Chapter 4
Games, Learning, and Assessment

Valerie J. Shute and Fengfeng Ke

4.1 Introduction

Scholars from various disciplines have recently shown increasing interest in using well-designed digital games to support learning (e.g., Gee, 2003; Prensky, 2006; Shaffer, Squire, Halverson, & Gee, 2005; Shute, Rieber, & Van Eck, 2011). A common motivation for studying games as vehicles to support learning is frustration with the current education system and a desire for alternative ways of teaching—ways that increase student engagement and yield a rich, authentic picture of the learner(s).

Frustration stems from the fact that most schools in the U.S. are not adequately preparing kids for success in the twenty-first century (e.g., Partnership for 21st Century Skills, 2006). Learning in school is still heavily geared toward the acquisition of content within a teacher-centered model, with instruction too often abstract and decontextualized and thus not suitable for this age of complexity and interconnectedness (Shute, 2007). One downside of this outdated pedagogy is that other developed countries of the world are surpassing the U.S. on measures of important competencies (e.g., mathematics problem solving) as assessed by international tests such as the PISA and TIMSS (Gonzales et al., 2008; Howard, Paul, Marisa, & Brooke, 2010).

To make the problem with today's schools clearer, consider the following scenario involving a prototypical student. Maya (13 years old) is sitting in her bedroom with two of her friends. They are playing *Little Big Planet*—a digital game involving sack-person characters, clever and complex problems to solve, and compelling music and graphics. The game can not only be played (for countless hours), but it also provides tools to develop one's own levels and worlds which can then be shared

V.J. Shute (✉) • F. Ke
Florida State University, 3205C Stone Building, 1114 West Call Street,
Tallahassee, FL 32306-4453, USA
e-mail: vshute@fsu.edu; fke@fsu.edu

and played with the rest of the Internet community. Fully engaging in the game requires problem solving skills, persistence, and creativity—i.e., competencies which are increasingly critical to success in the twenty-first century but are not supported by our current educational system.

Like so many young people today, Maya and her friends are bored with school, and their mediocre grades reflect that attitude. But if Maya's teachers could see what she was doing in Little Big Planet, their views of her as a "slacker" would be quite different. For instance, Maya created and uploaded a new level in the game and is showing it to her friends—both in her bedroom and all over the world via the Internet. Several weeks ago, she began by writing a creative storyline, and used the in-game toolbox to create a visually-stunning environment complete with actions and reactions in the environment that reflect highly sophisticated physics understanding (as well as a good command of AI programming skills that goes beyond what most of her teachers are capable of doing). She regularly contributes detailed descriptions of how she solved her various coding problems to the Little Big Planet discussion forum, crafting her messages so they communicate clearly to all of the Little Big Planet players. Is Maya completely wasting her time with this game when she could be studying for her science test (e.g., memorizing the parts of a cell) or writing an expository essay for English class (e.g., on "why someone you care about is important to you")?

To answer the question above and to be able to make the claim that Maya is indeed developing valuable skills like problem solving, creativity, and writing, we need to employ some type of valid assessment to understand what Maya is learning from playing the game, to what degree, and in which contexts. The main challenges involved with creating such an assessment is that it must be suitable for the dynamic nature of digital games, unobtrusive to the player, while not sacrificing reliability and validity in the process.

The purpose of this chapter is to take a closer look at issues relating to game-based assessment and learning. What are the core elements of a good game? Can good games be used to support learning, based on the cumulative findings of the literature? How can game-based learning be assessed without interrupting the engagement? To address these questions, we begin by defining games and learning, provide some examples of learning from games, and then present a new approach to dynamically and validly assess learning within game environments (i.e., evidence-based stealth assessment).

4.2 Games

According to Klopfer, Osterweil, and Salen (2009), games refer to structured or organized play. Play is voluntary, intrinsically motivating, and involves active cognitive and/or physical engagement that allows for the freedom to fail (and recover), to experiment, to fashion identities, and freedom of effort and interpretation (Klopfer et al., 2009; Pellegrini, 1995; Rieber, 1996). Different from "free

play," a game is usually a contest of physical or mental skills and strengths, requiring the player to follow a specific set of rules to attain a goal (Hogle, 1996).

A more succinct definition of "games" comes from Suits (1978), who describes games as, "unnecessary obstacles we volunteer to tackle." To illustrate this idea, he used the game of golf where the objective is to get the ball into the hole. The most obvious (and easiest) way to accomplish that goal is to just pick up the ball and put it in the hole. But when you include the rules of the game (e.g., you must hit the ball with a stick that has a small piece of metal on the end, while standing 200 yards or so away from the hole) and other challenges (e.g., sand traps), this makes the game much more difficult and thus all the more compelling. In games, these unnecessary obstacles become something that we want to overcome because reaching for goals and ultimately succeeding is highly rewarding. Games and their associated obstacles also create a positive kind of stress, called eustress, which is actually good for us, providing us with a sense of motivation and desire to succeed (McGonigal, 2011).

Taking a more componential tack, Prensky (2001) has argued that a game consists of a number of key elements: rules, goals and objectives, outcomes and feedback, conflict (or competition, challenge, opposition), interaction, and representation or story. Using Prensky's definition, a game differs from a simulation in that a game is intrinsically motivating and involves competition. A competitive format does not, however, require two or more participants (Dempsey, Haynes, Lucassen, & Casey, 2002). That is, if a simulation enables a learner to compete against him/herself by comparing scores over successive attempts at the simulation, or has a game structure imposed on the system, it is regarded as a type of game. If the focus of a simulation involves the completion of an event only, the simulation is not a game. In addition, a simulation generally requires representing certain key characteristics or behaviors of a selected real-world phenomenon or system. But not all games are created to simulate dynamic systems in reality. For instance, fantasy may be part of the game design.

4.2.1   Core Elements of Good Games

Diverse perspectives exist in the literature on what a good game should be. Gee (2009) recently defined six key properties for good digital games to promote deep learning: (a) an underlying rule system and game goal to which the player is emotionally attached; (b) micro-control that creates a sense of intimacy or a feeling of power; (c) experiences that offer good learning opportunities; (d) a match between affordance (allowing for a certain action to occur) and effectivity (the ability of a player to carry out such an action), (e) modeling to make learning from experience more general and abstract, and (f) encouragement to players to enact their own unique trajectory through the game (p. 78).

Other gaming scholars have focused on the playability of the game and player motivation in describing a good game (e.g., Fabricatore, Nussbaum, & Rosas, 2002;

Kirkpatrick, 2007; Yee, 2006). For example, Sweetser and Wyeth (2005) developed and validated an analytic model of game engagement called the *GameFlow model*. This model captures and evaluates a game's enjoyment or engagement quality through eight game flow elements, including concentration, challenge, player skills, control, clear goals, feedback, immersion, and social interaction. Each element encompasses a list of design criteria.

Concentration prescribes that games should provide stimuli from different sources to grab and maintain players' attention, but not burden players with trivial tasks or overload them beyond their cognitive, perceptual, and memory limits. Challenge in a game should match the player's skill level, be increased as the player progresses through the game, and allow for player-centered pacing. The element of player skills suggests that games should have an easy and user-friendly interface, provide a tutorial or online help that enables players' skill development as they progress through the game, and reward players for skill development. The element of control indicates that players should have a sense of control over the characters and movements in the game world, the game interface, and gameplay (i.e., actions and strategies players take or use when playing the game). Games should also present clear overall and intermediate goals, as well as provide immediate feedback and score status during the gaming process. As a result, games should support players becoming fully immersed in the game, losing a sense of time and environment in the process. Finally, games should support social interactions (including competition and cooperation) between players, and support social communities inside and outside the game.

By synthesizing the aforementioned findings from the literature and other discussions on good games, we have derived seven core elements of well-designed games that are presented below.

- *Interactive problem solving*: Games require ongoing interaction between the player and the game, which usually involves the requirement to solve a series of problems or quests.
- *Specific goals/rules*: Games have rules to follow and goals to attain which help the player focus on what to do and when. Goals in games may be implicit or explicit.
- *Adaptive challenges*: Good games balance difficulty levels to match players' abilities. The best games and instruction hover at the boundary of a student's ability.
- *Control*: A good game should allow or encourage a player's influence over gameplay, the game environment, and the learning experience.
- *Ongoing feedback*: Good games should provide timely information to players about their performance. Feedback can be explicit or implicit, and as research has indicated, has positive effects on learning.
- *Uncertainty* evokes suspense and player engagement. If a game "telegraphs" its outcome, or can be seen as predictable, it will lose its appeal.
- *Sensory stimuli* refer to the combination of graphics, sounds, and/or storyline used to excite the senses, which do not require "professional" graphics or sound to be compelling.

4.2.2 Good Games as Transformative Learning Tools

As many researchers have argued, good games can act as transformative digital learning tools to support deep and meaningful learning. Based on the situated learning theory (Brown, Collins, & Duguid, 1989), learning in a mindful way results in knowledge that is considered meaningful and useful, as compared to the inert knowledge that results from decontextualized learning strategies.

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). 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. With simulated visualization and authentic problem solving with instant feedback, computer games can afford a realistic framework for experimentation and situated understanding, hence can act as rich primers for active learning (Gee, 2003; Laurel, 1991).

In this chapter, learning is defined as a lifelong process of accessing, interpreting, and evaluating information and experiences, then translating the information/ experiences into knowledge, skills, values, and dispositions. It also involves change—from one point in time to another—in terms of knowing, doing, believing, and feeling. Prior research on games for learning usually focused on *content* learning in schools, such as learning the subjects of reading, writing, and mathematics. For example, major literature reviews on educational gaming research (Dempsey, Rasmussen, & Lucassen, 1996; Emes, 1997; Hays, 2005; Ke, 2008; Randel, Morris, Wetzel, & Whitehill, 1992; Vogel et al., 2006; Wolfe, 1997) have indicated that the majority of gaming studies have focused on content-specific learning. Learning in game studies encompasses the following subject areas: science education, mathematics, language arts, reading, physics, and health, among others (Ke, 2008). Substantially fewer studies to date have examined the development of cognitive processes in games (e.g., Alkan & Cagiltay, 2007; Pillay, 2002; Pillay, Brownlee, & Wilss, 1999).

While games can support content learning, we believe that games are actually better suited to support more complex competencies. As many researchers have pointed out (e.g., Gee, 2003; Malone & Lepper, 1987; Rieber, 1996), games, as a vehicle for play, can be viewed as a natural cognitive tool or toy for both children and adults (Hogel, 1996). And rather than being used as a means to achieve an external goal (e.g., learning mathematics), games are often made to align with players' intrinsic interests and challenge learners to use skills they would not otherwise tend to use (Malone & Lepper, 1987), thus enabling the design of intrinsically motivating environments, with knowledge and skill acquisition as a positive by-product of gameplay.

Besides providing opportunities for play, games enable extensive and multiple types of cognitive learning strategies. For example, games can be used as an anchor for learning-by-design to reinforce creativity of learners (Kafai, 2005). Games can involve players in forming, experimenting with, interpreting, and adapting playing strategy in order to solve problems, thus enabling players to practice persistent

problem solving (Kiili, 2007). Games can also be developed as dynamic systems with which players can observe and play out key principles inherent in the systems, and hence develop organizational and systemic thinking skills (Klopfer et al., 2009). Finally, games can express and inspire certain underlying epistemic frames, values, beliefs, and identities (Shaffer, 2005).

There is a convergence between the core elements of a good game and the characteristics of productive learning. The constructivist problem-based and inquiry learning methods indicated the success of learning in the context of challenging, open-ended problems (Hmelo-Silver, 2004). Goal-based scenarios have long been viewed as an active primer for situated learning (Bransford et al., 2000). Correspondingly, in a good game a player is involved in an iterative cycle of goal-based, interactive problem solving. 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. 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. Moreover, a good game reinforces a sense of control—a critical metacognitive component for self-regulated learning (Zimmerman & Schunk, 2001). Similarly, both well-designed games and productive learning processes employ ongoing feedback as a major mechanism of play/learning support. Finally, the literature on the contribution of curiosity for learning motivation (Krapp, 1999) and the critical role of sensory memory in information processing (Anderson, 1995) is closely connected with the discussion of uncertainty and sensory stimuli in good games.

The problem with offering a game as a transformative learning tool to support complex competencies is that its effectiveness often cannot be directly or easily measured by traditional assessment instruments (e.g., multiple-choice tests). Implicit learning occurs when players are not consciously intending to learn some content. Therefore, focusing solely on knowledge-test-scores as outcomes is too limited since the games' strength lies in supporting emergent complex skills.

4.3 Evidence of Learning from Games

Following are four examples of learning from digital games that represent commercial as well as educational games. Preliminary evidence suggests that students can learn deeply from such games, and acquire important twenty-first century competencies.

4.3.1 Deep Learning in Civilization

Our first example illustrates how a commercial digital game can be used to support deep learning of history. Kurt Squire, at the University of Wisconsin, used a strategy

game called Civilization in a high school world history class (Squire, 2004). The goal of this game is to build, advance, and protect a civilization. This game starts with kids picking a civilization that they want to build (e.g., ancient Mesopotamia). Kids make many decisions about how to build and grow their civilization. Sometimes their decisions can be as simple as deciding where to put a new bridge, but they can be as complex as deciding whether to start a nuclear war. To make successful decisions, a player needs to consider important elements of human history, including economy, geography, culture, technology advancement, and war.

So what do kids learn from playing this game? Squire reported that players mastered many historical facts (e.g., where Rome was located), but more importantly, at the end of the game, they took away a deep understanding about the intricate relationships involving geographical, historical, and economic systems within and across civilizations.

4.3.2 Gamestar Mechanic and Systems Thinking

Our next example illustrates how digital games can be used to support systems thinking skill. Systems thinking skill refers to a particular way of looking at the world which involves seeing the "big picture" and the underlying interrelationships among the constituent elements rather than just as isolated bits. Gamestar Mechanic is an online game that is intended to teach kids basic game design skills and also allows them to actually build their own games for themselves, friends, and family to play. To design a functioning and challenging game in Gamestar Mechanic, players need to think hard about various game elements, parameters, and their interrelationships. If they think too simply, and just change a few elements of the game without considering the whole system, the game will not work.

For example, consider a player who included too many enemies in her game (each one with full strength). The consequence of this decision would be that other players would not be able to beat the game, so it would not be any fun. With a little reflection, she would realize the impact that the number/strength of enemies feature of the game would have on other elements of the game, and revise accordingly. Torres (2009) recently reported on his research using Gamestar Mechanic. He found that kids who played the game did, in fact, develop systems thinking skills along with other important skills such as innovative design.

4.3.3 Epistemic Games

Another example of a type of digital game that supports learning is the epistemic game. An epistemic game is a unique game genre where players virtually experience the same things that professional practitioners do (e.g., urban planner, journalist, and engineer). Epistemic games are being developed by Shaffer and his research team at the University of Wisconsin-Madison (Shaffer, 2007). These games are

based on the idea that learning means acquiring and adopting knowledge, skills, values, and identities that are embedded within a particular discipline or professional community. For example, to really learn engineering means being able to think, talk, and act like an engineer.

One example of an epistemic game is Urban Science. In Urban Science, players work as interns for an urban and regional planning center. Players as a group develop landscape planning proposals for the mayor of the city where they live. As part of the game play process, they first conduct a site visit interviewing virtual stakeholders in the area to identify different interests. For instance, some stakeholders may want a parking garage while others want affordable housing. Players need to consider various social and economic impacts of their decisions. They also use a special mapping tool called iplan (which is a tool similar to an actual Geographic Information System) to come up with their final planning. Towards the end of the game, they write their final proposal to the mayor discussing strengths and weaknesses of their final planning ideas.

4.3.4   Taiga Park and Science Content Learning

Our last example illustrates how kids learn science content and inquiry skills within an online game called Quest Atlantis: Taiga Park. Taiga Park is an immersive digital game developed by Barab et al. at Indiana University (Barab, Gresalfi, & Ingram-Goble, 2010; Barab et al., 2007). Taiga Park is a beautiful national park where many groups co-exist, such as the fly-fishing company, the Mulu farmers, the lumber company, and park visitors. In this game, Ranger Bartle calls on the player to investigate why the fish are dying in the Taiga River. To solve this problem, players are engaged in scientific inquiry activities. They interview virtual characters to gather information, and collect water samples at several locations along the river to measure water quality. Based on the collected information, players make a hypothesis and suggest a solution to the park ranger.

To move successfully through the game, players need to understand how certain science concepts are related to each other (e.g., sediment in the water from the loggers' activities causes an increase to the water temperature, which decreases the amount of dissolved oxygen in the water, which causes the fish to die). Also, players need to think systemically about how different social, ecological, and economical interests are intertwined in this park. In a controlled experiment, Barab et al. (2010) found that the middle school students learning with Taiga Park scored significantly higher on the posttest (assessing knowledge of core concepts such as erosion and eutrophication) compared to the classroom condition. The same teacher taught both treatment and control conditions. The Taiga Park group also scored significantly higher than the control condition on a delayed posttest, thus demonstrating retention of the content relating to water quality.

As these examples show, digital games appear to support learning. But how can we more accurately measure learning, especially as it happens (rather than after the

fact)? The answer is not likely to be via multiple choice tests or self-report surveys as those kinds of assessments cannot capture and analyze the dynamic and complex performances that inform twenty-first century competencies. A new approach to assessment is needed.

4.4 Assessment in Games

In a typical digital game, 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 another object increases your inventory of goods. In addition, solving major problems in games permits players to gain rank or "level up." 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 current levels of systems thinking skill, creativity, and teamwork, where each of these competencies is further broken down into constituent knowledge and skill elements (e.g., teamwork may be broken down into cooperating, negotiating, and influencing skills). If the estimated values of those competencies got too low, the player would likely feel compelled to take action to boost them.

4.4.1 Evidence-Centered Design

One main challenge for educators who want to employ or design games to support learning involves 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 can yield real-time estimates of students' competency levels across a range of knowledge and skills (Mislevy, Steinberg, & Almond, 2003).

ECD is a conceptual framework that can be used to develop assessment models, which in turn support the design of valid assessments. The goal is to help assessment designers coherently align (a) the claims that they want to make about learners, and (b) the things that learners say or do in relation to the contexts and tasks of interest (for an overview, see Mislevy & Haertel, 2006; Mislevy et al., 2003). There are three main theoretical models in the ECD framework: competency, evidence, and task models.

The competency model consists of student-related variables (e.g., knowledge, skills, and other attributes) on which we want to make claims. For example, suppose that you wanted to make claims about a student's ability to "design excellent

presentation slides" using MS PowerPoint. The competency model variables (or nodes) would include technical as well as visual design skills. The evidence model would show how, and to what degree, specific observations and artifacts can be used as evidence to inform inferences about the levels or states of competency model variables. For instance, if you observed that a learner demonstrated a high level of technical skill but a low level of visual design skill, you may estimate her overall ability to design excellent slides to be approximately "medium"—if both the technical and aesthetic skills were weighted equally.

The task model in the ECD framework specifies the activities or conditions under which data are collected. In our current PowerPoint example, the task model would define the actions and products (and their associated indicators) that the student would generate comprising evidence for the various competencies.

There are two main reasons why we believe that the ECD framework fits well with the assessment of learning in digital games. First, in digital games, people learn in action (Gee, 2003; Salen & Zimmerman, 2005). That is, learning involves continuous interactions between the learner and the game, so learning is inherently situated in context. Therefore, the interpretation of knowledge and skills as the products of learning cannot be isolated from the context, and neither should assessment. The ECD framework helps us to link what we want to assess and what learners do in complex contexts. Consequently, an assessment can be clearly tied to learners' actions within digital games, and can operate without interrupting what learners are doing or thinking (Shute, 2011).

The second reason that ECD is believed to work well with digital games is because the ECD framework is based on the assumption that assessment is, at its core, an evidentiary argument. Its strength resides in the development of performance-based assessments where what is being assessed is latent or not apparent (Rupp, Gushta, Mislevy, & Shaffer, 2010). In many cases, it is not clear what people learn in digital games. However in ECD, assessment begins by figuring out just what we want to assess (i.e., the claims we want to make about learners), and clarifying the intended goals, processes, and outcomes of learning.

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 (e.g., Csikszentmihalyi, 1900) and Vygotsky's zone of proximal development (Vygotsky, 1978).

4.4.2   Stealth Assessment

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 is where "stealth assessment" comes in (Shute, 2011; Shute, Ventura, Bauer, & Zapata-Rivera, 2009). Very simply, stealth assessment refers to ECD-based assessments that are

woven directly and invisibly into the fabric of the learning environment. During game play, 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., scientific inquiry skills, creative problem solving). Evidence needed to assess the skills is thus provided by the players' interactions with the game itself (i.e., the processes of play), which can be contrasted with the product(s) of an activity—the norm in educational 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 the subsequent actions later on. Second, in traditional tests, questions are often designed to measure particular, individual pieces 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), stealth assessments within game environments can infer what learners know and do not know at any point in time. 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 above, evidence-based stealth assessments can address these problems.

As a brief example of stealth assessment, Shute et al. (2009) used a commercial video game called Oblivion (i.e., *The Elder Scrolls*® *IV*: *Oblivion*©, 2006, by Bethesda Softworks) and demonstrated how assessment can be situated within a game environment and the dynamic student data can be used as the basis for diagnosis and formative feedback. A competency model for creative problem solving was created, which was divided into two parts—creativity and problem solving. These, in turn, were divided into novelty and efficiency indicators which were tied to particular actions one could take in the game. Different actions would have different impacts on relevant variables in the competency model. For instance, if a player came to a river in the game and dove in to swim across it, the system would recognize this as a common (not novel) action and automatically score it accordingly (e.g., low on novelty). Another person who came to the same river but chose to use a spell to freeze the river and slide across would be evidencing more novel (and efficient) actions, and the score for the creative variable in the competency model would be updated accordingly.

The models are updated via Bayesian inference networks (or Bayes nets). That is, the model of a student's game-play performance (i.e., the "student model") accumulates and represents probabilistic belief about the targeted aspects of skill, expressed as probability distributions for competency-model variables (Almond & Mislevy, 1999). 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 models express situations that can evoke required evidence.

One upside of the evidence-based stealth assessment approach relates to its ability to assess general and content-specific learning in games. That is, stealth assessment is able to assess a range of attributes—from general abilities or dispositions (e.g., problem solving, creativity, and persistence) to content-specific learning (e.g., water quality, physics concepts), or even current beliefs.

4.5   Conclusion

At the beginning of this chapter we listed several questions and attempted to answer them throughout. That is, we (a) described a set of core elements of a well-designed game distilled from the literature, (b) presented examples of research studies where games were shown to support learning, and (c) discussed an approach to game-based learning using stealth assessment techniques. Our stealth assessment approach involves the use of ECD which enables the estimation of students' competency levels and further provides the evidence supporting claims about competencies. Consequently, ECD has built-in diagnostic capabilities that permits a stakeholder (i.e., the teacher, student, parent, and others) to examine the evidence and view the current estimated competency levels. This in turn can inform instructional support or provide valuable feedback to the learner.

While there seems to be a lot of promise in relation to the evidence-based stealth assessment idea, what are some of the downsides or possible limitations of this approach? First, Rupp et al. (2010) noted that when developing games that employ ECD for assessment design, the competency model must be developed at an appropriate level of granularity to be implemented in the assessment. Too large a grain size means less specific evidence is available to determine student competency, while too fine a grain size means a high level of complexity and increased resources to be devoted to the assessment. Second, the development costs of ECD-based assessments can be relatively high for complex competencies. To counter this obstacle, we are currently exploring ways to create stealth assessment models that can be used in related but different games (i.e., in a plug-and-play manner). Creating such cross-platform models for digital games would be useful and cost effective for educators interested in using games for assessment and support of learning. Finally, some people may not be "into games" thus there may be individual (or cultural) differences relating to prior game experience or differential interests that affect learning. That is, certain personal or cultural variables may be identified that interact, mediate, or moderate the effects of gameplay on learning. This is all valuable future research to pursue.

In conclusion, the world is changing rapidly but education is not. Preparing our kids to succeed in the twenty-first century requires fresh thinking on how to foster new competencies. There's an associated need to design and develop valid and reliable assessments of these new skills. We have suggested that ECD should be used

as the framework for developing new assessments that can yield valid measures; provide accurate estimates of complex competencies embedded in dynamic performances; and aggregate information from a variety of sources. We also believe that well-designed games can serve as one excellent type of learning environment because games are intrinsically motivating and can facilitate learning of academic content and twenty-first century competencies within complex and meaningful environments. Such games can also promote social skills (like communication, collaboration, negotiation, and perspective taking), higher-order thinking skills (like problem solving and critical reasoning), and ownership of learning.

Designing evidence-based stealth assessments and weaving them directly within digital games will allow all kids to become fully engaged, to the point where they want (perhaps even demand) to play/learn, even outside of school. That is a lovely vision, especially in contrast with often frequent struggles to get kids to do their homework.

References

Alkan, S., & Cagiltay, K. (2007). Studying computer game learning experience through eye tracking. *British Journal of Educational Technology, 38*(3), 538–542.

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

Anderson, J. R. (1995). *Learning and memory: An integrated approach*. New York: Wiley.

Barab, S. A., Gresalfi, M., & Ingram-Goble, A. (2010). Transformational play. *Educational Researcher, 39*(7), 525–536.

Barab, S. A., Zuiker, S., Warren, S., Hickey, D., Ingram-Goble, A., Kwon, E.-J., et al. (2007). Situationally embodied curriculum: Relating formalisms and contexts. *Science Education, 91*(5), 750–782.

Bethesda Softworks (2006). *Elder schools VI: Oblivion*. Retrieved April 9, 2012, from http://www.bethsoft.com/games/games_oblivion.html.

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

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32–42.

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

Csikszentmihalyi, M. (1990). Flow: *The psychology of optical experience*. New York: Harper Perrennial.

Dempsey, J. V., Haynes, L. L., Lucassen, B. A., & Casey, M. S. (2002). Forty simple computer games and what they could mean to educators. *Simulation & Gaming, 33*(2), 157–168.

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

Emes, C. E. (1997). Is Mr Pac Man eating our children? A review of the effect of digital games on children. *Canadian Journal of Psychiatry, 42*(4), 409–414.

Fabricatore, C., Nussbaum, M., & Rosas, R. (2002). Playability in action videogames: A qualitative design model. *Human Computer Interaction, 17*(4), 311–368.

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.), *Fourth international conference on formal concept analysis* (Lecture notes in computer science, Vol. 3874, pp. 61–79). New York: Springer.

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

Gee, J. P. (2009). Deep learning properties of good digital games: How far can they go? In U. Ritterfeld, M. Cody, & P. Vorderer (Eds.), *Serious games: Mechanisms and effects* (pp. 65–80). New York: Routledge.

Gonzales, P., Williams, T., Jocelyn, L., Roey, S., Kastberg, D., & Brenwald, S. (2008). *Highlights from TIMSS 2007: Mathematics and science achievement of U.S. fourth- and eighth-grade students in an international context (NCES 2009–001)*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Hays, R. T. (2005). *The effectiveness of instructional games: A literature review and discussion*. Retrieved May 10, 2006, from http://adlcommunity.net/file.php/23/GrooveFiles/Instr_Game_Review_Tr_2005.pdf.

Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review, 16*(3), 235–266.

Hogle, J. G. (1996). *Considering games as cognitive tools: In search of effective "Edutainment"*. Retrieved January 12, 2005, from ERIC, ED 425737.

Howard, L. F., Paul, J. H., Marisa, P. P., & Brooke, E. S. (2010). *Highlights from PISA 2009: Performance of U.S. 15-year-old students in reading, mathematics, and science literacy in an international context (NCES 2011–004)*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Kafai, Y. B. (2005). The classroom as "living laboratory": Design-based research for understanding, comparing, and evaluating learning science through design. *Educational Technology, 65*(1), 28–34.

Ke, F. (2008). A qualitative meta-analysis of computer games as learning tools. In R. E. Ferdig (Ed.), *Handbook of research on effective electronic gaming in education* (pp. 1–32). New York: IGI Global.

Kiili, K. (2007). Foundation for problem-based gaming. *British Journal of Educational Technology, 38*(3), 394–404.

Kirkpatrick, G. (2007). Between art and gameness: Critical theory and computer game aesthetics. *Thesis Eleven, 89*, 74–93.

Klopfer, E., Osterweil, S., & Salen, K. (2009). *Moving learning games forward: Obstacles, opportunities & openness*. Cambridge, MA: The Education Arcade.

Krapp, A. (1999). Interest, motivation and learning: An educational-psychological perspective. *European Journal of Psychology of Education, 14*(1), 23–40.

Laurel, B. (1991). *Computers as theatre*. Reading, MA: Addison-Wesley.

Malone, T. W., & Lepper, M. R. (1987). Making learning fun: A taxonomy of intrinsic motivations for learning. In R. E. Snow & M. J. Farr (Eds.), *Aptitude, learning and instruction: III. Cognitive and affective process analyses* (pp. 223–253). Hilsdale, NJ: Erlbaum.

McGonigal, J. (2011). *Reality is broken: Why games make us better and how they can change the world*. New York: Penguin Press.

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

Partnership for 21st Century Skills. (2006). *Results that matter: 21st century skills and high school reform*. Retrieved from April 28, 2012. http://www.p21.org/documents/RTM2006.pdf.

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

Pillay, H. (2002). An investigation of cognitive processes engaged in by recreational computer game players: Implications for skills of the future. *Journal of Research on Technology in Education, 34*(3), 336–350.

Pillay, H., Brownlee, J., & Wilss, L. (1999). Cognition and recreational computer games: Implications for educational technology. *Journal of Research on Computing in Education, 32*(1), 203.

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

Prensky, M. (2006). *Don't bother me mom, I'm learning!: How computer and digital games are preparing your kids for 21st century success and how you can help!* St. Paul, MN: Paragon House.

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

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

Rieber, L. P. (1996). Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games. *Educational Technology Research and Development, 44*(1), 43–58.

Rupp, A. A., Gushta, M., Mislevy, R. J., & Shaffer, D. W. (2010). Evidence-centered design of epistemic games: Measurement principles for complex learning environments. *The Journal of Technology, Learning, and Assessment, 8*(4). Retrieved April 9, 2012, from http://escholarship. bc.edu/jtla/vol8/4.

Salen, K., & Zimmerman, E. (2005). Game design and meaningful play. In J. Raessens & J. Goldstein (Eds.), *Handbook of computer game studies* (pp. 59–80). Cambridge, MA: MIT Press.

Shaffer, D. W. (2005). *Studio mathematics: The epistemology and practice of design pedagogy as a model for mathematics learning*. Wisconsin Center for Education Research Working paper, No. 2005-3.

Shaffer, D. W. (2007). *How computer games help children learn*. New York: Palgrave.

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

Shute, V. J. (2007). Tensions, trends, tools, and technologies: Time for an educational sea change. In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 139–187). New York: Lawrence Erlbaum/Taylor & Francis.

Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. In S. Tobias & J. D. Fletcher (Eds.), *Computer games and instruction* (pp. 503–524). Charlotte, NC: Information Age.

Shute, V. J., Rieber, L., & Van Eck, R. (2011). Games … and … learning. In R. Reiser & J. Dempsey (Eds.), *Trends and issues in instructional design and technology* (3rd ed., pp. 321–332). Upper Saddle River, NJ: Pearson Education.

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.

Squire, K. (2004). *Replaying history: Learning world history through playing Civilization III*. ProQuest Dissertations, Indiana University.

Steinberg, L. S., & Gitomer, D. G. (1996). Intelligent tutoring and assessment built on an understanding of a technical problem-solving task. *Instructional Science, 24*, 223–258.

Suits, B. H. (1978). *The grasshopper: Games, life and utopia*. Toronto, ON: University of Toronto Press.

Sweetser, P., & Wyeth, P. (2005). GameFlow: A model for evaluating player enjoyment in games. *ACM Computers in Entertainment, 3*(3), 1–24.

Torres, R. J. (2009). Using Gamestar Mechanic within a nodal learning ecology to learn systems thinking: A worked example. *International Journal of Learning and Media, 1*(2), 1–11.

Vogel, J. F., Vogel, D. S., Cannon-Bowers, J., Bowers, C. A., Muse, K., & Wright, M. (2006). Computer gaming and interactive simulations for learning: A meta-analysis. *Journal of Educational Computing Research, 34*(3), 229–243.

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.

Wolfe, J. (1997). The effectiveness of business games in strategic management course work. *Simulation & Gaming, 28*(4), 360–376.

Yee, N. (2006). The demographics, motivations, and derived experiences of users of massively multi-user online graphical environments. *Presence: Teleoperators and Virtual Environments, 15*(3), 309–329.

Zimmerman, B. J., & Schunk, D. H. (2001). *Self-regulated learning and academic achievement: Theoretical perspectives*. Mahwah, NJ: Lawrence Erlbaum.