) was excellent according to Fleiss (1981) for each of the coded variables, with a coefficients ranging from .82 to .98 for each of the moderator variables. All coding discrepancies were discussed until a consensus was reached.

Calculating Effect Size Statistic (d) and Analyses

The Hedges and Olkin (1985) approach was used to analyze the data. The effect size calculated for each study was *d*, the difference between the simulation game and the comparison group, divided by the pooled standard deviation. When means and standard deviations were not available, effect sizes were calculated from a correlation, *t*-test, or univariate *F*-test using formulas reported in Glass, McGaw, and Smith (1981) and Hunter and Schmidt (1990). Effect sizes were corrected for small-sample bias using formulas provided by Hedges and Olkin (1985). The self-efficacy effect sizes were corrected for attenuation using the scale reliabilities reported in each study. When a study failed to provide a coefficient alpha reliability estimate, I used the average self-efficacy reliability across all samples from this study and from Sitzmann, Brown, et al. (2008). The average reliability was .83. I did not correct the learning outcomes effect sizes for attenuation due to the lack of available test–retest or alternate forms reliability coefficients.

Occasionally a single study would report data from two simulation game training groups and/or two comparison groups. In these situations, an effect size was calculated for all possible simulation game–comparison group pairs and averaged by weighting each of the effect sizes by the sum of the sample size of the independent group and one half of the sample size of the nonindependent group. Thus, the nonindependent sample was weighted according to its sample size in the overall effect size. In addition, whenever a single study reported multiple effect sizes based on the same sample for a single criterion, the effect size that was most similar to the other assessments of that particular relationship was used in the meta-analysis. Studies that included multiple independent samples were coded separately and treated as independent.

Finally, 95% confidence intervals were calculated around the weighted mean *d*s. Confidence intervals assess the accuracy of the estimate of the mean effect size and provide an estimate of the extent to which sampling error remains in the weighted mean effect size (Whitener, 1990).

Outlier Analyses

Prior to finalizing the analyses, a search for outliers was conducted using Huffcutt and Arthur's (1995) sample-adjusted meta-analytic deviancy (SAMD) statistic. Based on the results of these analyses, Blunt (2007) was identified as a potential outlier with a SAMD value of 7.79. A review of the descriptive statistics reported in Blunt revealed that the learning outcome (grade in course) included zeros for multiple participants in the comparison group. Thus, Blunt included participants in the learning analysis even if they dropped out of training. As such, the three samples reported in Blunt were removed from all analyses.

Moderator Analyses

Hedges and Olkin's (1985) homogeneity analysis was used to determine whether the effect sizes were consistent across studies. For main effect analyses, the set of effect sizes was tested for homogeneity with the Q_T statistic. Q_T has an approximate χ^2 distribution with $k - 1$ degrees of freedom, where k is the number of effect sizes. If Q_T exceeds the critical value, then the null hypothesis of homogeneity is rejected. Rejection indicates that there is more variability in effect sizes than expected by chance, suggesting that it is appropriate to test for moderators.

The goal of the moderator analyses was to examine if the effectiveness of simulation games, relative to the comparison group, differed based on features of simulation games and the comparison group, characteristics of the instructional context, and methodological factors. Moderating effects were tested by classifying studies according to the moderator categories and testing for homogeneity between and within categories (Lipsey & Wilson, 2001). For each moderator, a between-class goodness-of-fit statistic, Q_B, was calculated to test for homogeneity of effect sizes across moderator categories. It has an approximate χ^2 distribution with $j - 1$ degrees of freedom, where j is the number of moderator categories. If Q_B exceeds the critical value, it indicates that there is a significant difference across moderator categories; this is analogous to a significant main effect in an analysis of variance. In addition, a within-class goodness-of-fit statistic, Q_W, was calculated to test for homogeneity of effect sizes within each moderator category. It has an approximate χ^2 distribution with $k - j$ degrees of freedom, where k is the number of effect sizes included in the analysis. If Q_W exceeds the critical value, it indicates that the effect sizes within the moderator categories are heterogeneous.

The moderating effect of year of publication, dissertation, or presentation was tested with a correlation weighted by the inverse of the sampling-error variance between the moderator variable and the effect sizes. A limitation of the subgroup approach to moderator analyses is the inability to account for the joint effect of correlated moderators. Thus, I also utilized weighted least squares (WLS) regression to examine the joint effect of the moderators on learning. Effects sizes were weighted by the inverse of the sampling error variance as described by Steel and Kammeyer-Mueller (2002).

TABLE 1
Meta-Analytic Results for Self-Efficacy and Cognitive Learning Outcomes Comparing Trainees Taught With Simulation Games to a Comparison Group

	d	Standard error	k	N	95% Confidence interval Lower	Upper	Q_T
Self-efficacy	0.52	.10	8	506	0.32	0.72	38.33*
Declarative knowledge	0.28	.04	39	2,758	0.20	0.36	283.99*
Procedural knowledge	0.37	.07	22	936	0.23	0.50	85.66*
Retention	0.22	.08	8	824	0.07	0.37	67.03*

Note. d = inverse variance weighted mean effect size; k = number of effect sizes included in the analysis; N = sum of the sample sizes for each effect size included in the analysis; Q_T = homogeneity statistic.
*indicates the Q_T value is statistically significant at the .05 level and the effect sizes are heterogeneous.

Results

Fifty-five research reports contributed data to the meta-analysis, including 39 published reports, 12 dissertations, and 4 unpublished reports. These reports included data from 65 samples and 6,476 trainees. Learners were undergraduate students in 77% of samples, graduate students in 12% of samples, employees in 5% of samples, and military personnel in 6% of samples.[1] Across all samples providing demographic data, the average age of trainees was 23 years and 52% were male. A majority of the researchers who contributed data to the meta-analysis were in the fields of education (25%) and psychology (25%), whereas 12% were in business; 11% were in educational technology; 9% were in medicine; 6% were in computer science, math, or engineering; 5% were in science; and 7% were in other disciplines.

Main Effect Analyses

The main effects are presented in Table 1. The first hypothesis predicted that trainees receiving instruction via a simulation game would have higher levels of posttraining self-efficacy than trainees in the comparison group. Across eight studies, self-efficacy was 20% higher for trainees

[1]The effect of simulation games, relative to a comparison group, on learning did not significantly differ across undergraduate, graduate, employee, or military populations, $Q_B = \chi^2(3) = 3.63, p > .05$.

receiving instruction via a simulation game than trainees in a comparison group ($d = .52$). In addition, the confidence interval for self-efficacy excluded zero, supporting Hypothesis 1.

Hypotheses 2 through 4 predicted that trainees receiving instruction via a simulation game would learn more than trainees in a comparison group. Trainees receiving instruction via a simulation game had higher levels of declarative knowledge ($d = .28$),[2] procedural knowledge ($d = .37$), and retention ($d = .22$) than trainees in the comparison group. On average, trainees in the simulation game group had 11% higher declarative knowledge levels, 14% higher procedural knowledge levels, and 9% higher retention levels than trainees in the comparison group. Moreover, the confidence intervals for all three learning outcomes excluded zero, providing support for Hypotheses 2 though 4.

Based on the similarity of the learning effect sizes and the overlapping confidence intervals, I tested the homogeneity of effect sizes for the three cognitive learning outcomes. The Q_B was not significant ($\chi^2(2) = 2.21$, $p > .05$), suggesting that the mean effect sizes for declarative knowledge, procedural knowledge, and retention did not differ by more than sampling error. As such, the three learning outcomes were combined for the moderator analyses.

Moderator Analyses

Table 2 presents the mean effect sizes and estimates of homogeneity between (Q_B) and within (Q_W) the moderator subgroups. A significant Q_B indicates that the mean effect sizes across categories of the moderator variable differ by more than sampling error, suggesting that the moderator variable is having an effect (Lipsey & Wilson, 2001). The Q_B statistic was significant for all of the hypothesized moderators except entertainment value.

Hypothesis 5 predicted that, relative to the comparison group, trainees will learn more from simulation games that are high rather than low in entertainment value. Relative to the comparison group, trainees learned the same amount from simulation games that had high ($d = .26$) or low ($d = .38$) entertainment value, failing to support Hypothesis 5.

Hypothesis 6 proposed that trainees will learn more from simulation games that actively engage them in learning rather than passively conveying the instructional material, relative to the comparison group. When the majority of the instruction in the simulation game was passive, the comparison group learned more than the simulation game group

[2]One outlier was identified in this study—Blunt (2007)—who reported data from three samples. When Blunt was included in the declarative knowledge analysis $d = .55$ ($k = 42$, $N = 4{,}920$).

TABLE 2
Meta-Analytic Moderator Results Comparing Learning From Simulation Games to a Comparison Group

	d	Standard error	k	N	95% Confidence interval Lower	Upper	Homogeneity of effect sizes Q_B	Q_W
Hypothesized moderators								
Entertainment value								
High	.26	.08	9	809	.11	.41	1.84	350.57*
Low	.38	.04	51	3,216	.31	.45		
Simulation game instruction was active or passive								
Active	.49	.04	46	3,260	.41	.56	36.41*	251.56*
Passive	−.11	.09	9	521	−.29	.07		
Unlimited access to the simulation game								
Yes	.68	.07	10	925	.54	.82	21.11*	265.06*
No	.31	.04	45	2,738	.23	.38		
Simulation game as sole instructional method								
Game is a supplement	.51	.04	40	3,109	.43	.58	64.59*	304.19*
Game is standalone	−.12	.07	21	946	−.26	.01		
Activity level of instruction the comparison group received as a substitute for simulation game								
Active	−.19	.07	14	832	−.33	−.05	84.48*	277.35*
Hands on practice	−.13	.12	4	282	−.37	.11		
Computerized tutorial	−.70	.13	4	273	−.95	−.45		
Assignment	.86	.29	3	55	.30	1.42		
Discussion	.13	.18	1	130	−.21	.48		
Group activities & discussion	−.11	.28	1	50	−.67	.44		
Computerized tutorial & assignment	.12	.31	1	42	−.49	.72		
Passive	.38	.07	18	970	.24	.51		
Lecture	.45	.10	8	457	.25	.65		
Reading	.42	.10	8	419	.22	.62		
Watching a video	.50	.41	1	24	−.31	1.32		
Video & reading	−.30	.24	1	70	−.77	.17		
Combination of lecture & active instructional methods	.38	.10	6	425	.18	.58		
No instruction	.61	.05	25	1,844	.51	.71		
Methodological moderators								
Random assignment to experimental conditions								
Yes	.35	.05	34	1,997	.26	.45	1.37	301.89*
No	.43	.05	25	1,931	.34	.53		
Rigor of the study design								
Pretest–posttest	.36	.05	28	1,832	.26	.46	.00	352.41*
Posttest only	.36	.05	32	2,193	.27	.45		
Publication status								
Published	.52	.04	44	3,032	.44	.59	64.91*	287.50*
Unpublished	−.10	.07	16	993	−.23	.03		

Note. d = inverse variance weighted mean effect size; k = number of effect sizes included in the analysis; N = sum of the sample sizes for each effect size included in the analysis; Q_B = between-class goodness-of-fit statistic; Q_W = within-class goodness-of-fit statistic.
*indicates the Q value is statistically significant at the .05 level.

($d = -.11$). However, when the majority of the instruction in the simulation game was active, the simulation game group learned more than the comparison group ($d = .49$). These findings provide support for Hypothesis 6 and suggest that simulation games are more effective when they actively engage trainees in learning the course material.

Hypothesis 7 predicted that, relative to the comparison group, trainees will learn more from simulation games when they can utilize the simulation game as many times as desired than when trainees have limited access to the simulation game. In support of Hypothesis 7, trainees in the simulation game group outperformed the comparison group to a greater extent when they had unlimited access to the simulation game ($d = .68$) than when trainees had limited access to the simulation game ($d = .31$).

Hypothesis 8 predicted that trainees will learn more from simulation games that are embedded in a program of instruction than when simulation games are the sole instructional method, relative to the comparison group. Consistent with Hypothesis 8, when simulation games were used as a supplement to other instructional methods, the simulation game group had higher knowledge levels than the comparison group ($d = .51$). However, when simulation games were used as standalone instruction, trainees in the comparison group learned more than trainees in the simulation game group ($d = -.12$).

Hypothesis 9 predicted that, relative to trainees taught with simulation games, the comparison group will learn more when they are taught with active rather than passive instructional methods. In support of Hypothesis 9, the comparison group learned more than the simulation game group when they were taught with active instructional methods ($d = -.19$). However, the simulation game group learned more than the comparison group when the comparison group was taught with passive instructional methods ($d = .38$) or a combination of lecture and active instructional methods ($d = .38$). In addition, the effect size was largest when the comparison group did not receive instruction as a substitute for the instruction received in the simulation game ($d = .61$). Follow-up analyses examined which instructional methods have been compared to simulation games and whether the specific instructional methods used to teach the comparison group influenced the effect size. Hands-on practice of the information taught in the simulation game, computerized tutorials, and assignments were the active instructional methods that have been compared to simulation games with the greatest frequency, and the effect sizes varied greatly across these three instructional methods. Computerized tutorials were much more effective than simulation games ($d = -.70$), and hands-on practice was slightly more effective than simulation games ($d = -.13$). In contrast, simulation games were much more effective than assignments ($d = .86$). With regards to passive instructional methods, simulation games were more effective than lecture (d = .45) and reading (d = .42), which were the two most common instructional methods used to teach the comparison group.

Next, I calculated effect sizes for studies where the simulation game and comparison groups were matched in terms of the activity level of the instruction. When both instructional methods utilized active instruction, simulation games were 1% less effective than comparison instructional methods ($d = -.02$, $k = 10$, $N = 627$); when both instructional methods utilized passive instruction, simulation games were 3% less effective than comparison instructional methods ($d = -.07$, $k = 6$, $N = 383$). Thus, the effectiveness of simulation games and alternative instructional methods is similar when the instructional methods are matched in terms of the extent to which they actively engage trainees in learning the material.

Turning now to methodological variables, random assignment to experimental conditions (d = .35 and .43 for studies with and without random assignment, respectively) and the rigor of the study design (d = .36 for both pretest–posttest and posttest only designs) did not moderate learning from simulation games, relative to the comparison group ($Q_B = 1.37$ and .00, respectively, $p > .05$). However, effect sizes were much larger for published (d = .52) than unpublished ($d = -.10$) studies ($Q_B = 64.91$, $p < .05$). Finally, the inverse of the sampling error variance weighted correlation between the year of the publication, dissertation, or presentation and the effect size was not statistically significant (r = .16). This suggests that the effect of simulation games, relative to the comparison group, on learning has not changed over time.

Overall the results indicated that four of the five hypothesized moderators had an effect on learning from simulation games relative to the comparison group (the exception is entertainment value). Trainees in the simulation game group learned more, relative to the comparison, when simulation games actively rather than passively conveyed course material, trainees had unlimited access to the simulation game, and the simulation game was used as a supplement to other instructional methods rather than as standalone instruction. The comparison group learned more than the simulation game group when the comparison group received instruction as a substitute for the simulation game that actively engaged them in the learning experience. Furthermore, the extent to which the simulation game group learned more than the comparison group was greater in published than unpublished studies. However, the Q_W was significant for all of the moderator results, indicating that there was more variation within the moderator categories than would be expected by subject-level sampling error alone (Lipsey & Wilson, 2001). That is, none of the moderator variables independently accounted for all of the variability in the learning effect sizes across studies.

Joint moderator effects. A limitation of the subgroup approach for examining moderators is that it is restricted to testing individual hypotheses and does not control for possible confounds between correlated moderators (Hedges & Olkin, 1985; Miller, Glick, Wang, & Huber, 1991). To address this concern, I used WLS regression to test the joint effect of the moderators on the learning effect sizes. In block one, I controlled for publication status, given that it was the only methodological factor with a significant moderating effect. In block two, I entered the five hypothesized moderators. Only 55 studies provided the information necessary to code all six moderators so statistical significance was interpreted at the .10 level. Publication status accounted for a significant 9.8% of the variance in learning ($\beta = .31; p < .05$). After controlling for publication status, the five hypothesized moderators accounted for an additional 18.0% of the variance in learning ($p < .10$). However, the only moderator with a significant main effect was simulation game as the sole instructional method ($\beta = .28; p < .10$). Trainees learned more from simulation games, relative to the comparison group, when the simulation game was used as a supplement to other instructional methods rather than as stand-alone instruction. This suggests that some of the observed moderator effects have multiple determinants, and whether the simulation game is the sole instructional method may be driving some of the other moderator results. However, the results also suggest that there is some added utility in considering the effects of the other hypothesized moderators. After controlling for both publication status and supplement versus stand-alone instruction, entertainment value, activity level of simulation game instruction, unlimited access to the simulation game, and the activity level of comparison groups' instruction accounted for an additional 5.2% of the variance in learning effect sizes. Thus, there is some value added by considering the effects of all of the hypothesized moderators.

Discussion

Organizations and universities are investing millions of dollars in computer-based simulation games to train their employees and college students (Bell et al., 2008; Summers, 2004). Previous reviews of the literature provide mixed evidence of the extent to which simulation games are more effective than other instructional methods and have the inherent assumption that the instructional effectiveness of simulation games is the same for children and adults (e.g., Lee, 1999; O'Neil et al., 2005; Randel et al., 1992; Vogel et al., 2006). Moreover, many researchers have speculated as to which features of simulation games and the instructional context increase trainees' motivation and learning (e.g., Malone, 1981;

Tennyson & Jorczak, 2008; Tobias & Fletcher, 2007; Wilson et al., 2009), but there is little consensus on which features are the most important. This meta-analysis addressed these debates by examining both how much adult trainees learn from computer-based simulation games, relative to a comparison group, and the instructional and contextual factors that contribute to higher levels of learning.

Instructional Effectiveness of Simulation Games

The meta-analytic results are favorable regarding the use of simulation games in training. Self-efficacy, declarative knowledge, procedural knowledge, and retention results all suggest that training outcomes are superior for trainees taught with simulation games relative to the comparison group. These results provide some support for Garris et al.'s (2002) input–process–output model and Tennyson and Jorczak's (2008) interactive cognitive complexity theory. Simulation games may have a positive effect on these outcomes because they aim to influence both affective and cognitive processes. Moreover, repeatedly engaging in the simulation game cycle may increase trainees' confidence in their ability to remember and apply the information taught in training.

Two key simulation game theories—Malone (1981) and Garris et al. (2002)—propose that the primary benefit of using simulation games in training is their motivational potential. Thus, it is ironic that a dearth of research has compared posttraining motivation for trainees taught with simulation games to a comparison group. A number of studies have compared changes in motivation and other affective outcomes from pre- to posttraining for trainees taught with simulation games (e.g., Jarmon, Traphagan, & Mayrath, 2008; Orvis, Horn, & Belanich, 2008; Venkatesh & Speier, 2000), but this research design suffers from numerous internal validity threats, including history, selection, and maturation (Cook & Campbell, 1979). Also, the use of pre-to-post comparisons may result in an upward bias in effect sizes (Lipsey & Wilson, 2001), leading researchers to overestimate the effect of simulation games on motivational processes.

Several previous training meta-analyses have examined the effectiveness of technology-delivered instruction. Comparing these results with the results of previous meta-analyses provides a comprehensive understanding of the value of utilizing technology to deliver training and how simulation games compare to other forms of technology-delivered instruction. Zhao, Lei, Lai, and Tan (2005) compared the effectiveness of distance education courses (i.e., courses where the instructor and students are physically separated) to face-to-face courses and found no difference in the overall effectiveness of the two delivery media. However, several previous meta-analyses have reported positive effect sizes for various forms of technology-delivered instruction relative to classroom instruction, including computer-assisted training (Kulik, 1994; Kulik & Kulik, 1991) and hypermedia systems (Liao, 1999). Recently, Sitzmann et al. (2006) found online instruction was 6% more effective than classroom instruction for teaching declarative knowledge, but the two delivery media were equally effective for teaching procedural knowledge, and trainees were equally satisfied with online and classroom instruction. In addition, blended learning was 13% more effective than classroom instruction for teaching declarative knowledge and 20% more effective than classroom instruction for teaching procedural knowledge, but trainees reacted 6% more favorably toward classroom instruction. In comparison, this meta-analysis found trainees in the simulation games group had 11% higher declarative knowledge levels, 14% higher procedural knowledge levels, and 9% higher retention levels than trainees in the comparison group. Furthermore, simulation games were 17% more effective than lecture and 5% more effective than discussion, the two most popular instructional methods in classroom instruction. Thus, the current results are in line with the results of other meta-analyses on technology-delivered instruction and suggest that, when properly employed, technology can enhance learning outcomes.

The main effect results support the continued investment in simulation games. However, computer-based simulation games are more expensive to develop than other forms of technology-delivered training, with complex simulation games costing between $5 and $20 million to create (Jana, 2006; T+D, 2007). Traditional online training takes an average of 220 hours to create each hour of instructional content, whereas online simulations require 750 to 1,500 hours to create each hour of instructional content (Bell et al., 2008; Summers, 2004). However, if some of the assumptions about simulation game play hold true—that a preponderance of employees will choose to play the simulation game, employees are willing to devote their free time to playing work-related simulation games, and simulation games reduce attrition from training (DeRouin-Jessen, 2008; Garris et al., 2002; Jana, 2006)—organizations may realize that investing in simulation games is a sound use of their training dollars. Furthermore, the cost of developing simulation games may be offset by the reduction in travel costs for training that used to be delivered via classroom instruction. In order to maximize the utility of simulation games, game designers need to focus on content reuse or utilizing software that streamlines the game development process in order to reduce the cost of developing simulation games. In addition, more research is needed to investigate the return on investment for simulation games relative to other instructional methods.

Effect of Simulation Game Design and the Instructional Situation on Learning

Most simulation game models and review articles propose that the entertainment value of the instruction is a key feature that influences instructional effectiveness (e.g., Garris et al., 2002; Tennyson & Jorczak, 2008; Wilson et al., 2009). Contrary to popular assumption, the empirical summary of the literature suggests that this feature did not impact learning.

However, whether the simulation game was implemented as stand-alone instruction or as a supplement to other instructional methods had a strong effect on learning from simulation games, relative to the comparison group. Furthermore, this was the only moderator that had a significant effect on learning while controlling for the other hypothesized moderators. Best practices in simulation game design recommend integrating this instructional method in a program of instruction rather than as a stand-alone instructional technique (Hays, 2005; Tobias & Fletcher, 2007). Garris et al.'s (2002) theory proposes that a debriefing session after game play mediates the effect on learning. Simulation games may be an ineffective stand-alone training tool because people do not naturally learn complex relationships from experience alone (Garris et al., 2002; Simons, 1993). This is consistent with Dewey's (1938) assumption that experience plus reflection is required for learning and Hays' (2005) recommendation that simulation games should be embedded in an instructional program. Rarely are simulation games so well designed that trainees can learn an instructional topic through game play alone (Tennyson & Jorczak, 2008), and even the best simulation games do not guarantee learning will occur (Salas, Bowers, & Cannon-Bowers, 1995; Salas, Bowers, & Rhodenizer, 1998).

Ensuring that the vast majority of the simulation game content is delivered via active, rather than passive, instruction also enhanced learning from simulation games, relative to the comparison group. Theory suggests that the advantage of simulation games is that they actively engage trainees in learning the training material (Garris et al., 2002; Malone, 1981; Ricci et al., 1996). Utilizing passive instruction in simulation games is contrary to recommendations for simulation game design (Ricci et al., 1996) and general best practices in training design (Jonassen, 2002; Northup, 2002; Sitzmann et al., 2006; Webster & Hackley, 1997). However, about 16% of the simulation games included in the meta-analysis conveyed the majority of the instructional content in a passive manner via text or audio explanations of the course material. This suggests that researchers and practitioners need to ensure that the instructional experience is accurately labeled and instruction delivered via simulation games is always actively conveyed. Otherwise the training is merely conventional instruction that is supplemented with a simulation experience. Thus, simulation game designers must utilize creative techniques for teaching work-related knowledge and skills during game play. This can best be accomplished by including instructional designers with pedagogical expertise on simulation game development teams (Fletcher & Tobias, 2006). It is critical to remember that simulation games are just tools for training, and learning principles must be incorporated in the design of simulation games to ensure that they are effective learning tools (Salas et al., 1998; Salas & Cannon-Bowers, 1997).

Learning was also enhanced when trainees could choose to utilize the simulation game as many times as desired. Extensive time to engage in game play is essential for the game cycle in Garris et al.'s (2002) model to occur. There is a cyclical relationship among enjoyment, the decision to continue playing the simulation game, and system feedback. When limits are placed on trainees' interaction with the simulation game, it stunts the game cycle, thereby limiting the learning potential.

One feature of the comparison group—whether the comparison group received instruction as a substitute for the content covered in the simulation game and, if so, whether the instruction was active or passive—was instrumental in determining the relative effectiveness of simulation games. The extent to which the simulation game group learned more than the comparison group was greatest when the comparison group did not receive an alternative form of instruction as a substitute for game play. However, learning was greater for the comparison than the simulation games group when the comparison group was actively engaged in learning the training material, while the treatment group utilized the simulation game. Consistent with Clark's (Clark, 1983, 1994) theory and previous meta-analytic findings (Sitzmann et al., 2006), this confirms that technology is a means for delivering training but does not have a direct effect on learning. Rather, computer-based simulation games and other instructional methods must actively engage learners in the instructional experience to maximize their learning potential.

Consistent with previous meta-analyses on psychological interventions (e.g., Lipsey & Wilson, 1993), I also found evidence of publication bias in the simulation games literature. Publication bias is often referred to as the "file drawer problem" and occurs when the probability that a study is published is dependent on the magnitude, direction, or significance of a study's results (Begg, 1994). Two previous meta-analyses in this area (Lee, 1999; Vogel et al., 2006) included a very limited number of unpublished studies. Thus, Vogel et al. may have overestimated the cognitive gains from simulation games. Furthermore, Lee's effect size of .41 for academic achievement may have been greater than the current learning effect sizes due to the failure to include sufficient unpublished research.

Accounting for the upward bias in published studies adds credibility to these findings.

Recommendations for Practitioners

The results support Salas and Cannon-Bowers' (2001) notion that it is misleading to conclude that a simulation game "(in and of itself) leads to learning" (p. 484). Simulation games should not be employed in training simply because the technology exists, but rather, careful consideration is required to determine training needs and which instructional features should be included in the simulation game (Salas et al., 1998). The entertainment value of simulation games did not affect how much trainees learned from simulation games relative to the comparison group. Rather, avoiding teaching trainees with passive instruction during game play was the simulation game feature that was instrumental in enhancing learning.

Simulation games need to actively engage the learner as they are reviewing the instructional material. The comparison group learned more than the simulation game group when the majority of the material covered in the simulation game utilized passive instructional techniques. For example, DeRouin-Jessen (2008) had trainees maneuver a game character through a fantasy world. Trainees were active as they maneuvered their characters, but they were not learning the course material at this point. Rather, trainees were expected to read a passage or listen to a character explain the course material once they reached their destination. Utilizing passive instruction in simulation games violates a fundamental course design principle—"learning is best accomplished through the active involvement of the students" (Webster & Hackley, 1997, p. 1284).

It is also important to consider the role of the instructional context in determining how much trainees learn from simulation games. Learning was maximized from simulation games, relative to the comparison group, when trainees had unlimited access to the simulation game. As such, organizations may benefit from providing trainees with a copy of the simulation game so that they can learn during their free time. Anecdotal and empirical evidence from Cold Stone Creamery, Canon Inc., and Cisco Systems suggests that employees are willing to utilize simulation games on their own time and improve their work-related competencies as a result of game play (Aldrich, 2007; Jana, 2006). However, to maximize the potential of computer-based simulation games, they should be used as a supplement to lecture, discussion, tutorials, or other instructional methods. Simulation games are beneficial for practicing work-related skills, but trainees must first learn work-related knowledge in order to apply it during game play. Moreover, a debriefing session after game play is beneficial for ensuring that trainees realize how their experience in the simulation game is applicable to the work environment (Garris et al., 2002; Salas & Cannon-Bowers, 2001).

Study Limitations and Directions for Future Research

Consistent with Wilson et al. (2009) and the limitations noted in previous meta-analyses (e.g., Sitzmann, Ely, Brown, & Bauer, 2010), I found the level of detail reported in primary research restricted which attributes of simulation games and the instructional context could be meta-analytically examined. Detailed descriptions of the training course and instructional situation would have enabled coding whether there was a debriefing session at the end of the simulation game as well as whether the training goals were salient to trainees during game play. It would also have enabled coding some of the moderators on a continuous rather than dichotomous scale, increasing the variability across studies in the moderator effects. In addition, psychological fidelity was not included in the meta-analysis due to its high correlation with entertainment value. However, Vogel et al. (2006) found the level of picture realism did not moderate the effectiveness of simulation games. Future research should include detailed descriptions of the training courses and instructional context in order to advance meta-analytic research.

Moreover, the hypothesized moderators were interrelated such that simulation games that were well designed tended to implement multiple hypothesized moderators. This is a common problem in meta-analytic research because studies that implement certain instructional features are likely to have other co-occurring features (Lipsey, 2003). These interrelationships represent statistical confounds that make it difficult to tease apart of the role of a single moderator. In this meta-analysis, when the hypothesized moderators were considered jointly, whether the simulation game was implemented as stand-alone instruction or a supplement to other instructional methods was the only hypothesized moderator that had a significant effect on learning. Although the other moderator effects were not significant, follow-up analyses revealed that they had a meaningful influence on the effect size, after controlling for publication status and supplement/stand alone. Additional primary research is needed to examine the joint and independent effects of characteristics of simulation games and the instructional context on affective, behavioral, and cognitive training outcomes. This research should investigate the effects of each of the components of entertainment value (e.g., challenge and fantasy) as well as other course design features (e.g., learner control, feedback, adaptive guidance), which previous research has demonstrated influence the effectiveness of technology-delivered instruction (Bell & Kozlowski, 2002; Kraiger & Jerden, 2007; Sitzmann et al., 2006). Researchers should also examine whether the amount of time that trainees spend reviewing with simulation games, rather than the level of access that they have, moderates the effectiveness of game play.

Another limitation is that there was insufficient primary research on simulation games to run the moderator analyses separately for the three learning outcomes. It is possible that the moderator variables may have different effects on these three outcomes. In order to overcome this limitation, additional primary research is needed on the effectiveness of simulation games. In addition, only one study was identified that compared posttraining motivation for trainees taught with simulation games to a comparison group (DeRouin-Jessen, 2008); two studies compared effort exerted (DeRouin-Jessen, 2008; Sukhai, 2005); three studies compared trainee reactions (DeRouin-Jessen, 2008; Parchman et al., 2000; Ricci et al., 1996); and one study compared transfer (Meyers et al., 1989). This is a monumental gap in our collective understanding of the effectiveness of simulation games and suggests the need for additional research on the affective and skill-based outcomes of simulation games, relative to other instructional methods. This research should randomly assign trainees to experimental conditions in order to eliminate potential confounds (e.g., trainee motivation) that may influence the level of effort that trainees exert, their satisfaction with the instruction, or their training transfer.

Hays (2005) proposed that simulation games are motivational, but they motivate trainees to play the game rather than to enhance their work-related knowledge and skills. Anecdotal evidence partially supports this claim—adults get absorbed when playing their favorite games and experience a loss of time when engaged in game play (Wood et al., 2007). However, the instructional benefits of simulation games would be maximized if trainees were also motivated to utilize the knowledge and skills taught in simulation games on the job. Confirming that simulation games enhance work-related motivation is a critical area for future research.

Research is also needed to assess the level of cognitive load imposed by simulation games and techniques for ensuring that the cognitive load does not exceed the cognitive capacity of trainees. Simulation games tend to utilize discovery learning environments (Munro, 2008), and both Mayer (2004) and Sweller (1999, 2004) have acknowledged that discovery learning imposes a heavy cognitive load on trainees. Simulation game players must make numerous choices, recall game rules, and develop simulation game strategies while also increasing their work-related knowledge (Tobias & Fletcher, 2007). Researchers should investigate how guidance and advanced organizers can be incorporated in simulation games to increase instructional effectiveness and minimize the cognitive load of training. Furthermore, active learning theory suggests that core training design elements (e.g., exploration and training frame) directly influence self-regulatory processes and indirectly influence adaptive transfer (Bell & Kozlowski, 2008). It is possible that incorporating these design elements in simulation games will enhance training effectiveness, and this is an essential avenue for future research.

Additional research is also needed to examine the utility of other advanced training technologies. For example, how can organizations incorporate virtual worlds in training to enhance learning outcomes? Furthermore, peer production of training content (e.g., Youtube, Wikipedia) is becoming increasingly commonplace (Brown & Sitzmann, 2011). How can an organization ensure that the information exchanged is accurate and fosters the organization's best interest? Finally, intelligent tutoring systems are becoming more practical to develop and deploy. Research is needed to examine whether trainees feel the same level of connection with an intelligent tutor as with a human tutor such that they will engage in the conversation and reach a deeper level of understanding of the training material (Sitzmann & Ely, 2008).

Conclusion

Simulation games have the potential to enhance the learning of work-related knowledge and skills. Overall, declarative knowledge was 11% higher for trainees taught with simulation games than a comparison group; procedural knowledge was 14% higher; retention was 9% higher; and self-efficacy was 20% higher. Characteristics of simulation games and the instructional context were instrumental in determining the amount that trainees learned from simulation games relative to a comparison group. Specifically, learning from simulation games was maximized when trainees actively rather than passively learned work-related competencies during game play, trainees could choose to play as many times as desired, and simulation games were embedded in an instructional program rather than serving as stand-alone instruction. The ultimate goal for simulation game design teams is to exploit the motivational capacity of simulation games to enhance employees' work-related skills. Thus, additional research is needed to examine the dynamic interplay of affective and cognitive processes during game play and, ultimately, their effect on training transfer.

REFERENCES

References marked with an asterisk indicate studies included in the meta-analysis.

Aldrich C. (2007, July). Engaging mini-games find niche in training. *T+D,* 22–24.

Baldwin TT, Ford JK. (1988). Transfer in training: A review and directions for future research. PERSONNEL PSYCHOLOGY, *41*, 63–105.

Bandura A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191–215. doi:10.1037/0033-295X.84.2.191

Bandura A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes, 50*, 248–287.

Bandura A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*, 117–148.

Bandura A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

*Bayrak C. (2008). Effects of computer simulations programs on university students' achievement in physics. *Turkish Online Journal of Distance Education, 9*, Article 3, 53–62. Retrieved from http://tojde.anadolu.edu.tr/tojde32/pdf/article_3.pdf

Begg CB. (1994). Publication bias. In Cooper HM, Hedges LV (Eds.), *The handbook of research synthesis* (pp. 399–409). New York: Russell Sage.

Bell BS, Kanar AM, Kozlowski SWJ. (2008). Current issues and future directions in simulation-based training in North America. *The International Journal of Human Resource Management, 19*, 1416–1436.

Bell BS, Kozlowski SWJ. (2002). Adaptive guidance: Enhancing self-regulation, knowledge, and performance in technology-based training. PERSONNEL PSYCHOLOGY, *55*, 267–306.

Bell BS, Kozlowski SWJ. (2008). Active learning: Effects of core training design elements on self-regulatory processes, learning, and adaptability. *Journal of Applied Psychology, 93*, 296–316.

*Blunt R. (2007, November). *Does game-based learning work? Results from three recent studies*. Paper presented at the Interservice/Industry Training, Simulation & Education Conference (I/ITSEC), Orlando, FL.

*Bower RW. (1998). *An investigation of a manipulative simulation in the learning of recursive programming*. Unpublished doctoral dissertation, Iowa State University, Ames.

*Boyd BL, Murphrey TP. (2002). Evaluation of a computer-based, asynchronous activity on student learning of leadership concepts. *Journal of Agricultural Education, 43*, 36–45.

Brannick MT, Prince C, Salas E. (2005). Can PC-based systems enhance teamwork in the cockpit? *The International Journal of Aviation Psychology, 15*, 173–187.

*Brant G, Hooper E, Sugrue B. (1991). Which comes first the simulation or the lecture? *Journal of Educational Computing Research, 7*, 469–481.

Brown KG, Ford JK. (2002). Using computer technology in training: Building an infrastructure for active learning. In Kraiger K (Ed.), *Creating, implementing, and maintaining effective training and development: State-of-the-art lessons for practice* (pp. 192–233). San Francisco: Jossey-Bass.

Brown, KG, Sitzmann, T. (2011). Training and employee development for improved performance. In S. Zedeck (Ed.), *Handbook of industrial and organizational psychology* (Vol. 2, pp. 469–503). Washington, DC: APA.

Bruner J. (1962). *On knowing: Essays for the left hand*. Cambridge, MA: Harvard University Press.

*Cameron B, Dwyer F. (2005). The effect of online gaming, cognition and feedback type in facilitating delayed achievement of different learning objectives. *Journal of Interactive Learning Research, 16*, 243–258.

*Cataloglu E. (2006). Open source software in teaching physics: A case study on vector algebra and visual representations. *The Turkish Online Journal of Educational Technology, 5*, Article 8, 68–74. Retrieved from http://www.tojet.net/articles/518.pdf

Chen H-HC, O'Neil HF. (2008). A formative evaluation of the training effectiveness of a computer game. In O'Neil HF, Perez RS (Eds.), *Computer games and team and individual learning* (pp. 39–54). Oxford, UK: Elsevier.

*Chou C-H. (1998). *The effectiveness of using multimedia computer simulations coupled with social constructivist pedagogy in a college introductory physics classroom*. Unpublished doctoral dissertation, Columbia University, New York.

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

Clark RE. (1994). Media will never influence learning. *Educational Technology Research and Development, 42*, 21–29.

CNET Networks Entertainment. (2004, January 23). The greatest games of all time: The Oregon Trail. Retrieved from http://www.gamespot.com/gamespot/features/all/greatestgames/p-34.html

Cook TD, Campbell DT. (1979). *Quasi-experimentation: Design and analysis for field settings*. Chicago: Rand McNally.

Cordova DI, Lepper MR. (1996). Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology, 88*, 715–730.

Csikszentmihalyi M. (1990). *Flow: The psychology of optimal experience*. New York: Harper & Row.

Deci E, Ryan R. (1985). The support of autonomy and control of behavior. *Journal of Personality and Social Psychology, 53*, 1024–1037.

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

*Desrochers MN, Clemmons T, Grady M, Justice B. (2000). Practice in behavioral assessment and treatment decisions. *Journal of Technology in Human Services, 17*, 15–27.

Dewey J. (1938). *Experience and education*. New York: MacMillan.

Driskell JE, Dwyer DJ. (1984). Microcomputer videogame based training. *Educational Technology, 24*, 11–17.

*Ebner M, Holzinger A. (2007). Successful implementation of user-centered game based learning in higher education: An example from civil engineering. *Computers & Education, 49*, 873–890.

*Ellinger R-S, Frankland P. (1976). Computer-assisted and lecture instruction: A comparative experiment. *Journal of Geography, 75*, 105–120.

*Farrell MJ, Arnold P, Pettifer S, Adams J, Graham T, MacManamon M. (2003). Transfer of route learning from virtual to real environments. *Journal of Experimental Psychology: Applied, 9*, 219–227.

*Faryniarz JV, Lockwood LG. (1992). Effectiveness of microcomputer simulations in simulating environmental problem solving by community college students. *Journal of Research in Science Teaching, 29*, 453–470.

Federation of American Scientists. (2006). *Summit on educational games: Harnessing the power of video games for learning*. Washington, DC: Author.

Fisher SL, Ford JK. (1998). Differential effects of learner effort and goal orientation on two learning outcomes. PERSONNEL PSYCHOLOGY, *51*, 397–420.

Fleiss JL. (1981). *Statistical methods for rates and proportions*. New York: John Wiley.

Fletcher JD, Tobias S. (2006, February). *Using computer games and simulations for instruction: A research review*. Paper presented at the Society for Applied Learning Technology Meeting, New Learning Technologies, Orlando, FL.

*Frear V, Hirschbuhl JJ. (1999). Does interactive multimedia promote achievement and higher level thinking skills for today's science students? *British Journal of Educational Technology, 30*, 323–329.

Frese M, Brodbeck FC, Heinbokel T, Mooser C, Schleiffenbaum E, Thiemann P. (1991). Errors in training computer skills: On the positive function of errors. *Human Computer Interaction, 6*, 77–93.

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

Glass GV, McGaw B, Smith ML. (1981). *Meta-analysis in social research*. Newbury Park, CA: Sage.

*Gokhale A. (1991). Effectiveness of computer simulation versus lab and sequencing of instruction in teaching logic circuits. *Journal of Industrial Teacher Education, 28*, 22–23.

*Gopher D, Weil M, Bareket T. (1994). Transfer of skill from a computer game trainer to flight. *Human Factors, 36*, 387–405.

*Gorrell J, Downing H. (1989). Effects of computer-simulated behavior analysis on preservice teacher's problem solving. *Journal of Educational Computing Research, 5*, 335–347.

Gredler ME. (1996). Educational games and simulations: A technology in search of a research paradigm. In Jonassen DH (Ed.), *Handbook of research for educational communications and technology* (pp. 521–539). New York: MacMillan.

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

Hedges LV, Olkin I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

*Holzinger A, Kickmeier-Rust MD, Wassertheurer S, Hessinger M. (2009). Learning performance with interactive simulations in medical education: Lessons learned from results of learning complex physiological models with the HAEMOdynamics SIMulator. *Computers & Education, 52*, 292–301.

*Hsu Y-S. (1997). *Development and formative evaluation of an instructional simulation of adiabatic processes*. Unpublished doctoral dissertation, Iowa State University, Ames.

Huffcutt AI, Arthur W. (1995). Development of a new outlier statistic for meta-analytic data. *Journal of Applied Psychology, 80*, 327–334.

*Hughes IE. (2001). Alternatives to laboratory practicals—Do they meet the needs? *Innovations in Education and Teaching International, 38*, 3–7.

Hunter JE, Schmidt FL. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage.

*Ivers K, Barron AE. (1994, November). *Teaching telecommunications: A comparison between video and computer-based instruction*. Paper presented at TelEd '94: Third International Conference on Telecommunications in Education, Albuquerque, NM.

Jana R. (2006, March 27). On-the-job video gaming. *Business Week, 3977*, 43.

Jarmon L, Traphagan T, Mayrath M. (2008). Understanding project-based learning in second life with a pedagogy, training, and assessment trio. *Educational Media International, 45*, 153–171.

Jonassen D. (2002). Integration of problem solving into instructional design. In Reiser R, Dempsey J (Eds.), *Trends and issues in instructional design and technology* (pp. 107–120). Upper Saddle River, NJ: Merrill/Prentice Hall.

Keith N, Frese M. (2005). Self-regulation in error management training: Emotion control and metacognition as mediators of performance effects. *Journal of Applied Psychology, 90*, 677–691.

Keith N, Frese M. (2008). Effectiveness of error management training: A meta-analysis. *Journal of Applied Psychology, 93*, 59–69.

*Kim JH, Kim WO, Min KT, Yang JY, Nam YT. (2002). Learning by computer simulation does not lead to better test performance on advanced cardiac life support than textbook study. *Journal of Clinical Anesthesia, 14*, 395–400.

*Kinkade RE, Matthews CT, Draugalis JR, Erstad BL. (1995). Evaluation of a computer simulation in a therapeutics case discussion. *American Journal of Pharmaceutical Education, 59*, 147–150.

*Korfiatis K, Papatheodorou E, Stamou GP. (1999). An investigation of the effectiveness of computer simulation programs as tutorial tools for teaching population ecology at university. *International Journal of Science Education, 21*, 1269–1280.

*Kos I. (1996). *Teaching clinically oriented embryology with computer simulations*. Unpublished doctoral dissertation, New York University, New York.

Kozlowski SWJ, Bell BS. (2006). Disentangling achievement orientation and goal setting: Effects on self-regulatory processes. *Journal of Applied Psychology, 91*, 900–916.

Kraiger K, Ford JK, 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.

Kraiger K, Jerden E. (2007). A meta-analytic investigation of learner control: Old findings and directions. In Fiore SM, Salas E (Eds.), *Toward a science of distributed learning* (pp. 65–90). Washington, DC: APA.

Kulik CC, Kulik JA. (1991). Effectiveness of computer-based instruction: An updated analysis. *Computers in Human Behavior, 7*, 75–94.

Kulik JA. (1994). Meta-analytic studies of findings on computer-based instruction. In Baker EL, O'Neil HF Jr (Eds.), *Technology assessment in education and training* (pp. 9–33). Hillsdale, NJ: Erlbaum.

*Lawson KA, Shepherd MD, Gardner VR. (1991). Using a computer-simulation program and a traditional approach to teach pharmacy financial management. *American Journal of Pharmaceutical Education, 55*, 226–235.

Lee J. (1999). Effectiveness of a computer-based instructional simulation: A meta-analysis. *International Journal of Instructional Media, 26*, 71–85.

Liao Y-K. C. (1999). Effects of hypermedia on students' achievement: A meta-analysis. *Journal of Educational Multimedia and Hypermedia, 8*, 255–277.

Lipsey MW. (2003). Those confounded moderators in meta-analysis: Good, bad, and ugly. *The Annals of the American Academy of Political and Social Science, 587*, 69–81.

Lipsey MW, Wilson DB. (1993). The efficacy of psychological, educational, and behavioral treatment. *American Psychologist, 48*, 1181–1201.

Lipsey MW, Wilson DB. (2001). *Applied social research methods series: Vol. 49. Practical meta-analysis*. Thousand Oaks, CA: Sage.

*Loar RS. (2007). *The impact of a computer simulated case study on nurse practitioner students' declarative knowledge and clinical performance*. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.

Malone TW. (1981). Toward a theory of intrinsically motivating instruction. *Cognitive Science, 4*, 333–369.

Mayer RE. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *American Psychologist, 59*, 14–19.

Meyers SC, Strang HR, Hall EL. (1989). Impact of microcomputer simulation training on student-clinicians' ability to effectively intervene with preschool stutterers. *Journal of Fluency Disorders, 14*, 135–151.

Miller CC, Glick WH, Wang YD, Huber GP. (1991). Understanding technology structure relationships: Theory development and meta-analytic theory testing. *Academy of Management Journal, 34*, 370–399.

*Mitchell RC. (2004). Combining cases and computer simulations in strategic management courses. *Journal of Education for Business, 79*, 198–204.

*Moshirnia AV. (2008). *The influence of information delivery systems in modified video games on learning*. Unpublished doctoral dissertation, University of Kansas, Lawrence.

*Moslehpour S. (1993). *A comparison of achievement resulting from learning electronics concepts by computer simulation versus traditional laboratory instruction*. Unpublished doctoral dissertation, Iowa State University, Ames.

Munro A. (2008). Foundations for software support of instruction in game contexts. In O'Neil HF, Perez RS (Eds.), *Computer games and team and individual learning* (pp. 39–54). Oxford, UK: Elsevier.

Newell A, Rosenbloom PS, Laird JE. (1989). Symbolic architectures for cognition. In Posner MI (Ed.), *Foundations of cognitive science* (pp. 93–131). Cambridge, MA: Bradford Books/MIT Press.

Noe RA. (1986). Trainees' attributes and attitudes: Neglected influences on training effectiveness. *Academy of Management Review, 11*, 736–749.

*North MM, Sessum J, Zakalev A. (2003). Immersive visualization tool for pedagogical practices of computer science concepts: A pilot study. *JCSC, 19*, 207–215.

Northup PT. (2002). Online learners' preference for interaction. *Quarterly Review of Distance Education, 3*, 219–226.

O'Neil HF, Wainess R, Baker EL. (2005). Classification of learning outcomes: Evidence from the computer games literature. *The Curriculum Journal, 16*, 455–474.

*Ortiz GA. (1994). Effectiveness of PC-based flight simulation. *The International Journal of Aviation Psychology, 4*, 285–291.

Orvis KA, Horn DB, Belanich J. (2008). The roles of task difficulty and prior videogame experience on performance and motivation in instructional videogames. *Computers in Human Behavior, 24*, 2415–2433.

*Parchman SW, Ellis JA, Christinaz D, Vogel M. (2000). An evaluation of three computer-based instructional strategies in basic electricity and electronics training. *Military Psychology, 12*, 73–87.

Parker LE, Lepper MR. (1992). Effects of fantasy contexts on children's learning and motivation: Making learning more fun. *Journal of Personality and Social Psychology, 62*, 625–633.

Piaget J. (1951). *Psychology of intelligence*. London: Routledge and Kegan Paul.

Prensky M. (2001). *Digital game-based learning*. New York: McGraw-Hill.

Randel JM, Morris BA, Wetzel CD, Whitehill BV. (1992). The effectivenses of games for educational purposes: A review of recent research. *Simulation & Gaming, 23*, 261–276.

*Randell T, Hall M, Bizo L, Remington B. (2007). DTkid: Interactive simulation software for training tutors of children with autism. *Journal of Autism and Developmental Disorders, 37*, 637–647.

Rapp TL, Mathieu JE. (2007). Evaluating an individually self-administered generic teamwork skills training program across time and levels. *Small Group Research, 38*, 532–555.

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

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

*Ross WH, Jr, Pollman W, Perry D, Welty J, Jones K. (2001). Interactive video negotiator training: A preliminary evaluation of the McGill negotiation simulator. *Simulation & Gaming, 32*, 451–468.

*Rowland PM. (1988). *The effect of two modes of computer-assisted instruction and individual learning differences on the understanding of science concept relationships*. Unpublished doctoral dissertation, New Mexico State University, Las Cruces.

*Rowland PM, Stuessy CL. (1987, April). *Effects of modes of computer-assisted instruction on conceptual understanding and achievement of college students exhibiting individual differences in learning: A pilot study*. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, Washington, DC.

Salas E, Bowers CA, Cannon-Bowers JA. (1995). Military team research: 10 years of progress. *Military Psychology, 7*, 55–75.

Salas E, Bowers CA, Rhodenizer L. (1998). It's not how much you have but how you use it: Toward a rational use of simulation to support aviation training. *The International Journal of Aviation Psychology, 8*, 197–208.

Salas E, Cannon-Bowers JA. (1997). Methods, tools, and strategies for team training. In Quiñones MA, Ehrenstein A (Eds.), *Training for a rapidly changing workplace: Applications of psychological research* (pp. 249–279). Washington, DC: APA.

Salas E, Cannon-Bowers JA. (2001). The science of training: A decade of progress. *Annual Review of Psychology, 52*, 471–499.

Schenker JD. (2007). *The effectiveness of technology use in statistics instruction in higher education: A meta-analysis using hierarchical linear modeling*. Unpublished doctoral dissertation, Kent State University, Ohio.

*Schloss PJ, Cartwright GP, Smith MA, Polka BA. (1986). Effectiveness of computer-assisted special education administration simulations. *The Journal of Special Education, 20*, 503–513.

Schollmeyer J. (2006). Games get serious. *Bulletin of the Atomic Scientists, 62*, 34–39.

*Schwid HA, Rooke GA, Ross BK, Sivarajan M. (1999). Use of a computerized advanced cardiac life support simulator improves retention of advanced cardiac life support guidelines better than a textbook review. *Critical Care Medicine, 27*, 821–824.

*Shute VJ, Glaser R. (1990). A large-scale evaluation of an intelligent discovery world: Smithtown. *Interactive Learning Environments, 1*, 51–77.

Siemer J, Angelides MC. (1995). Evaluating intelligent tutoring with gaming simulations. *Proceedings of the 1995 Winter Simulation Conference,* 1376–1382.

Simons R-J. (1993). *The learner as designer of his own learning environment*. Unpublished manuscript, University of Nijmegen, Netherlands.

Sitzmann T, Brown KG, Casper W, Ely K, Zimmerman RD. (2008). A review and meta-analysis of the nomological network of trainee reactions. *Journal of Applied Psychology, 93*, 280–295.

Sitzmann T, Ely K. (2008). Do second-generation models have more to offer? *Industrial and Organizational Psychology, 1*, 494–495.

Sitzmann T, Ely K. (2010). Sometimes you need a reminder: The effects of prompting self-regulation on regulatory processes, learning, and attrition. *Journal of Applied Psychology, 95*, 132–144.

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

Sitzmann T, Ely K, Wisher R. (2008). Designing Web-based training courses to maximize learning. In Orvis KL, Lassiter AR (Eds.), *Computer supported collaborative learning: Best practices and principles for instructors* (pp. 1–19). Hershey, PA: Idea Group.

Sitzmann T, Kraiger K, Stewart D, Wisher R. (2006). The comparative effectiveness of Web-based and classroom instruction: A meta-analysis. PERSONNEL PSYCHOLOGY, *59*, 623–664.

Slagle M. (2006, May 8). Poll: 4 in 10 Americans play video games [Electronic Version]. *Washington Post*. Retrieved from http://www.washingtonpost.com/wpdyn/content/article/2006/05/07/AR2006050700172.html.

Steel PD, Kammeyer-Mueller JD. (2002). Comparing meta-analytic moderator estimation techniques under realistic conditions. *Journal of Applied Psychology, 87*, 96–111.

*Sterling J, Gray MW. (1991). The effect of simulation software on students' attitudes and understanding in introductory statistics. *Journal of Computers in Mathematics and Science Teaching, 10*, 51–56.

*Stockburger DW. (1982). Evaluation of three simulation exercises in an introductory statistics course. *Contemporary Educational Psychology, 7*, 365–370.

*Sukhai DR. (2005). *The effects of Web-based animated demonstration and guided simulation on time and performance in business software applications training*. Unpublished doctoral dissertation, University of South Alabama, Mobile.

Summers GJ. (2004). Today's business simulation industry. *Simulation & Gaming, 35*, 208–241.

*Surface EA, Dierdorff EC, Watson AM, Cahoon M. (2009). *Relative effectiveness of Tactical Iraqi(tm) training and lecture-based classroom training* (Tech. Rep. No. 2009010606). MacDill Air Force Base, FL: Special Operations Forces Language Office, United States Special Operations Command.

Sweller J. (1999). *Instructional design in technical areas*. Camberwell, Australia: ACER Press.

Sweller J. (2004). Instructional design consequences of an analogy between evolution by natural selection and human cognitive architecture. *Instructional Science, 32*, 9–31.

T+D. (2007, June). What do simulations cost? *T+D, 88*.

*Taylor RS, Chi MTH. (2006). Simulation versus text: Acquisition of implicit and explicit information. *Journal of Educational Computing Research, 35*, 289–313.

Tennyson RD, Breuer K. (1997). Psychological foundations for instructional design theory. In Tennyson R, Schott F, Seel N, Dijkstra S (Eds.), *Instructional design: International perspectives: Vol. 1. Theory, research, and models* (pp. 113–134). Mahwah, NJ: Erlbaum.

Tennyson RD, Jorczak RL. (2008). A conceptual framework for the empirical study of instructional games. In O'Neil HF, Perez RS (Eds.), *Computer games and team and individual learning* (pp. 39–54). Oxford, UK: Elsevier.

Tobias S, Fletcher JD. (2007). What research has to say about designing computer games for learning. *Educational Technology, 47*, 20–29.

*Tompson GH, Daas P. (2000). Improving students' self-efficacy in strategic management: The relative impact of cases and simulations. *Simulation & Gaming, 31*, 22–41.

*Upah S, Thomas P. (1993). An investigation of manipulative models on the learning of programming loops. *Journal of Educational Computer Research, 9*, 397–412.

Venkatesh V, Speier C. (2000). Creating an effective training environment for enhancing telework. *International Journal of Human Computer Systems, 52*, 991–1005.

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

*Waddick J. (1994). Case study: The use of a hypercard simulation to aid in the teaching of laboratory apparatus operation. *Educational and Training Technology International, 31*, 295–301.

Webster J, Hackley P. (1997). Teaching effectiveness in technology-mediated distance learning. *Academy of Management Journal, 40*, 1282–1309.

Whitener EM. (1990). Confusion of confidence intervals and credibility intervals in meta-analysis. *Journal of Applied Psychology, 75*, 315–321.

*Willis DAS. (1989). *Development and evaluation of IRIS: A computer simulation to teach preservice teachers to administer a formal reading inventory*. Unpublished doctoral dissertation, Texas Tech University, Lubbock.

Wilson KA, Bedwell WL, Lazzara EH, Salas E, Burke CS, Estock JL, . . . Conkey C. (2009). Relationships between game attributes and learning outcomes: Review and research proposals. *Simulation & Gaming, 40*, 217–266.

Wood RTA, Griffiths MD, Parke A. (2007). Experiences of time loss among videogame players: An empirical study. *CyberPsychology & Behavior, 10*, 38–44.

*Yares A. (2005). *The effects of diagnosis training software on the diagnostic fluency and self-confidence of school psychology graduate trainees*. Unpublished doctoral dissertation, Temple University, Philadelphia, PA.

*Yip FWM, Kwan ACM. (2006). Online vocabulary games as a tool for teaching and learning English vocabulary. *Educational Media International, 43*, 233–249.

*Zacharia ZC. (2007). Comparing and combining real and virtual experimentation: An effort to enhance students' conceptual understanding of electric circuits. *Journal of Computer Assisted Learning, 23*, 120–132.

Zhao Y, Lei J, Lai BYC, Tan HS. (2005). What makes the difference? A practical analysis of research on the effectiveness of distance education. *Teachers College Record, 107*, 1836–1884.