Games … and … Learning

Valerie J. Shute, Lloyd Rieber, & Richard Van Eck

Introduction

Anyone who makes a distinction between games and learning doesn't know the first thing about either.
~Marshall McLuhan

In this chapter, we aim to connect the dots between games and learning. Gee (2008) has suggested, and we agree, that game design has a lot to teach us about learning, and contemporary learning theory has something to teach us about designing better games. One link already in place between these two realms is formative feedback—a critical part of any learning effort (e.g., Shute, 2008), and also a key component in good game design which adjusts challenges and gives feedback so that different players feel the game is challenging and their effort is paying off.

Our thesis in this chapter is that (a) learning is at its best when it is active, goal-oriented, contextualized, and interesting (e.g., Bransford, Brown, & Cocking, 2000; Bruner, 1961; Quinn, 2005; Vygotsky, 1978); and (b) instructional environments should thus be interactive, provide ongoing feedback, grab and sustain attention, and have appropriate and adaptive levels of challenge—i.e., the features of good games. Along the same line, Gee (2003) has argued that the secret of a good game is not its 3D graphics and other bells and whistles, but its underlying *architecture* where each level dances around the outer limits of the player's abilities, seeking at every point to be hard enough to be just doable. Similarly, psychologists (e.g., Falmagne, Cosyn, Doignon, & Thiery, 2003; Vygotsky, 1987) have long argued that the best instruction hovers at the boundary of a student's competence. More recent reports (e.g., Thai, Lowenstein, Ching, & Rejeski, 2009) contend that well-designed games can act as *transformative digital learning tools* to support the development of skills across a range of critical educational areas. In short—well designed games have the potential to support meaningful learning across a variety of content areas and domains.

Why aren't games used more widely in classrooms? While time constraints, cost of games, and a lack of prescriptive guidance are all possible reasons for this (e.g., Charsky, in press; Van Eck, 2007), one *major* hurdle is the lack of good research on games and learning (Van Eck, 2008). Compared to other types of instructional systems, there are currently too few experimental studies examining the range of effects of gaming environments on learning, and a corresponding lack of theory and practice for their design and implementation.

In their seminal book, *Rules of Play*, Salen and Zimmerman (2004) define a game as "a system in which players engage in an artificial conflict, defined by rules, that results in a quantifiable outcome" (p. 80). In addition to conflict, rules, and outcomes, Prensky (2001) includes goals, feedback, interaction, and representation (or story) into the mix of essential game elements. Pulling from each, our list of educational-game "must haves" includes: (a) conflict or challenge (i.e., a problem to be solved), (b) rules of engagement, (c) particular goals or outcomes to achieve (which often includes many sub-goals), (d) continuous feedback (mostly implicit, but may be explicitly cognitive and/or affective), (e) interaction within the environment, and (f) compelling storyline. These game elements are actually quite similar to those underlying good instructional design, which we hope will be clear by the end of the chapter.

Because this chapter is housed within a book on instructional system design (ISD), our focus is on games designed for educational purposes. Specifically, we limit the scope of our discussion to *interactive, digital games that are intended to support learning and/or skill acquisition*. We'll refer to these systems as educational games, and those who play them as students (to emphasize our educational focus).

This chapter consists of three main parts. First, we describe theories about and features of games that make them so engaging and thus suitable as instructional vehicles. Second, we examine the architectural elements and functionality needed in games to enable them to support learning, using a case study to illustrate our point. Third, we describe stealth assessment--an evidence-based method and set of tools that enable us to assess learning during game play while not disrupting the fun or state of flow. We conclude with a view toward the future of games-and-learning research.

Games and Play Theory

The question of what makes a game fun to play is very similar to the question of what makes a joke funny. On one level, we all know the answer, but articulating it well, if at all, is surprisingly difficult. In one episode of the television show, *Star Trek: The Next Generation*, Commander Data decided to confront the question "What makes something funny?" As an android who aspired to become human, this question was very perplexing to him and he set about answering it as a programmed machine, or a very analytical engineer might, by breaking the construct "funny" into all of the conceivable rules. Mr. Data erroneously tried to come up with a grand "if/then" tree for "funny" (i.e. if I say this, then say that, in this way, etc., *then* it is funny). In contrast, the best answer, to paraphrase Garrison Keillor, is that something is funny simply because people laugh. We can chuckle at Mr. Data's misguided attempt, but many people in the ISD field seem to be following a similar rule-based "engineering" path to try to understand how to design a game that is fun and also leads to learning. Similar to Keillor's definition of funny, a game is fun to play if people enjoy playing it. More specifically, we argue that a game is engaging, or fun to play if it triggers the play phenomenon in the player. So, we must take some time to understand the play phenomenon. Fortunately, much research has been done on play from a multitude of disciplines, such as education, psychology, sociology, and anthropology.

Making play an objective of an educational game requires a paradigm shift for most designers – one that is very learner-centered and constructivist in nature. To understand this paradigm, you need to understand the difference between merely playing a game and being "at play." The former can be mandated by a teacher to students, or a trainer to a group of employees, and these participants can dutifully "play the game." That is, one can watch and track their behavior or performance from beginning to end and even declare that one or more has won the game. However, these individuals may never have been "at play," meaning that they never entered the conceptual cognitive or cultural space in which play occurs (Huizinga, 1950).

So, what is play? Everyone reading this chapter already knows what play is and you yourself have probably experienced it within the last 24 hours, even though you may resist, as many adults do, in using the word "play" to describe it. It probably happened during your leisure time, although if you are fortunate to love your job, it may have happened at work. It was definitely something you wanted to do and you would say that you did it voluntarily. You found the activity intrinsically motivating and so you were not concerned about "getting something" out of it. You were also doing something actively and probably physically. Finally, you were likely in a state where you were not conscious of yourself or of your place in the world, but rather felt wholly absorbed in the activity. This state also carried a feeling of being very free from risks. You felt free to try new things or to experiment with different ways of doing or thinking – after all, it was only play. Your awareness of time likely disappeared and you were probably surprised by how much time had passed when the activity had ended (see Pellegrini, 1995; Rieber, 1996; and Sutton-Smith, 1997 for formal definitions and attributes of play). Some of you may have experienced play while engaged in a hobby, such as gardening, woodworking, photography, painting, or some craft. Others may have experienced it while caring for a son or daughter and enjoying each other's company. Yet others experienced it while reading a book, playing a musical instrument, or playing a video game. A lucky few have experienced it while writing a chapter in a book.

Educators and other educational stakeholders (e.g., parents, state legislators) are quick to ask--What good is play? Does it lead to some productive outcome or result? The seminal work of Jean Piaget remains an important starting point for such questions (Phillips, 1981; Piaget, 1951). Piaget felt that play and imitation were core and innate human strategies for cognitive development. With play, a child could rehearse a newly formed concept to make it fit within what they already knew and understood (assimilation). As a child experiences or encounters new events, activities, ideas, or rituals, imitation is used to build entirely new mental models (accommodation). The child continues in this way to achieve an orderly balanced world while constantly confronting a changing, shifting environment. Just as the mental processes of assimilation and accommodation continue throughout life, so too do play and imitation remain important cognitive tools for people from childhood through adulthood.

There are other examples of research literature, while not overtly aligning with play, that are clearly in the same camp. The research on self-regulated learning (Zimmerman, 1990, 2008) is one example, especially with its emphasis on an individual actively working towards goals within intrinsic motivating activities. However, the attributes of flow theory proposed by Csikszentmihalyi (1990) are the most similar to that of play, especially in the context of game design. For example, flow theory specifically addresses the need to optimize challenge, so as to continually avoid anxiety and boredom. Activities that induce flow have clear goals, coupled with clear and consistent feedback about whether a person is reaching these goals. Another important attribute of flow is that it takes effort to attain a state of flow, requiring a clear and deliberate investment of sustained attention.

The psychologist Brian Sutton-Smith (Sutton-Smith, 1997) has proposed many ways to think about play--what he calls the rhetorics of play. Among the most alluring of these rhetorics for educators is the idea that play leads to something productive (i.e., play as progress). However, Sutton-Smith refers to the ambiguity of play in being able to "deliver the goods." Although there are tantalizing reasons for believing that play is by and large a good thing, one should be very careful in attributing positive results directly to it. There is evidence that positive outcomes and play go together (i.e., correlational effects), but one cannot say that play *caused* these outcomes. Another ambiguity of play is that experiencing it itself may be its own reward and that the goal of getting something out of play is misguided. But, the presence of the play state may at least be evidence that the person is in a good state for subsequently experiencing cognitive and social growth and this alone may be good enough reason to make play a goal for any learning environment.

A developmental theory of play and its role in the human life cycle comes from the renowned child psychologist David Elkind (2007). His theory is quite different from that of Brian Sutton-Smith. Elkind's theory posits three instinctual drives that are the root of all human cognition and behavior throughout a person's lifetime: Love, work, and play. Love refers to a person's disposition to express one's desires, feelings, and emotions. Work refers to a person's disposition to adapt to the demands of his or her physical and social worlds. Play, modeled after that of Piaget, is a person's need to adapt the world to one's self and create novel learning experiences.

To become a well-adjusted person living compatibly within a complex social system, one must balance the demands and goals of each of these three elements in ways that change throughout life. For example, we begin life with play dominating our experiences, but then work dominates us as we proceed through the early elementary school years. Love and play, though diminished, do not disappear, but take on a supporting role to work. In adolescence, love dominates, but again with the other two supporting it. Love, work, and play become fully separate in adulthood, but each can be manifested in combination with the others. Bringing love, work, and play into harmony with each other at points throughout one's life is an important goal and one that parents and teachers (and instructional designers) should work to facilitate. Achieving a balance between love, work, and play is similar to Csikszentmihalyi's concept of psychological growth during flow (Csikszentmihalyi, 1990), where an individual becomes more complex or advanced by balancing the need to be a unique individual with a unique identity (differentiation) while at the same time feeling connected to other people and social groups (integration). Furthermore, Elkind's theory of the relationship between love, work, and play can easily be mapped onto the current interest in "21st century" skills in which the ability to work creatively and effectively with others sharing a common goal or purpose is much valued.

This chapter is about digital, or computer, games for learning. Among those conducting research in this area, the prevailing interest tends to be focused on the immersive games, such as massively multi-player online role-playing games, or MMORPGs. The technology underlying these highly visual, persistent virtual worlds is very impressive and the technical sophistication of these "high-tech" gaming environments can only increase, leading to new game genres and models of interaction which we cannot even imagine now. Would-be educational game designers, however, would do well to consider low (i.e., non-digital) and middle-tech (i.e., the span of digital games up to high-tech examples) approaches to gaming in addition to the high-tech games, if only to understand that the fundamentals of a game extend outside of the specific technology of any single game. Regardless of the degree of technology infusion in the game, we believe that the play phenomenon is always eager to emerge.

Having laid a theoretical foundation for learning and play in relation to educational games, we next examine architectural issues, viewed through the lens of instructional system design.

Game Architecture—A Case Study

To illustrate the learning principles and architectural elements we have described, we present a case study of the design of an educational game to teach scientific problem-solving skills in the context of ecological problems facing a mythical city. A more complete description can be found elsewhere (see Van Eck, Hung, Bowman, & Love, 2009) as space does not permit a full discussion here. Instead we focus on some of the most significant elements of the design process as it relates to our central thesis of good games and learning. We begin by addressing the theory and models that drove the design of the game, then describe how those theories and models led to specific game architectural elements. Finally, we discuss how the learning principles outlined in the first part of this chapter are addressed generally throughout the game.

In the game we describe below, students take on the role of an apprentice environmental scientist trying to solve a series of nine environmental “mysteries” within a mythical city. In each scenario, the player interacts with individual community members, an in-game mentor, a committee of environmental experts, and a community council.

Theories and Models

Problem solving is largely context- and domain-specific (e.g., Bransford, Franks, Vye, & Sherwood, 1989; Brown, Collins, & Duguid, 1989), so problem solving in one domain often fails to transfer to others. However, repeated practice across multiple problems can improve problem solving within a given domain (Gagné, Wager, Golas, & Keller, 2005). Because good games are engaging and interactive, they can potentially provide repeated problem solving practice. In addition, because games are frequently cited as examples of situated problem-solving in general (Gee, 2007, Van Eck 2006), and in domains like science specifically (Gaydos & Squire, in press), it should be possible to design a game to promote scientific problem solving skills. Van Eck and colleagues set out to design such a game (Van Eck et al., 2009).

Any such endeavor must be driven by theory and instructional design principles, and do so in a way that is sensitive to the ludic (play) nature of games. Researchers (e.g., Hung, 2006; Jonassen, 1997, 2000, & 2002; Jonassen & Hung, 2008) have made advances in both the delineation and definition of problem types and models for designing effective problems. The first challenge the designers faced for this game, then, was to identify what was meant by scientific problem solving, and how to design effective problems for each game scenario.

Scientific Problem Solving. Scientific inquiry comprises two types of problem solving processes or methods: “scientific” and “engineering” (see: Society for Science and the Public, 2008). The scientific method is used to gather information and answer questions about the problem itself, and the engineering method is used to design, implement, and evaluate solutions to the problem. The designers combined these two approaches along with the National Science Education Standards (1996) into a model for scientific problem solving which will be described in more detail later (see Figure 1).

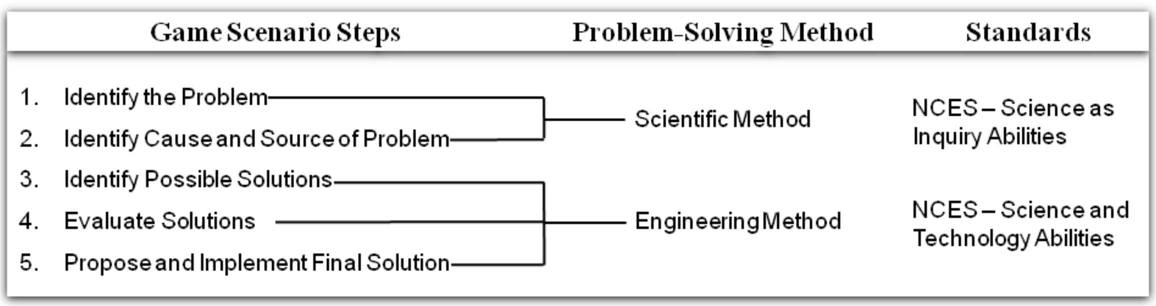

Figure 1. Alignment of Game Steps with Scientific Problem Solving

Problem Design. One of the key challenges for problem-based learning is the design of good problems (e.g., Lee, 1999). This, in turn, requires an understanding of different problem typologies, their attendant cognitive processing requirements, and the way each problem type may be supported best by different types of gameplay (see Hung & Van Eck, in press; and Jonassen, 2000 for more on this topic).

The designers adopted the 3C3R model (Hung, 2006) and the nine-step process for its application (Hung, 2009) to guide the systematic design of effective problems. The

3C3R model consists of content, context, and connection (the core components) which are primarily concerned with the issues of appropriateness and sufficiency of content knowledge, knowledge contextualization, and knowledge integration. The 3C3R model also comprises what are called processing components, which include researching, reasoning, and reflecting, to facilitate mindful and meaningful engagement. The 3C3R model and the 9-step design process have been validated (Goodnough & Hung, 2008) thus ensuring that the quality of the problems designed for the game would be high.

Theory into Practice

We have articulated the theories and models the game designers used to structure scientific problem solving and to design problems within the game. We turn now to a description of how the learning principles outlined in the beginning of this chapter and the steps specified by the theory and model for scientific problem solving led to specific architectural elements within the game. We remind readers that this list is not exhaustive and that a large body of literature on games and learning exist and should be consulted. Our purpose here is to illustrate some of the principles as they can be instantiated within a gaming environment designed to promote learning.

Learning Should be Goal Oriented. According to Jonassen (2002), all good problems share two characteristics. First, they have some kind of goal, or unknown. The goal/unknown requires the generation of new knowledge. Second, all problems should have some value to the learner in solving them. Like problems, games have a goal/unknown which requires the learner to generate new knowledge. Games, (at least, good ones) also have a value to the learner in achieving the goal. So a game that focuses on problem solving will, by definition, be goal oriented.

The designers of the game in our case study specified a series of nine problems to solve, each of which has its own goal and solution strategies, so each problem is goal oriented. These problems/scenarios are the means by which the player achieves the overall goal of the game, which is to become an initiate of the inner circle of scientists, so the game itself is goal oriented. Finally, because different players have different goal orientations (e.g., performance vs. mastery), several different strategies are possible in solving the problems. For example, players can pursue side quests such as solving a PH problem for a community pond. Depending on the number and type of successful solutions to these side quests, the player can earn different awards such as Water Wizard or Earth Engineer. Thus, mastery goal-oriented students can pursue multiple quests, while performance goal-oriented students can pursue optimal paths in the fastest time possible for social rewards. While not central to the problems themselves, such features allow for multiple goal orientations among players.

Learning Should be Contextualized. Everything a player is asked to learn in an educational game should be relevant and contextualized such that players should not have to learn something that is not used, nor use something they cannot learn. Thus educational games employ contextualized learning. In our case-study game, scientific problem solving is learned in the context of solving actual problems that face citizens in a city rather than through studying a set of rules, propositions, or heuristics, and players/learners are only asked to do so when it is relevant to the problem scenario they currently face. For example, players do not need to know anything about neurological effects of lead and other toxic elements until they are trying to figure out why there has been an increase in learning disabilities at a local school.

Learning Should be Active and Interactive. Problems in the 21st century, like the challenges in games, are solved in a distributed, iterative fashion. Such problems are often ill-structured, non-linear, and require data gathering from a variety of sources. Problems themselves are also rarely presented in a complete fashion, but instead, often have several elements missing. The problems for the game in this case study were designed as a complete case first, then had key elements removed. Those missing pieces were distributed across multiple resources, including the mayor, reporter, neurologist, school superintendent, mentor, committee of scientists, PDA, and technical manuals and references. Solving the problems requires that the player seek out different resources throughout the game. Resources provide different types of information depending on where the player has been and what information they currently possess. Thus, the game provides multiple opportunities for interaction and requires active participation for the player to solve the problem. This is a common feature of commercial games and ensures that the learner is an active participant.

Learning Should Provide Adaptive Challenge and Support. In order to provide varying levels of challenge and support according to different levels of expertise, the designers organized the nine problems in the game into three levels of increasing complexity and decreasing support. The first three game scenarios (Level 1) are designed to be the easiest to complete and provide the most support. The problems for levels two and three (each level containing three scenarios) are gradually more complex, require more contributions from the learner, and are accompanied by less support.

Like the problems themselves, support is also distributed and contextualized. The designers adopted Vygotsky's (1978) concept of scaffolding to provide the minimal support needed for learners to achieve at a level beyond their independent ability. Scaffolding takes many forms within the game, including dialog with the committee of eco-scientists, interactions with the mentor/advisor and other characters, the PDA, and various texts and references. Challenge and support are also adaptive in the sense that behavior patterns and actions (like the Evidence Model that we refer to in the final section of this chapter does) within the game may trigger support or challenge options. For example, too much elapsed time since the last action and repeatedly exploring dead-end branches of the game may trigger a communication from the mentor or an early summons by the eco-scientist committee. Likewise, completion of optimal paths in short periods of time may result in fewer offers of support from the mentor and/or a decrease in time allowed for each scenario.

Learning Should Incorporate Feedback. Every action in the game results in some form of feedback, but the nature of that feedback is again contextualized. Taking soil samples at the schoolyard results in flag markers to indicate the location and identifying number, which are later used to identify the sample results returned from the testing laboratory. Speaking to characters always results in some form of response, either as additional information (in which case the player knows they are on the right track) or a canned response like "I don't have anything to add to what I've already said" (in which case the learner knows that it's time to move on). What is key is that the feedback provides contextualized hints and prompts rather than direct answers or instruction, and it is contextually sensitive to the game narrative, problem, and environment.

Putting it all Together

We have described how the designers used theories and models to articulate a game architecture, and discussed how the learning principles we outlined in the beginning of the chapter are contextually embedded in the game. We now describe an initial step in the first scenario (about lead pollution) to further illustrate the problem-solving process and learning principles already described. This step is the first of the five listed in the scientific problem-solving model, shown earlier in Figure 1.

Step 1: Identify the Problem. The scenario opens with a TV news story called "Learning Disabilities On the Rise?" The news story talks about increased learning disabilities in the schools. "Some people say learning disabilities are on the rise" [quote from a prominent local educator], "but others say we are simply better at detecting what has always been there" [quote from the Mayor]. "If disabilities are on the rise, what's causing the increase?" [quote from a neurologist]. "Some say that the issue is quite complicated, and what we are seeing is really an increase in ADHD" [quote from the mayor].

The game mentor contacts the player and provides basic hints about possible next steps. The player gets names of sources and additional background information from the reporter. From this, the player devises a plan for gathering more data by talking to different people. In tracking down an educator at one school, the player learns there have been increased numbers of learning disabilities and that enrollment in special needs curricula seems to be higher. After following up with another educator and the school superintendent, the player is told there has *not* been an increase. The mayor believes it is just the result of better testing, and that the school is reporting higher learning disability cases as an excuse for failing. A neurologist quoted in the news story says there is no data or information on causes, and that she doesn't even know if there IS a problem because she hasn't seen the research. Each conversation results in notes and facts in the player's in-game PDA for later reference. Once the player has spoken with all the relevant people, the mentor contacts the player to ask for a status report. The player delivers this by selecting data and facts in the PDA notes (from prior conversations) and placing the list of information in order of relevance. In response, the mentor says it sounds like it might be enough to warrant a trip to the committee of eco-scientists.

The player and mentor then go to visit the eco-scientists. During this meeting, the committee asks a series of questions designed to get the player to think critically about the key aspects of the issue (as determined by the 3C3R model in generating the problem for the game). The player must again be selective about the data she's gathered by checking off the key facts and assigning relative weights to them. The scientists ask questions about key factors not identified by the player, and about factors that are identified but are irrelevant to the problem. The player must choose from a list of possible dialog responses (a common feature of commercial games that has also been shown to be as effective as the learner generating their own questions) designed to force her to think about the information critically. If the player successfully navigates this process, the committee members begin to confer, but their conversation can be overheard.

As the committee speaks, key facts show up in the player's PDA. The scenario proceeds, with the player having to identify her approach, who she plans to talk to, and how she'll know if she's on the right track. The player then leaves the committee headquarters to collect additional data; one educator shows increased enrollments in special needs classes, while another does not. The superintendent says there is an overall increase, but that it mirrors the general population. If the player has already spoken to both educators, an additional line of questions opens up with the superintendent regarding changes at specific schools. Following that line of questions prompts the superintendent to provide a report of incidences by school, which the player can use to identify one particular school with a "cluster" of cases. Following up with the neurologist, the player learns the timeframe from contamination to neurological effects, and *if* the player has spoken to the educators and superintendent, an additional line of questions about specific schools opens up. The player then returns to the committee and the process continues.

This is approximately half of the first step, and each successive step in the five-step model proceeds in a similar fashion. It should be noted that what is described here is only the optimal path and does not include all the possible elements of the game play experience.

This case study shows how theory and practice from ISD are important to the design of good educational games. One final aspect of games-and-learning research that we want to address is the issue of assessment. That is, given the design and development of an educational game that is based on a theoretical foundation, how can we determine if, in fact, the game is succeeding in its goal of engendering deep, meaningful learning relative to important competencies or content? We now turn our attention to the important issue of assessment within educational games, specifically “stealth assessment” intended to accurately and dynamically measure how students are progressing, while not interrupting the fun of the game.

Assessment in Games

In games, as players interact with the environment, the values of different game-specific variables change. For instance, getting injured in a battle reduces health and finding a treasure or other object increases your inventory of goods. In addition, solving major problems in games permits players to gain rank or "level up," such as getting the thumbs-up from the committee of eco-scientists described in the aforementioned case study. One could argue that these are all "assessments" in games—of health, personal goods, and rank. But now consider monitoring educationally-relevant variables at different levels of granularity in games. In addition to checking health status, players could check their systems-thinking, creativity, and teamwork skills, where each of these competencies is further broken down into constituent knowledge and skill elements. If the values of those skills got too low, the player would likely feel compelled to take action to boost them.

One main challenge for educators who want to employ or design games to support learning is making *valid inferences* – about what the student knows, believes, and can do – at any point in time, at various levels, and without disrupting the flow of the game (and hence engagement and learning). One way to increase the quality and utility of an assessment is to use evidence-centered design (ECD) which informs the design of valid assessments and yields real-time estimates of students' competency levels across a range of knowledge and skills (Mislevy, Steinberg, & Almond, 2003). Accurate information about the student can be used as the basis for (a) delivering timely and targeted feedback, as well as (b) presenting a new task or quest that is right at the cusp of the student's skill level, in line with flow theory and Vygotsky's zone of proximal development. ECD will be described in more detail, shortly.

Given the goal of using educational games to support learning in school settings (and elsewhere) we need to ensure that the assessments are valid, reliable, and also pretty much invisible (to keep engagement intact). That's where "stealth assessment" comes in (see Shute, in press; Shute, Ventura, Bauer, & Zapata-Rivera, 2009). During gameplay, students naturally produce rich sequences of actions while performing complex tasks, drawing on the very skills or competencies that we want to assess (e.g., specific scientific inquiry skills). Evidence needed to assess the skills is thus provided by the players' interactions with the game itself (i.e., the processes of play), which may be contrasted with the product(s) of an activity, comprising the norm within educational and training environments.

Making use of this stream of evidence to assess students' knowledge, skills, and understanding (as well as beliefs, feelings, and other learner states and traits) presents problems for traditional measurement models used in assessment. First, in traditional tests the answer to each question is seen as an independent data point. In contrast, the individual actions within a sequence of interactions in a game are often highly dependent on one another. For example, what one does in a particular game at one point in time affects subsequent actions later on. Second, in traditional tests, questions are often designed to get at one particular piece of knowledge or skill. Answering the question correctly is evidence that one may know a certain fact: one question – one fact. But by analyzing a *sequence* of actions within a quest (where each response or action provides incremental evidence about the current mastery of a specific fact, concept, or skill), instructional environments are able to infer what learners know and do not know overall. Now, because we typically want to assess a whole cluster of skills and abilities from evidence coming from learners' interactions within a game, methods for analyzing the sequence of behaviors to infer these abilities are not as obvious. As suggested earlier, ECD can address these problems.

Evidence-centered design

The fundamental ideas underlying ECD came from Messick (1994) and were then formalized by Mislevy and colleagues (e.g., Mislevy & Haertel, 2006; Mislevy, Almond, & Lukas, 2004; Mislevy, et al., 2003). A game that includes evidence-based assessment must be able to elicit behavior from the students that bears evidence about the targeted knowledge and skills (i.e., the competencies), and it must additionally provide principled interpretations of that evidence in relation to the purpose of the assessment. Figuring out these variables and their interrelationships is a way to answer a series of questions posed by Messick (1994) that get at the very heart of assessment design generally, and ECD specifically.

Competency Model: What collection of knowledge and skills should be assessed? A given assessment is meant to support inferences for some purpose (e.g., grading, diagnosis, guidance for further instruction). Variables in the competency model (CM) are usually called 'nodes' and describe the set of knowledge and skills on which inferences are to be based. The term 'student model' is used to denote a student-instantiated version of the CM—like a profile or report card, only at a more refined grain size. Values in the student model express the current belief about a learner's level on each variable in the CM. For instance, in the case study described earlier, suppose that a student was seriously struggling with stating a hypothesis about an environmental problem. The relevant node in the "scientific inquiry" competency model would be "hypothesis generation" and may be estimated as *p*(HypothGen is `LOW` given current set of evidence) = .85 (presuming a set of three discrete levels: low, medium, and high). This level of skill is about the right grain size for diagnosis and instructional support.

Evidence Model: What behaviors or performances should reveal those competencies? An evidence model expresses how the learner's interactions with, and responses to a given problem constitute evidence about CM variables. The evidence model (EM) attempts to answer two questions: (a) What behaviors or performances reveal targeted competencies? and (b) What is the functional (or statistical) connection between those behaviors and the CM variable(s)? Basically, an evidence model lays out the argument about why and how the observations in a given task situation (i.e., learner performance data) constitute evidence about CM variables. In the hypothesis-generation estimate above, the evidence model calculates the probabilities, per competency, given a learner response or sequence of actions as input.

Task Model: What tasks should elicit those behaviors that comprise the evidence? Task-model variables, used in typical assessment design, describe features of situations that will be used to elicit performance. A task model (TM) provides a framework for characterizing and constructing situations with which a student will interact to provide evidence about targeted aspects of knowledge related to competencies. Task specifications establish what the learner will be asked to do, what kinds of responses are permitted, what types of formats are available, and so on. The main purpose of tasks or quests is to elicit evidence (which is observable) about competencies (which are unobservable). For stealth assessment in games, we use the term "action model" instead of task model. This reflects the fact that we are dynamically modeling students' action sequences. These action sequences form the basis for drawing evidence and inferences and may be compared to simpler task responses as with typical assessments. The action model in a gaming situation defines the sequence of actions, and each action's indicators of success. Actions represent the things that students do to complete the quest or solve a problem. To continue with the hypothesis-generation example, the game would need to have the means for a learner to input their hypothesis, as well as rubrics for scoring them.

In games with stealth assessment, the competency model for a given student dynamically accumulates and represents belief about the targeted aspects of skill, expressed as probability distributions for competency-model variables (Almond & Mislevy, 1999; Shute et al., 2009). Evidence models identify what the student says or does that can provide evidence about those skills (Steinberg & Gitomer, 1996) and express in a psychometric model how the evidence depends on the competency-model variables (Mislevy, 1994). Task or action models express situations that can evoke required evidence. One effective tool that has been used in such competency and evidence modeling efforts is Bayesian networks (e.g., Pearl, 1988). That is, Bayes nets may be used within student models to handle uncertainty by using probabilistic inference to update and improve belief values (e.g., regarding learner competencies). Examples of Bayes net implementations for student models may be seen in: Conati, Gertner, and VanLehn (2002); Shute, Graf, and Hansen (2005); and VanLehn et al. (2005).

In short, using ECD and Bayes nets to craft stealth assessments embedded directly in the game along with automated data collection and analysis tools, can not only collect valid evidence of students' competency states, but also reduce teachers' workload in relation to managing the students' work (or "play") products. If a particular game was easy to employ and provided integrated and automated assessment tools, then teachers would more likely want to utilize the game to support student learning across a range of educationally valuable skills. Stealth assessment is intended to help teachers facilitate learning, in a fun and engaging manner, of educationally valuable skills not currently supported in school. It also, of course, is intended to facilitate the flow state for students engaged in gameplay.

Summary and Conclusion

Our goal for this chapter was to begin to connect the dots between games and learning. Toward that end, we described how well-designed games provide an environment in which people are more receptive to learning, especially as compared to traditional environments like the classroom (see earlier discussion on Piaget's and Sutton-Smith's theories). We compiled a list of elements comprising a "well-designed" game, including: problems to solve, rules of engagement, objectives or outcomes, feedback, interaction, and storyline, which we linked to good instructional design (e.g., providing feedback and opportunities for interaction). Next, we presented a case study illustrating many of those elements within an educational game. Finally, we noted the importance of explicitly connecting games' processes and outcomes to educationally-valuable competencies to maximize the odds of games being used in educational contexts. Our final section on stealth assessment was intended to highlight the need for accurate, dynamic assessment and diagnosis of educationally-valuable skills during gameplay. We discussed how stealth assessment can support instructional decisions while operating beneath the radar in terms of monitoring and measuring these competencies. See Figure 2 for an illustration of these relationships.

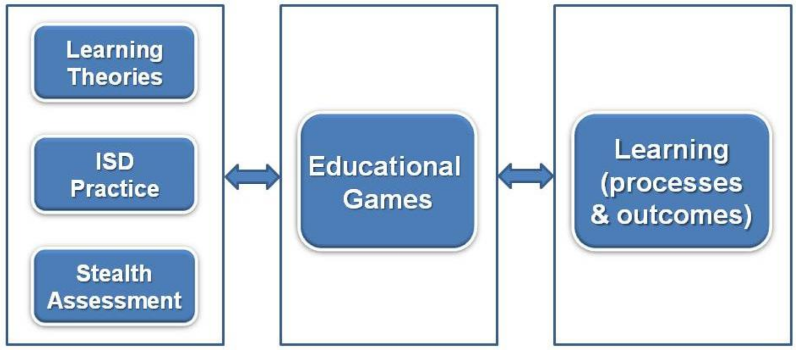

Figure 2. Relationships among factors relating to games and learning

This chapter can be viewed as the beginning of a gaming-for-learning model that supports new competencies, and incorporates socio-constructivist learning theory, flow theory, play theory, principles of instructional design, and stealth assessment. Fleshing out such a model (comprising future research) would help us to identify specific game-design elements, their interactions with one another, with the learner(s), and with the content and competencies being supported. For instance, does the type and timing of feedback differentially affect learners or types of learners? Is it better to explicate goals or design them to be emergent or induced? Does that decision about goals depend on the learner and/or the content being instructed? What is the best grain size of competencies to monitor in a game to maximize learning? And similarly, how can we optimally match the level of a challenge to a learner's competency level (i.e., at the cusp of do-ability)? For many of these research questions, learning theory may inform design elements that facilitate the students' in-game experiences and enhance learning within and from educational games. The gaming-for-learning model itself may be used to design and analyze a variety of educational games answering general questions such as: what works, for whom, to what degree, under which conditions, and for what competencies or domains?

In conclusion, well-designed games are a potentially powerful vehicle to support learning – particularly in relation to new competencies not currently embraced by our educational system but needed to succeed in the 21st century (e.g., work productively within diverse teams, identify and solve complex problems with innovative solutions, communicate effectively, think critically, use technology efficiently, understand system dynamics, and engage in evidence-based reasoning). There are simply too few experimental studies examining the range of effects of gaming environments on learning (e.g., Van Eck, 2007). We believe that the new games-and-learning research stream is highly relevant and important to the field of ISD, which can both inform and be informed by the research.

We close as we began with a relevant quote.

> "*Games are thus the most ancient and time-honored vehicle for education… We don't see mother lions lecturing cubs at the chalkboard; we don't see senior lions writing their memoirs for posterity. In light of this, the question, 'Can games have educational value?' becomes absurd. It is not games but schools that are the newfangled notion, the untested fad, the violator of tradition.*"
> - Chris Crawford

Gaming and Learning: Summary of Key Principles and Practices

1. Good games trigger the play phenomenon in the players.

2. Good games for learning, like all good learning activities, should be active, goal-oriented (with goals valued by the players), contextualized, and designed with adaptive challenge and support.

3. The fundamentals of designing a good game for learning extend beyond any specific technology for a single game.

4. Principles of instructional design and problem-based learning can support and inform the design of good games for learning.

5. Good games for learning provide opportunities for real-time, unobtrusive (stealth) assessment leading to evidence-centered design making it possible to deliver timely and targeted feedback to players and present new game tasks that are at the upper boundary of the student's skill level.

6. The ability to work creatively and effectively with others toward a common goal is an important 21st century skill that is emphasized in good games.

Gaming and Learning: Application Questions

1. Try to design a non-digital game with everyday objects found in your home or classroom (e.g. paper cups, paper clips, ping pong balls, etc.). Ask a friend or two to play it, then ask them if they think the game is any fun. Ask them for ideas to improve the game. Using any of their ideas, and those you thought of, redesign the game and ask another group of friends to play this new version. Is the game more fun? Try to list or chart out the design process you experienced. Does the game have any value for learning? If not, what is missing?

2. Choose a learning theory that you feel is compatible with games. What kind of game (MMORPG, puzzle game, adventure game, first-person shooter, etc.) do you think it would be most compatible with? Why? What are the design implications of adopting that theory for a given game? Name one example of a specific design element in a game that was designed according to your theory.

3. In the game architecture case study of this chapter, the authors described some of the architecture of a game under development. What would you do to ensure the game is as engaging as it is effective? Design a formative evaluation protocol for assessing engagement, including the different audiences and mechanisms by which you would evaluated engagement and make design modifications.

4. Using the game described in the architecture case study, describe an approach to stealth assessment that could be built into the game. Be specific in addressing how it aligns with some learning outcome, how you would measure it, how you could integrate it surreptitiously, and how it could be used for assessment, to modify game performance in some way, or both.

References

Almond, R. G., & Mislevy, R. J. (1999). Graphical models and computerized adaptive testing. *Applied Psychological Measurement, 23*,(3), 223-237.

Bransford, J., Brown, A. L., & Cocking, R. R. (2000). *How People Learn: Brain, Mind, Experience, and School* (expanded edition) , Washington: National Academies Press.

Bransford, J. D., Franks, J. J., Vye, N. J., & Sherwood, R. D. (1989). New approaches to instruction: Because wisdom can't be told. In S. Vosniadou & A. Ortany (Eds.), *Similarity and analogical reasoning* (pp. 470-497). New York: Cambridge University Press.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*, 32-42.

Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review, 31*(1), 21–32.

Charsky, D. (in press).Making a connection: Where game design meets instruction design. In R. Van Eck (Ed.), *Gaming and cognition: Theories and perspectives from the learning sciences*,. Hershey, PA: IGI Global.

Conati, C., Gertner, A., & VanLehn, K. (2002). Using Bayesian networks to manage uncertainty in student modeling. *User Modeling & User-Adapted Interaction*, *12*(4), 371-417.

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

Elkind, D. (2007). *The power of play: How spontaneous, imaginative activities lead to happier, healthier children*. Cambridge, MA: Da Capo Lifelong.

Falmagne, J.-C., Cosyn, E., Doignon, J.-P., & Thiery, N. (2003). The assessment of knowledge, in theory and in practice. In R. Missaoui & J. Schmidt (Eds.), *Lecture notes in computer science*: Vol. 3874: 4th international conference on formal concept analysis (pp. 61–79). New York: Springer-Verlag.

Gagné, R. M., Wager, W. W., Golas, K. C., & Keller, J. M. (2005). *Principles of instructional design* (5th ed.). Belmont, CA: Wadsworth/Thomson Learning.

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

Gee, J. P. (2007). Games and learning: Issues, perils and potentials. In J. P. Gee (Ed.), *Good video games and good learning: Collected essays on video games, learning and literacy* (pp. 129-174). New York: Palgrave/Macmillan.

Gee, J. P. (2008). Video games, learning, and "content." In C. Miller (Ed.), *Games: Purpose and Potential in Education*. Boston, MA: Springer.

Goodnough, K. & Hung, W. (2008). Designing effective problems: Evaluation of 3C3R 9-step design process, *Interdisciplinary Journal of Problem-based Learning, 2*(2), 61-90.

Gaydos, M., & Squire, K. D. (in press). Citizen science: Designing a game for the 21st century. In *Interdisciplinary Models and Tools for Serious Games: Emerging Concepts and Future Directions*, Richard Van Eck, Editor. Hershey, PA: IGI Global.

Huizinga, J. (1950). *Homo Ludens: A study of the play element in culture*. Boston, MA: Beacon Press.

Hung, W. (2006). The 3C3R model: A conceptual framework for designing problems in PBL, *Interdisciplinary Journal of Problem-based Learning, 1*(1), 55-77.

Hung, W. (2009). The 9-Step process for designing PBL problems: Application of the 3C3R model, *Educational Research Review, 4*(2), 2009, 118-141.

Hung, W., & Van Eck, R. (in press). Aligning problem solving and gameplay: A model for future research & design. In R. Van Eck (Ed.) *Interdisciplinary Models and Tools for Serious Games: Emerging Concepts and Future Directions*, Hershey, PA: IGI Global.

Jonassen, D. H. (1997). Instructional design models for well-structured and ill-structured problem solving learning outcomes. *ETR&D*, *45*(1), 65-94.

Jonassen, D. H. (2000). Toward a design theory of problem solving. *ETR&D, 48*(4), 63-85.

Jonassen, D. H. (2002). Integration of problem solving into instructional design. In R. A. Reiser & J. V. Dempsey (Eds.), *Trends and Issues in Instructional Design & Technology* (pp. 107-120). Upper Saddle River, NJ: Merrill Prentice Hall.

Jonassen, D. H., & Hung, W. (2008). All problems are not equal: Implications for PBL. *Interdisciplinary Journal of Problem-Based Learning, 2*(2), 6-28.

Lee, J. (1999). Problem-based learning: A decision model for problem selection, *Proceedings of selected research and development papers presented at the National Convention of the Association for Educational Communications and Technology, Houston, TX*.

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Education Researcher, 32*(2), 13–23.

Mislevy, R.J., (1994). Evidence and inference in educational assessment. *Psychometrika, 59*, 439-483

Mislevy, R. J., & Haertel, G. D. (2006). Implications of evidence-centered design for educational testing. *Educational Measurement: Issues and Practice, 25*(4), 6-20.

Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2004). *A brief introduction to evidence-centered design* (CSE Report 632). CA: Center for Research on Evaluation, Standards, and Student Testing. (ERIC Document Reproduction Service No. ED483399)

Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). On the structure of educational assessment. *Measurement: Interdisciplinary Research and Perspective, 1* (1) 3–62.

National Science Education Standards (1996). Washington, DC: National Academies Press.

Pellegrini, A. D. (Ed.). (1995). *The future of play theory: A multidisciplinary inquiry into the contributions of Brian Sutton-Smith*. Albany, NY: State University of New York Press.

Phillips, J. L. (1981). *Piaget's theory: A primer*. San Francisco, CA: W.H. Freeman.

Piaget, J. (1951). *Play, dreams, and imitation in childhood*. New York: W.W. Norton & Company.

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

Quinn, C. (2005). *Engaging learning: Designing e-learning simulation games.* Pfeiffer: San Francisco.

Rieber, L. P. (1996). Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games. *Educational Technology Research & Development, 44*(2), 43-58.

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

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153-189.

Shute, V. J., Graf, E. A., & Hansen, E. (2005). Designing adaptive, diagnostic math assessments for individuals with and without visual disabilities. In L. PytlikZillig, R. Bruning, and M. Bodvarsson (Eds.). *Technology-based education: Bringing researchers and practitioners together* (pp. 169-202). Greenwich, CT: Information Age Publishing.

Shute, V. J., Ventura, M., Bauer, M. I., & Zapata-Rivera, D. (2009). Melding the power of serious games and embedded assessment to monitor and foster learning: Flow and grow. In U. Ritterfeld, M. Cody, & P. Vorderer (Eds.), *Serious games: Mechanisms and effects* (pp. 295-321). Mahwah, NJ: Routledge, Taylor and Francis.

Society for Science and the Public (2008). *International rules for precollege science research: Guidelines for science and engineering fairs*. Washington, D.C.: Society for Science and the Public.

Sutton-Smith, B. (1997). *The ambiguity of play.* Cambridge, MA: Harvard University Press.

Thai, A., Lowenstein, D., Ching, D., & Rejeski, D. (2009). *Game changer: Investing in digital play to advance children's learning and health*. New York, NY: The Joan Ganz Cooney Center at Sesame Workshop.

Van Eck, R. (2006). Building intelligent learning games. In D. Gibson, C. Aldrich, & M. Prensky (Eds.) *Games and simulations in online learning: Research & development frameworks*. Hershey, PA: Idea Group.

Van Eck, R. (2007). Six ideas in search of a discipline. In B. Shelton & D. Wiley (Eds.), *The educational design and use of computer simulation games*. Boston, MA: Sense.

Van Eck, R. (2008). COTS in the classroom: A teachers guide to integrating commercial off-the-shelf (COTS) games. In R. Ferdig (Ed.) *Handbook of Research on Effective Electronic Gaming in Education*, Hershey, PA: Idea Group.

Van Eck, R., Hung, W., Bowman, R., & Love, S. (2009). 21st century game design: A model and prototype for promoting scientific problem solving. *Proceedings of the Twelfth IASTED International Conference on Computers and Advanced Technology in Education: Globalization of Education Through Advanced Technology*. Calgary, Canada: ACTA Press.

Van Lehn, K., Lynch, C., Schulze, K., Shapiro, J. A., Shelby, R., Taylor, L. Treacy, D. Weinstein, A., & Wintersgill, M. (2005). The Andes physics tutoring system: Lessons learned, *International Journal of Artificial Intelligence and Education*, *15* (3), 1-47.

Vygotsky, L.S. (1978). *Mind in society: The development of higher mental processes*. Cambridge, MA: Harvard University Press

Vygotsky, L. S. (1987). *The collected works of L. S. Vygotsky*. New York: Plenum.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 25*(1), 3-17.

Zimmerman, B. J. (2008). Investigating Self-Regulation and Motivation: Historical Background, Methodological Developments, and Future Prospects. *American Educational Research Journal, 45*(1), 166-183.