Running head: ASSESSMENT IN INSTRUCTIONAL GAMES

The Assessment of Learning in Instructional Games and Simulations

Rod Myers

Indiana University

R711 – Readings in Instructional Systems Technology

Professor Brush

December 13, 2007

Abstract

Games and simulations are of great interest to educators, in part because of their proven ability to engage and motivate players to recognize and solve difficult problems in situated contexts. They are compatible with many contemporary theories of learning and related methods of assessment. This review summarizes the research on the methods and difficulties of assessing learning outcomes from the use of games and simulations for instruction. Games and simulations are defined and their salient characteristics are described. They are then considered in terms of their purposes and functions within an educational context, followed by an analysis of the expected learning outcomes and methods of assessment. The review ends by addressing gaps in the literature and proposing beneficial areas for research.

The Assessment of Learning in Instructional Games and Simulations

The field of instructional systems technology has its origins in the visual instruction movement of the early twentieth century, and for many decades it was defined in terms of instructional media. One enduring hallmark of the field is the willingness of designers to experiment with new media and technologies and to study their effectiveness for teaching and learning. Each new medium—from film to radio and sound recording to television to computers and other digital devices—presents it own possibilities and challenges to the instructional designer, and each influences the theories and practices in the field.

Digital games are the latest mass medium of popular culture, and interest in their study and use for learning has visibly increased in the last decade. The digital game industry now has annual revenues greater than the movie industry, leading some to identify a Games Generation (Prensky, 2001). The majority of today’s learners (K-12 and traditional first-time college students) have grown up with computers, video games, mobile phones, and portable media players. As a result, they are used to a variety of stimuli, they are adept multi-taskers, and they are enthusiastic adopters of new technologies. This generation’s cognitive style differs from previous generations in its preference for speed, multiple tasks, graphics over text, immediacy, exploration, and connection with others (Prensky, 2001). In 2003 the Pew Internet and American Life Project released the results of a survey of college students. Everyone surveyed had played a video, computer, or online game. Seventy percent played at least once in a while, and 65% played occasionally or regularly. About one in ten admitted that playing was a way of avoiding studying. One third admitted to playing games that were not part of instructional activities during class. A majority (69%) reported no exposure to gaming for educational purposes in the classroom (Jones, 2003).

Teaching and learning in schools has changed little in the past one hundred years. As a result, learning in the classroom bears little resemblance to learning outside the classroom, leading to a greater degree of disengagement among learners. According to the most recent High School Survey of Student Engagement, 2 out of 3 students report that they are bored in class every day, and 17 percent say they are bored in every class (Yazzie-Mintz, 2007). Educators need to utilize the "cultural tools" of contemporary society to engage students in ways that are familiar to them (Strommen, 1992).

The purpose of this review is to summarize the research on the methods and difficulties of assessing learning outcomes from the use of games and simulations for instruction. To provide a foundation, games and simulations are defined and their salient characteristics are described. Games and simulations are then considered in terms of their purposes and functions within an educational context, followed by an analysis of the expected learning outcomes and methods of assessment. The review ends by addressing gaps in the literature and proposing beneficial areas for research.

Background and Definitions

A Brief History

The educational use of games and simulations dates back to war games in the seventeenth century and military training in the eighteenth century (Egenfeldt, 2005; Gredler, 2004). In the 1950s the practice was adapted for business management training, and in 1956 the American Management Association produced the first business game (Leemkuil, de Jong, & Ootes, 2000). As the capabilities of computing technologies have grown, increasingly complex and sophisticated games and simulations have been used for instruction in a variety of content areas, including medical education, the natural and social sciences, and corporate training. Cruickshank

(1988) described several media-based (audio-visual) simulations from the 1960s and 1970s designed for preservice teachers. He noted that in the 1980s computer-based simulations became more prevalent and also more specialized with regard to content or focus. He cited as an example William Harless of the National Library of Medicine, who developed “an interactive videodisk-based simulation designed to teach clinical problem solving by enabling voice input to ask questions [and] order lab tests” (p. 151). In 1987, Faria (cited in Dempsey, Lucassen, Gilley, & Rasmussen, 1993-1994) reported that a survey revealed that 8,755 instructors in 1,900 business schools used business games in their courses.

Research on the use of simulations and games for learning seems to be increasing. Rutter and Bryce (2006) compared the periods of 1995-1999 and 2000-2004 and found nearly twice as many peer-reviewed papers on digital games during the latter period. Bragge and Storgards (2007) used the ISI Web of Science to find 2,100 studies in more than 170 categories related to digital games between 1986 and 2006, with a significant increase beginning in 2003. However, much of the reporting on the use of games for learning is anecdotal, descriptive, or judgmental and not tied to theory or rigorous research (Gredler, 2004; Kirriemuir & McFarlane, 2003; Leemkuil et al., 2000; Washbush & Gosen, 2001; Wideman et al., 2007). A review of the literature by Dempsey, Rasmussen, and Lucassen (1996) consisted of 99 sources from the past 12 years. Building on an earlier article (Dempsey et al., 1993-94), the authors defined five categories of gaming articles: discussion (n=51), research (n=38), reviews (n=12), theory (n=11), and development (n=2).

There have been relatively few studies of the use of games for learning in K-12 settings, and these have been primarily case studies, often involving students’ and teachers’ perceptions of learning (McFarlane, Sparrowhawk, & Heald, 2002; Wideman et al, 2007). Based on their review of the literature, Wideman et al. (2007) concluded that disciplines with the most research in educational gaming are medical education and business management studies. Bragge and Storgards (2007) combined the 170 categories found in their review into larger domains to find that the three most prominent areas were social sciences, health sciences, and information and communication technologies and mathematics.

Definitions and Characteristics

Games and simulations. A variety of definitions for "game" and "simulation" are presented in the literature. Wolfe and Crookall (1998) note that despite several decades as a field, researchers and practitioners in simulation and gaming are still grappling to create a generally accepted taxonomy. Gredler (2004) defines games as "competitive exercises in which the objective is to win and players must apply subject matter or other relevant knowledge in an effort to advance in the exercise and win," while simulations are "open-ended evolving situations with many interacting variables … in which the participants take on bona fide roles with well-defined responsibilities and constraints" (p. 571). For Garris, Ahlers, and Driskell (2002), the key distinction is that a simulation represents reality and a game does not. However, a simulation may contain game-like features and may become a game if a performance goal is set. For example, SimCity is an endless simulation of a city. But if the player sets a particular goal, the simulation turns into a game. Salen and Zimmerman (2004) reviewed many of the major writers on games and simulations and synthesized their definitions: "A *game* is a system in which players engage in an artificial conflict, defined by rules, that results in a quantifiable outcome" (p. 80) and "[a] *simulation* is a procedural representation of aspects of 'reality'" (p. 423). They contend that some simulations are not games but that all games are some form of simulation.

Similarly, in their proposed taxonomy of simulations, Maier and Grossler (2000) distinguish between modeling-oriented and gaming-oriented simulations, the latter being single-user simulators or multi-user planning games. However, others have offered different schemas. Leemkuil et al. (2000) categorize simulations based on underlying models: conceptual models based on principles, concepts, and facts related to a system and operational models based on sequences of procedures applied to a system. Gredler (2004) distinguishes types of simulations based on participant roles. Experiential simulations are social microcosms in which learners take on roles and responsibilities and participate in authentic scenarios and tasks. In symbolic simulations the learner is not a functional component of the system but operates on the system, testing her conceptual model of the relationships among the variables of the system.

Games are often categorized by their genres, which include action, adventure, strategy, role-playing, racing, sports, shooting, word games, and puzzles. However, games may also be grouped by medium, such as board games, card games, video games, and miniature war games. Salen and Zimmerman (2004) even categorize games based on their underlying systems—emergent systems, information theory systems, cybernetic systems, systems of conflict, etc.

Perhaps Parlett (1999) is correct in insisting that “game” is a “slippery lexicological customer” (p. 1) and that there is no use in trying to propose a single definition. Rather than trying to define the distinctions between games and simulations, Aldrich (2005) pragmatically suggests that it is more productive to think about elements of the instructional experience: simulation elements, game elements, and pedagogy elements. For the purposes of this review, the distinction between games and simulations will be considered irrelevant, and they will be considered as interchangeable if not exactly synonymous. Instead the focus is on what make games and simulations effective modes of instruction and how their effectiveness is measured.

Fidelity and validity. Fidelity is the degree to which a simulation is faithful to that which it simulates. In discussing the development of physical skills, Romiszowski (1999) distinguishes between "technical fidelity" as defined by an expert and "perceived fidelity" as that which is simple enough to train a novice without undue complications. Reigeluth and Schwartz (1989) theorized that the most fundamental aspects of a simulation should have high fidelity, while lower fidelity is appropriate for the more superficial aspects that may otherwise lead to cognitive overload and impede learning and transfer. They suggested that factors to consider include the complexity of the real world environment, the potential for transfer, the motivational consequence of high fidelity, and the expense of achieving high fidelity. More recently, Feinstein and Cannon (2002) examined numerous studies from the 1960s and 1970s that focused on the effects of fidelity on training and education. They report that greater fidelity did not result in greater learning and may in fact reduce effectiveness through unnecessary complexity and overstimulation. Similarly, Winn (2002) notes that a virtual environment does not need to simulate the real world to be useful for instruction, and that high fidelity may lead to constrained and inflexible understanding and make it difficult to transfer knowledge and skills to new contexts.

Verification is "the process of assessing that a model is operating as intended" whereas validation is "the process of assessing that the conclusions reached from a simulation are similar to those reached in the real-world system being modeled" (Feinstein & Cannon, 2002, p. 427). The former is an aspect of internal validity and the latter is an aspect of external validity, and both relate to the development and the use of the simulation. In designing and developing the simulation, representational validation encompasses the game logic and structure (internal, construct validity) and the phenomena being modeled (external, content validity). In using the simulation, educational validation results in student insight (internal, conceptual validity) and demonstration of student learning (external, criterion validity) (Feinstein & Cannon, 2002).

Gosen and Washbush (2004) reviewed studies and meta-analyses that addressed the validity of business simulations. They concluded that there is not enough methodologically sound evidence to claim general validity for simulations. However, this may say more about the quality of the research that has been done than about the validity of the simulations.

The Use of Games and Simulations for Instruction

This section reviews the rationale and methods for using games in educational settings in order to provide a context for examining the ways in which learning is subsequently assessed. The traditional instructional paradigm through the 1960s was information transfer from a "knowledgeable educator who constructed and transmitted knowledge … using the accepted instructional technologies of the day—books, articles, and lectures" (Ruben, 1999, p. 498). Foreman (2004) notes that such a model is based on scarcity of quality materials and instructors. Furthermore this framework implies that teaching is a prerequisite for learning; it ignores the social, collaborative, and peer based nature of learning outside the classroom (Ruben, 1999). As educators explored more experience-based approaches to instruction—such as case studies, role playing, simulations, games, and other structured exercises—the traditional, didactic model gradually ceded prominence to a learner-centered model emphasizing active, experiential learning (Ruben, 1999; Garris et al., 2002). This approach accommodates more complex and diverse approaches to learning by allowing greater interactivity, collaboration, and peer learning. Ruben (1999) notes that what is important in learning is translating knowledge into behavior, which requires "reinforcement, application, repetition, and often practice in a variety of settings and contexts" (p. 499).

Why Use Games to Teach?

Two main reasons for using instructional games are their power to engage and motivate and their ability to facilitate learning through doing (Kirriemuir & McFarlane, 2004). According to Garris et al. (2002), there are several reasons why educators should be interested in using games in instruction, including the shift to a learner-centered model and the intensity of involvement and engagement in games. The memorization of facts and concepts that is easily measured on a standardized test has led to the presentation of abstract, decontextualized knowledge that is divorced from purpose and instrumentality. In contrast, games require players constantly to use what they have learned to solve situated problems (Shaffer, Squire, Halverson, & Gee, 2005; Wideman et al., 2007). Findings demonstrate that the kinds of experiential learning available in games improve learners' problem-solving skills and judgment (Feinstein & Cannon, 2002). Games can serve as immersive learning environments conducive to experiential learning and can encourage exploration along the lines of guided discovery. Instead of reading about something students can experience it. Children have also shown learning gains using games in content areas with specific stated objectives, such as math and language skills. In part this is because the active learning required in games facilitates integration of knowledge with existing cognitive structures (Randel, Morris, Wetzel, & Whitehill, 1992).

In their review of the literature, Mitchell and Savill-Smith (2004) found several frequently cited benefits of games in education. These include increases in perseverance, confidence, and self-esteem among learners; the ability to visualize, manipulate, and explore concepts; and greater academic, social, and computer literacy skills. Some studies cited improved metacognition, strategic thinking, problem recognition, and problem solving. In the health sciences, simulations enable students to diagnose and manage virtual patients' problems. In business education, teams manage virtual companies. In both areas, simulations are used to identify students' problem solving abilities and to bridge the gap between classroom instruction and real-world practice (Gredler, 2004).

Many of the attributes of games are also attributes of good instructional design. Games often involve problem solving, provide rapid feedback, and can adjust to optimal level of difficulty (Oblinger, 2003). Gee (2003, 2005) identified dozens of learning principles that are found in good games, including manipulation and control by the learner, scaffolding and elaboration, well-ordered problems, optimal challenge, skills as strategies and cycles of expertise, information as needed (just in time), systems thinking, and learning by doing.

Many studies of the benefits of playing games to learn have emphasized the motivational or social aspects rather than knowledge acquisition (Kafai, 2001). However, intrinsic motivation is generally considered a prerequisite for learning. Garris et al. (2002) describe the motivated learner as enthusiastic, engaged, focused, and persistent. The factors that make an activity intrinsically motivating are challenge, curiosity, and fantasy (Malone, 1981). Not surprisingly, these are all common elements of games. Garris et al. (2002) propose an input-process-output game model that facilitates intrinsic motivation. The input is a combination of instructional content and game features. The features promote a game cycle of user judgments, user behavior, and system feedback in an iterative loop which, when successful, results in increased engagement, greater persistence of effort, and greater likelihood of achieving intended learning outcomes.

The choice to use a game or simulation should be based on "a *detailed analysis of the learning requirements* and an analysis of the *tradeoffs* among alternate instructional approaches" [emphasis in the original] (Hays, 2006, p. 312). In a meta-analysis of studies that compared the instructional effectiveness of games with traditional classroom instruction over 28 years, only 68 empirical studies were found (Randel et al., 1992). Of those, 38 found no differences in effectiveness, 27 found games more effective, and 3 found classroom instruction more effective. However, the authors noted a lack of rigor in the research designs, including a lack of random sampling, failure to report reliability and validity, and failure to control confounding variables. In a quantitative meta-analysis of simulation gaming, Van Sickle (1986) found weak support for games over other approaches. Only five studies found simulation gaming more effective for immediate recall of knowledge and only two studies found that simulation gaming improved retention of knowledge. However, Hays (2006) criticizes Van Sickle’s methodology, noting that 6 of the 22 studies did not compare instructional approaches.

How Games Are Used for Instruction

Kirriemuir and McFarlane (2004) report that the use of mainstream games in K-12 education is and will probably remain rare for several reasons. Evaluating a game’s relevance to curriculum and accuracy of content is difficult and time-consuming. A mainstream game that is applicable to curriculum standards will likely have much irrelevant content. Furthermore, most teachers are not familiar with methods for using games in instruction. de Freitas and Oliver (2006) propose a framework with four dimensions to guide and support the evaluation of educational games: context, learner specification, pedagogic considerations, and mode of representation. However, even with a framework, choosing a game for use in an educational setting takes time and experience.

Gredler (2004) states that the purposes of games and simulations in education are to practice or refine existing knowledge and skills, to identify gaps or weaknesses in knowledge or skills, to develop new relationships among known concepts and principles, and to serve as a summation or review. These are consistent with reviews of the reported use of games, in which games were most frequently used to learn new skills and practice existing skills, generally after the learners had received some introductory instruction to prepare them for the game (Dempsey et al., 1993-1994; Dempsey et al., 1996). Options for integrating games into a curriculum include use as a pre-instructional strategy, a co-instructional strategy, and a post-instructional strategy (for assessment and synthesis) (Oblinger, 2006).

A review of the literature led Leemkuil et al. (2000) to conclude that there is some consensus that games and simulations will not be effective unless accompanied by instructional support, such as model progression, prompting, feedback (from the game/simulation or the instructor or peers), debriefing, and reflection. Gredler (2004) concurs that open-ended, discovery learning in a simulation is problematic. She recommends that students acquire required knowledge and capabilities (including metacognitive skills) prior to using a simulation. Research consistently concludes that students need some structure in order to learn in discovery-oriented simulations. Rieber (2005) recommends short explanations offered at the appropriate times within the simulation. He also suggests model progression in which the simulation becomes increasingly difficult based on the learner's mastery of required skills.

Many researchers of games and simulations emphasize the importance of debriefing in guiding the construction and integration of new knowledge (Dempsey et al., 1996; Garris et al., 2002; Hays, 2006). Historically debriefing has been used to obtain information from a participant (e.g. military debriefing of rescued hostages) and to desensitize a participant (or dehoax in the context of a psychological study involving deception) (Peters & Vissers, 2004). However, debriefing in the context of experience-based learning focuses on participant learning. Because participants in a simulation game may have different experiences and therefore derive different knowledge, debriefing is an important phase of the learning process. Debriefing involves a joint analysis of their experiences. The design of the debriefing should be tailored to the learning objectives and the participants' characteristics (Peters & Vissers, 2004). Debriefing should focus not just on content but on process, especially when the game is played by teams rather than individuals.

Learning Outcomes and Assessment

One common principle of all forms of assessment is that assessment is a process of reasoning from evidence. Assessment is comprised of three connected elements: cognition, observation, and interpretation (Pellegrino, Chudowsky, & Glaser, 2001). Traditional assessment often requires the recall and application of knowledge in contrived contexts that do not address the learner's ability to transfer learning to real-world situations. Such an approach is rarely effective in measuring learning that occurs through complex interactions with others and the environment. Contemporary assessment often includes open-ended, authentic tasks and the iterative creation of artifacts that require the learner to demonstrate not only what he knows but what he can do with that knowledge (Hooper & Reinartz, 2002; Winn, 2002). Pellegrino et al. (2001) prescribe the use of assessments that identify metacognitive strategies used by learners in solving problems and that place those strategies "on a developmental continuum of efficiency and appropriateness for a particular domain of knowledge and skill" (p. 103).

Given the previous description of the purposes and functions of games and simulations in education and the nature of contemporary theories of learning and assessment, it would be reasonable to conclude that the use of games and simulations in education results in improved learning achievement and better assessment of learning outcomes. Unfortunately it is difficult to support this assumption given the available research. Hays (2006) examined over 270 documents from the literature of instructional games and found only 48 that provided empirical evidence of the effectiveness of games. These studies involved different tasks, age groups, and types of games. Furthermore, he concluded that the literature lacks clearly defined terminology and is filled with methodological flaws. In their review of 51 journal articles related to instructional games, Dempsey et al. (1993-1994) found that the majority (n = 28) were descriptive and presented no empirical evidence. In fifteen of the articles, specific learning outcomes were not discussed. The most frequent learning outcome discussed was problem solving (n = 16). In an extension of that review (Dempsey et al., 1996) that included 99 journal articles, many (n=45) did not contain a discussion of learning outcomes. In those articles that did discuss learning outcomes, problem solving was still most prevalent (n=22).

Traditional Assessment Methods

Critical questions in gaming research include what to assess and how to measure it. Bredemeier and Greenblat (1982) identified three major areas that are often said to be addressed by games and simulations: substantive learning (cognitive or affective), motivation, and atmosphere (student-teacher relations). They reported that comparative studies offered conflicting findings regarding the effectiveness for substantive learning and that the majority found no significant differences with traditional instruction. In some cases, games and simulations were more effective “in facilitating positive attitude change toward the subject and its purposes” (p. 324). Most studies reported higher satisfaction and greater interest as a result of a simulation-gaming experience. Furthermore, there were many anecdotal claims about improved learning atmosphere during and after the use of games and simulations, including a more relaxed atmosphere, greater openness and participation, and more incisive inquiry.

Dempsey et al. (1996) reported that the dependent variables found in the review of the literature included “creativity, problem-solving ability, achievement, retention, attitude, self-image, self-efficacy, and continued motivation.” The most frequently mentioned variable was achievement (n=31). Others included problem-solving (n=13), attitude (n=12), continued motivation (n=10), and retention (n=8).

Anderson and Lawton (1992) surveyed 146 college business instructors on their use of eight evaluation methods: team performance versus other teams (92.5%), evaluation of written plan (76.7%), paper analyzing team’s performance (61.6%), oral presentation of team’s performance (52.7%), performance relative to written plan (51.4%), test on rules and procedures (41.1%), ability to predict results of decisions (26.0%), peer evaluation (13.9%), other methods (24.1%). As can be seen from the results, most instructors used more than one method of assessment, with an average of 4.4 methods used. The grade for the simulation was on average 30% of a student’s grade in the course.

Anderson and Lawton (1992) further analyzed which evaluation methods were used with which learning objectives (based on Bloom’s taxonomy of cognitive objectives). A test of rules and procedures was most commonly used to measure basic knowledge (84.8%). The majority tested comprehension using the evaluation of written plan (49.6%). A variety of methods were used to measure application, including team performance (58.7%), ability to predict results (56.8%), and performance relative to plan (55.4%). Similarly, analysis was measured by performance relative to plan (62.2%), a paper analyzing the team’s performance (59.6%), team performance (58.7%), and ability to predict results (56.8%). Synthesis was primarily measured by a paper analyzing the team’s performance (50.6%) and oral presentation (48.6%). Evaluation was measured by oral presentation (50.0%), a paper analyzing the team’s performance (49.4%), and other methods (45.5%). The authors note that peer evaluations were not highly used as a measure of learning, and they speculate that their use was to influence group dynamics.

Gosen and Washbush (2004) analyzed the results of twenty papers on computer-based simulation assessment that used control groups or pre-post designs and yielded quantitative results. Most used course exams to measure effectiveness on learning achievement, although some focused on perceptions of learning, self-efficacy, or attitudes.

Many researchers discuss the difficulties in determining what and how to assess learning outcomes that result from games and simulations. Gosen and Washbush (2004) note that it is difficult and time-consuming to create relevant, purposive learning objectives and a valid instrument with which to determine outcomes. The dependent variable—learning achieved as a result of participating in a simulation—is often difficult to quantify. Many of the potential benefits of games, including “self-management skills, the capacity to collaborate, or the ability to abstract and transfer problem-solving strategies” (Wideman et al, 2007, p. 17) are not easily assessed by traditional means. It might seem reasonable to suggest that the game or simulation is itself the assessment, i.e. one’s score or success is the measure of one’s learning. Washbush and Gosen (2001) conducted eleven experimental studies involving business simulations and found statistically significant evidence of learning based on pre- and posttest comparisons. However, they did not find a correlation between simulation performance and learning gains, leading them to suggest several possible explanations. For example, high performers may have discovered a beneficial strategy early in the simulation, or players may have learned by making poor decisions, or players may have been experimenting with alternative strategies. One of the distinguishing characteristics of modern digital games is that players fail more than they succeed. Therefore, equating a learner’s performance with learning gains seems dubious.

On the other hand, there may be ways around this dilemma. For example, Aldrich (2004) reports struggling to find a way to score performance in a leadership training simulation. He felt that too much emphasis on scores might cause players to take a conservative approach to the simulation instead of experimenting with different strategies. He decided to offer players the option of practicing as much as they desired and seeing their scores. Once they felt they were ready, they could decide to play an assessment round in which their scores were recorded. However, they could only play that round once, and when it was over they had to proceed to the next round.

Contemporary and Emerging Assessment Methods

Contemporary and emerging approaches to assessment emphasize the use of frequent feedback to improve the effectiveness and efficiency of instruction (Pellegrino et al., 2001). Feedback is a critical component of learning as it provides the learner with the opportunity to refine his understanding and performance. Perkins and Unger (1999) note that too often feedback only occurs at the end of instruction and with little indication of what is wrong with the learner’s performance or how it could be improved. Leemkuil et al. (2000) distinguish between outcome measures and process measures. Outcome measures reflect the overall performance but may be calculated using a variety of criteria. Process measures are based on the actions and decisions that led to the outcome.

Technology-based learning environments (including games and simulations) make it possible to assess a learner’s decisions and actions in problem solving and complex reasoning, and to track a learner’s progression over time and provide individualized and immediate feedback (Pellegrino et al. 2001). In discussing computer-based instruction, Young (1995) suggests the use of time-stamped logs to track a learner’s navigation patterns from screen to screen as a way of inferring the learner's goals and intentions. Gibson (2003) notes that network-based assessment systems can offer new methods of assessing learning that are more aligned with the cognitive demands of knowledge work in an information economy. These methods include support and documentation for complex performances, long-term documentation of change, comparison of novice and expert differences, development of metacognitive skills, identification and analysis of problem-solving strategies, customized feedback, and complex statistical analysis and visualization. Loh (2007) proposes the use of information trails in games to track a learner's path and actions at particular nodes in a virtual environment. These nodes must be incorporated into the design of the game so that they present opportunities for desired performance in terms of learning objectives and serve as evidence for assessment purposes.

Barab, Thomas, Dodge, Carteaux, and Tuzun (2005) suggest that portfolio assessment is particularly appropriate when evaluating experiential and inquiry-based learning outcomes. In Quest Atlantis (QA), which is based in a virtual world, learners explore and select quests (academic tasks) that require them to engage in real-world activities and produce artifacts based on their learning. In QA there is "a focus on inquiry-based activities that support the learner in generating information, in evaluating its relevance to real-world problems, in constructing meanings in authentic settings, and in justifying the credibility of assertions" (p. 97).

Wideman et al. (2007) recommend case studies that involve direct or indirect observation in authentic settings (i.e. classrooms rather than research laboratories). The "think-aloud" protocol that is often employed in usability testing may be used in case studies to gain insight into players' understanding and decision making. This approach to assessment can document "knowledge acquisition and application, pattern recognition, strategy deployment, and metacognitive functioning" (p. 18). It can also facilitate the examination of playing styles, interactions (among learners and the teacher), and instructional strategies and methods. However, this kind of research design requires time for both observation and coding and analysis of the data. Furthermore, sample sizes tend to be small and generalizability is limited.

Implications for Further Research

This review has found that while there is growing interest in the use of games and simulations for instruction, much of the past research is anecdotal or descriptive and lacks sound empirical evidence. Games and simulations are congruent with many prevalent theories of learning and employ numerous principles of good instruction. They engage and motivate learners by providing optimal challenges, leading to increased academic learning time. Furthermore, they require learners to think strategically and to recognize and solve complex problems based on acquired information and timely feedback. However, games and simulations need to be designed to facilitate meaningful learning and to provide opportunities for relevant assessment. If it is true that games and simulations can have a positive effect on learning, better research is needed to answer the following questions:

1) Which characteristics of games and simulations facilitate and which detract from learning?
2) What kinds of learning are best suited to games and simulations?
3) What theories, models, and practices can guide instructional designers in creating games that engage and instruct?
4) How can instructors incorporate games into their curricula and support the learning that occurs in games and simulations?
5) How can learning from games and simulations be assessed?

While these questions are essentially cumulative in that the answers to each inform the next, they may be addressed concurrently. In seeking to answer these questions, researchers should employ appropriate research designs such as design-based research (Wang & Hannafin, 2005), developmental research (Richey, Klein, & Nelson, 2004), and formative research (Reigeluth & Frick, 1999). These approaches can enable researchers to ensure the value of their work for real-world practice through the collaborative and iterative implementation and formative evaluation of their designs.

The contemporary and emerging methods of assessment described earlier hold some promise for identifying and measuring the learning outcomes from games and simulations. However, we must continue to develop new methods and tools and adapt relevant approaches from other fields. For example, given the interest in the social nature of learning and the situated, distributed nature of cognition in multiplayer games (Steinkuehler, 2006), current research on social network analysis seems potentially applicable to instructional games and simulations. New tools and techniques are also necessary to analyze and evaluate the patterns of a learner's decisions and actions over time in virtual learning environments. Frick (2006) has developed methods and related software for mapping and analyzing patterns and structures across time that have been used to measure classroom interactions and software usability. These may be highly applicable to the analysis of data generated by learner interactions in games and simulations. But virtual learning environments must be designed with assessment in mind in order to collect and store relevant data regarding their effectiveness for instruction.

References

Aldrich, C. (2004). *Simulations and the future of learning: An innovative (and perhaps revolutionary) approach to e-learning*. San Francisco, CA: John Wiley & Sons.

Aldrich, C. (2005). *Learning by doing: The essential guide to simulations, computer games, and pedagogy in e-learning and other educational experiences*. San Francisco, CA: John Wiley & Sons.

Anderson, P. H., & Lawton, L. (1992). A survey of methods used for evaluating student performance on business simulations. *Simulation & Gaming, 23*(4), 490-498.

Barab, S., Thomas, M., Dodge, T., Carteaux, R., & Tuzun, H. (2005). Making learning fun: Quest Atlantis, a game without guns. *Educational Technology Research and Development, 53*(1), 86-107.

Bragge, J., & Storgards, J. (2007). *Profiling academic research on digital games using text mining tools*. Paper presented at the Situated Play, Digra 2007 Conference.

Bredemeier, M. E., & Greenblat, C. S. (1981). The educational effectiveness of simulation games: A synthesis of findings. *Simulation & Gaming, 12*(3), 307-332.

Cruickshank, D. R. (1988). The uses of simulations in teacher preparation: Past, present, and future. *Simulation & Gaming, 19*(6), 133-156.

de Freitas, S., & Oliver, M. (2006). How can exploratory learning with games and simulations within the curriculum be most effectively evaluated? *Computers & Education, 46*(3), 249-264.

Dempsey, J. V., Lucassen, B., Gilley, W., & Rasmussen, K. (1993-1994). Since Malone's theory of intrinsically motivating instruction: What's the score in the gaming literature? *Journal of Educational Technology Systems, 22*(2), 173-183.

Dempsey, J. V., Rasmussen, K., & Lucassen, B. (1996). *The instructional gaming literature: Implications and 99 sources* (No. 96-1). Mobile, AL: University of South Alabama.

Egenfeldt-Nielsen, S. (2005). Beyond edutainment: Exploring the educational potential of computer games. Unpublished doctoral dissertation. IT University.

Feinstein, A. H., & Cannon, H. M. (2002). Constructs of simulation evaluation. *Simulation & Gaming, 33*(4), 425-440.

Foreman, J. (2004). Game-based learning: How to delight and instruct in the 21st century. *EDUCAUSE Review, 39*(5), 50-66

Frick, T. (2006). *Predicting education system outcomes: A scientific approach*. Paper presented at the Association for Educational Communications and Technology.

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & Gaming, 33*(4), 441-467.

Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. New York: Palgrave Macmillan.

Gee, J. P. (2005). Learning by design: Good video games as learning machines. *E-learning, 2*(1), 5-16.

Gibson, D. (2003). *Elements of network-based assessment.* Paper presented at the Society for Information Technology in Teacher Education conference, Albequerque, NM.

Gosen, J., & Washbush, J. (2004). A review of scholarship on assessing experiential learning effectiveness. *Simulation & Gaming, 35*(2), 270-293.

Gredler, M. E. (2004). Games and simulations and their relationships to learning. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (2nd ed., pp. 571-581). Mahwah, NJ: Lawrence Erlbaum & Associates.

Hays, R. T. (2006). *The science of learning: A systems theory perspective*. Boca Raton, FL: BrownWalker Press.

Hooper, S., & Reinartz, T. J. (2002). Educational multimedia. In R. A. Reiser & J. V. Dempsey (Eds.), *Trends and issues in instructional design and technology* (pp. 308-318). Upper Saddle River, NJ: Pearson Education, Inc.

Kafai, Y. (2001). The Educational Potential of Electronic Games: From Games-To-Teach to Games-To-Learn. Retrieved October 25, 2004 from http://culturalpolicy.uchicago.edu/conf2001/papers/kafai.html.

Kirriemuir, J., & McFarlane, A. (2004). *Literature review in games and learning: A report for NESTA Futurelab*. Retrieved September 1, 2004, from http://www.nestafuturelab.org/research/reviews/08_01.htm.

Leemkuil, H., de Jong, T., & Ootes, S. (2000). *Review of educational use of games and simulations.* Twente: University of Twente. Retrieved December 1, 2007, from http://kits.edte.utwente.nl/documents/D1.pdf

Loh, C. S. (2007). Designing online games assessment as "information trails". In D. Gibson, C. Aldrich & M. Prensky (Eds.), *Games and simulations in online learning: Research and development frameworks* (pp. 323-348). Hershey, PA: Information Science Publishing.

McFarlane, A., Sparrowhawk, A., & Heald, Y. (2002). *Report on the educational use of games*. Cambridge, UK: Teem. Retrieved October 23, 2007, from http://www.teem.org.uk/publications/teem_gamesined_full.pdf

Maier, F. H., & Grossler, A. (2000). What are we talking about--A taxonomy of computer simulations to support learning. *System Dynamics Review*, *6*(2), 135-148.

Mitchell, A., & Savill-Smith, C. (2004). *The use of computer and video games for learning: A review of the literature*. London: Learning Skills and Development Agency, Ultralab. Retrieved February 16, 2005, from http://www.lsda.org.uk/files/PDF/1529.pdf.

Oblinger, D. G. (2006). *Simulations, games, and learning* (No. ELI3004): EDUCAUSE. Retrieved October 30, 2007, from http://connect.educause.edu/library/abstract/SimulationsGamesandL/39338.

Parlett, D. (1999). *The Oxford history of board games*. New York: Oxford University Press.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.

Perkins, D. N., & Unger, C. (1999). Teaching and learning for understanding. In C. M. Reigeluth (Ed.), *Instructional design theories and models: A new paradigm of instructional theory* (Vol. II, pp. 93-114). Hillsdale, NJ: Lawrence Erlbaum Associates.

Peters, V. A. M., & Vissers, G. A. N. (2004). A simple classification model for debriefing simulation games. *Simulation & Gaming, 35, (1),* 70-84.

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

Randel, J. M., Morris, B. A., Wetzel, C. D., & Whitehill, B. V. (1992). The effectiveness of games for educational purposes: A review of recent research. *Simulation & Gaming, 23*(3), 261-276.

Richey, R. C., Klein, J. D., & Nelson, W. A. (2004). Developmental research: Studies in instructional design and development. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (2nd ed., pp. 1099-1130). Mahwah, NJ: Lawrence Erlbaum & Associates.

Reigeluth, C. M., & Frick, T. W. (1999). Formative research: A methodology for creating and improving design theories. In C. M. Reigeluth (Ed.), *Instructional design theories and models: A new paradigm of instructional theory* (Vol. II, pp. 633-651). Hillsdale, NJ: Lawrence Erlbaum Associates.

Reigeluth, C. M., & Schwartz, E. (1989). An instructional theory for the design of computer-based simulations. *Journal of Computer-Based Instruction, 16*(1), 1-10.

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

Romiszowski, A. (1999). The development of physical skills: Instruction in the psychomotor domain. In C. M. Reigeluth (Ed.), *Instructional design theories and models: A new paradigm of instructional theory* (Vol. II, pp. 460-481). Hillsdale, NJ: Lawrence Erlbaum Associates.

Ruben, B. D. (1999). Simulations, games, and experience-based learning: The quest for a new paradigm for teaching and learning. *Simulation & Gaming, 30*(4), 498-505.

Rutter, J., & Bryce, J. (2006). An introduction to understanding digital games. In J. Rutter & J. Bryce (Eds.), *Understanding digital games* (pp. 1-17). London: Sage Publications Ltd.

Salen, K., & Zimmerman, E. (2004). *Rules of play: game design fundamentals*. Cambridge, Massachusetts: The MIT Press.

Shaffer, D. W., Squire, K. R., Halverson, R., & Gee, J. P. (2005). Video games and the future of learning. *Phi Delta Kappan, 87*(2), 104-111.

Steinkuehler, C. (2006). Games as a highly visible medium for the study of distributed, situated cognition. In S. A. Barab, K. E. Hay, N. B. Songer, & D. T. Hickey (Eds.), *Proceedings* *of the International Conference of the Learning Sciences* (pp. 1048-1049). Mahwah NJ: Erlbaum.

Strommen, E., & Lincoln, B. (1992). Constructivism, technology, and the future of classroom learning. *Education and Urban Society, 24 (4)*, 468-469.

Van Sickle, R. (1986). A quantitative review of research on instructional simulation gaming: A twenty-year perspective. *Theory and Research in Social Education, 14*(3), 245-264.

Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development*, *53*(4), 5-23.

Washbush, J., & Gosen, J. (2001). An exploration of game-derived learning in total enterprise simulations. *Simulation & Gaming, 32*(3), 281-296.

Wideman, H. H., Owston, R. D., Brown, C., Kushniruk, A., Ho, F., & Pitts, K. C. (2007). Unpacking the potential of educational gaming: A new tool for gaming research. *Simulation & Gaming, 38*(3), 10-30.

Winn, W. (2002). Current trends in educational technology research: The study of learning environments. *Educational Psychology Review, 14*(3), 331-352.

Wolfe, J., & Crookall, D. (1998). Developing a scientific knowledge of simulation/gaming. *Simulation & Gaming, 29*(1), 7-19.

Yazzie-Mintz, E. (2007). *Voices of students on engagement: A report on the 2006 high school survey of student engagement*. Retrieved December 8, 2007, from http://ceep.indiana.edu/hssse/pdf/HSSSE_2006_Report.pdf

Young, M. (1995). Assessment of situated learning using computer environments. *Journal of Science Education and Technology, 4*(3), 89-96.