The Curriculum Journal
Vol. 16, No. 4, December 2005, pp. 455–474

Classification of learning outcomes: evidence from the computer games literature

Harold F. O'Neil[a]*, Richard Wainess[a] and Eva L. Baker[b]
[a]*University of Southern California/National Center for Research on Evaluation, Standards, and Student Testing (CRESST), USA;* [b]*UCLA/National Center for Research on Evaluation, Standards, and Student Testing (CRESST), USA*

Following up on an earlier issue of *The Curriculum Journal* (Vol. 16, No. 1), this article focuses on learning outcomes in the context of video games. Learning outcomes are viewed from two theoretical frameworks: Kirkpatrick's levels of evaluation and the CRESST model of learning. These are used to analyse the outcomes claimed in journal articles that report empirical work, indicating the usefulness of the frameworks, and the necessity to consider the role of affective learning. The article ends with some comments on the relationship of instructional design to effective games and learning outcomes.

Keywords: *Learning outcomes; Models of learning; Evaluation; Computer games; Motivation*

The purpose of this article is to review the last 15 years of empirical research on video games and learning outcomes for adults in the context of two theoretical frameworks: Kirkpatrick's four levels for evaluating training (1994) and the CRESST model of learning (Baker & Mayer, 1999). Based on this review, suggestions for classifications of learning outcomes will be offered.

Educators and trainers began to take notice of the power and potential of computer games for education and training back in the 1970s and 1980s (Malone, 1981; Ramsberger *et al.*, 1983; Malone & Lepper, 1987; Donchin, 1989; Thomas & Macredie, 1994; Ruben, 1999). Computer games were hypothesized to be potentially useful for instructional purposes and were also hypothesized to provide multiple benefits: (a) complex and diverse approaches to learning processes and outcomes; (b) interactivity; (c) ability to address cognitive as well as affective learning issues; and, perhaps most importantly, (d) motivation for learning.

*Corresponding author. 15366 Longbow Drive, Sherman Oaks, CA 91403, USA.
Email: honeil@usc.edu

ISSN 0958-5176 (print)/ISSN 1469-3704 (online)/05/040455–20

DOI: 10.1080/09585170500384529

While effectiveness of game environments can be documented in terms of intensity and longevity of engagement (participants voting with their money or time), as well as the commercial success of the games, there is much less solid information about what outcomes are systematically achieved by the use of individual and multiplayer games to train participants in acquiring knowledge and skills. For example, early studies of the game 'WEST' compared variations in coaching but found that the game itself produced little increment in learning (Baker *et al.*, 1985). Thus, the game was not instructionally more effective than traditional approaches. Similar results were found by Parchman *et al.* (2000), yet there is a persistent belief that games can provide an attractive venue for engaging participants in learning. They can produce rapid multiple trials, allow students to manage their feedback or to replay with a different strategy, and include social components, such as teamwork, of real value in the military.

What is missing is how games should be evaluated for education and training purposes. First is the degree to which they are designed to foster the desired key knowledge and skills. Second, the impact of game playing needs to be studied to determine what works (see O'Neil *et al.*, 2002; O'Neil & Fisher, 2004). Without an investment in evaluation and the accumulation of clear evidence of impact, there will be a tendency to dismiss game environments as motivational fluff.

There are multiple ways of assessing learning. For example, one could specify the assessment of the training effects of a game by examining trainees' ability to solve criterion problems, their application of declarative and procedural knowledge, their willingness to raise or lower game challenge conditionally, their self-reports and records of their play. Evaluation questions to be answered about the cognitive and affective effects of games should concern the four levels of Kirkpatrick's (1994) framework (these are explained in the next section): the evaluation system should include measures related to the attainment at different levels of expertise of the specific content and skill acquisition being trained. These may include skills to be learned along the way, as well as those of an intentional or unintentional outcome nature.

There is also a skill related to personal development—the 'learning to learn' or self-regulation outcome (O'Neil, 2002). To what degree do participants develop the strategic knowledge necessary to apply to the specific training topics? Do players develop more general predispositions and behaviours that support transfer of knowledge across different contexts or problem variations? The development of individual metacognitive skills, especially as used in time-constrained environments, should be estimated.

Maximizing skill and knowledge acquisition and retention in extremely short periods becomes tractable if there is comprehensive and accurate information on the trainee's background (e.g. quantitative and verbal aptitude, degree of prior knowledge, and experience in the training content), information on performance on the training task outcomes (e.g. quality of solution), and ongoing measures of behavioural (e.g. trainee's clickstream, that is, a series of clicks from a mouse and keyboard strokes) and conative (e.g. motivation, self-regulation) processes embedded within the task (e.g. measures of trainee understanding, or stress).

To examine some of these issues two literature searches were conducted using three search engines, PsycINFO, EducationAbs and SocialSciAbs. The purpose of the first search was to locate articles that reported on research about the use of video games in general, for training adults. The second search was to locate articles that reported research specifying the use of multiplayer or massively multiplayer video games for training adults. (Massively multiplayer games are an environment in which people role-play either real or imaginary characters in a real or imaginary setting. The number of players can vary from 40 to 100,000.) The literature review that follows is structured in the following manner. First, two theoretical frameworks for evaluation and learning are provided and related to each other. Then we define critical terminology (i.e. games, simulation and simulation games). Next, we provide the literature review results and a summary of the research. We view the game literature through two major theoretical frameworks: (a) Kirkpatrick's (1994) four levels for evaluating training and (b) Baker and Mayer's (1999) CRESST model of learning.

Kirkpatrick's four levels for evaluating training

Kirkpatrick described four levels that represent a sequence of ways to evaluate programs (see Figure 1). Level 1 is *Reaction*, Level 2 is *Learning*, Level 3 is *Behaviour* and Level 4 is *Results*. *Reaction* (Level 1) is an assessment of learner satisfaction and measures 'how those who participate in the program react to it' (1994, p. 21). Kirkpatrick commented that if learners do not react favourably to the training, they will probably not be motivated to learn. According to Kirkpatrick, 'Positive reactions may not ensure learning, but negative reaction almost certainly reduces the possibility of its occurring' (p. 22). He described *Learning* (Level 2) as 'the extent to which participants change attitudes, improve knowledge, and/or increase skill as a result of attending the program' (p. 22). Level 2 is evaluated within the context of the training session.

Behaviour (Level 3), which is evaluated on the job, after training has occurred, is defined as 'the extent to which change in behaviour has occurred because the participant attended the training program' (Kirkpatrick, 1994, pp. 22–23). Level 3 concerns the concept of *transfer*, which Brunken *et al.* (2003) described as the ability to apply acquired knowledge and skills to new situations. In Kirkpatrick's levels, new situations occur in the job context.

The final level, *Results* (Level 4), refers to the benefits from the company's perspective and can be defined as the 'final results that occurred because the participant attended the program' (Kirkpatrick, 1994, p. 25). Final results are related to cost effectiveness of training or return on investment and include increased production, improved quality, decreased costs, reduced frequency and/or severity of accidents, increased sales, reduced turnover and higher profits. Generally these results are the reason for attending the program. According to Kirkpatrick, when moving from level to level, the evaluation process becomes more difficult and time consuming, but provides more valuable information. Kirkpatrick further argued that

Four levels for evaluating training	
Level 1: REACTION	Trainee's reaction to the program: level of satisfaction
Level 2: LEARNING	Trainee's attitude change, increased knowledge, and/or increased skill, due to the training
Level 3: BEHAVIOUR	On the job change in behaviour because of program participation, i.e. transfer of learning to the job setting
Level 4: RESULTS	How the organization benefited from the learner's participation in the program (e.g. increased production or profits, improved quality, decreased costs, fewer accidents)

Figure 1. Kirkpatrick's four levels for evaluating training (1994)

all levels should be included in an evaluation and should be assessed in order of the levels. The Kirkpatrick framework is the dominant one in training evaluation but is seldom used in the education sector.

Baker and Mayer's CRESST model of learning

The CRESST model of learning (Baker & Mayer, 1999; see Figure 2) is composed of five families of cognitive demands: content understanding, collaboration or teamwork, problem solving, communication and self-regulation. In the CRESST model, 'each family consists of a task that can be used as a skeleton for the design of instruction and testing' (Baker & Mayer, 1999, p. 275). For example, content understanding involves explanation, which in turn involves a variety of actions such as having students read opposing views (of the content), invoking prior knowledge, and organizing and writing a valid explanation. Such knowledge and skills can be measured by multiple-choice tests, essays or knowledge maps. This framework supports many different learning domains, such as history or science. Problem solving is a family that is a superset of other families (i.e. problem solving consists of content understanding, self-regulation and problem-solving strategies; O'Neil, 2002). Each of these aspects of problem solving would have specific measures. Each of the other cognitive demands is also measured, either quantitatively or qualitatively (Baker & Mayer, 1999; Baker & O'Neil, 2002).

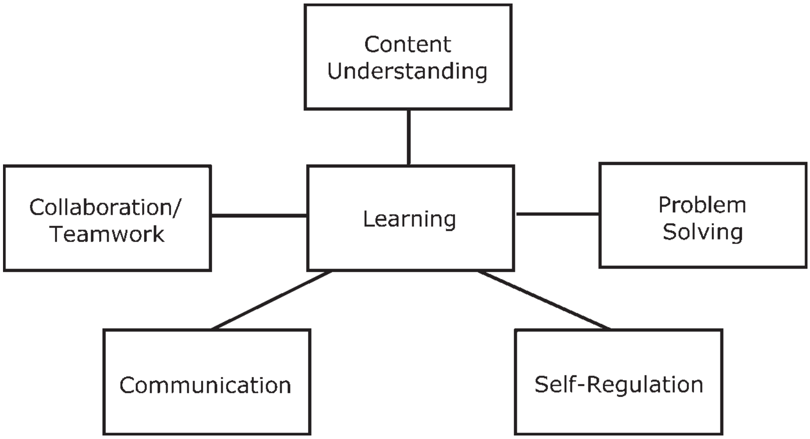

Figure 2. Baker and Mayer's CRESST model of learning: families of cognitive demands (1999)

Relating the frameworks

Kirkpatrick's (1994) four levels for evaluating training and Baker and Mayer's (1999) CRESST model of learning were not designed for the evaluation of game studies' outcomes *per se*. However, both frameworks were designed to evaluate learning. Since the goal of many game studies, and in particular the studies in this review, is to examine characteristics that can affect learning outcomes, these two frameworks are particularly relevant and complementary. Kirkpatrick's is the macro view in evaluation, and the CRESST model of learning is the micro view in the area of learning.

There is no relationship between Kirkpatrick's Level 1 (*Reaction*) and CRESST's (Baker & Mayer) model of learning, as Kirkpatrick's Level 1 refers to satisfaction, not learning. Likewise, there is no relation in Baker and Mayer's model to Kirkpatrick's Level 4 (*Results*), since *results* refers to institutional benefits, not learner benefits *per se*. However, Kirkpatrick's Level 2 (*Learning*) equates to the entire CRESST model, and Kirkpatrick's Level 3 (*Behaviour*) is reflected in transfer measures on the job, some of which would reflect the CRESST model of learning (e.g. problem solving). We will use the common elements of both frameworks in this article to review the literature in computer games.

Defining games, simulations and simulation games

A major problem area with reviews of research on games and simulations is terminology. The three most commonly used terms in the literature are *game*, *simulation* and *simulation game*, yet there is little consensus in the education and training literature on how these terms are defined. Because the goals and features of games and simulations (as well as the hybrid, simulation games) differ, it is important when examining the potential effects of the two media—games and simulations—to be clear about which one is being examined. Clearly defining the differences between games and simulations also provides a foundation for accurately defining the mixed mode environment of simulation games.

This article will use the definitions of games, simulations and simulation games as stated by Gredler (1996). These definitions combine the most common features cited by the various researchers, and yet provide clear distinctions between the three media. According to Gredler, 'games consist of rules that describe allowable player moves, game constraints and privileges (such as ways of earning extra turns) and penalties for illegal (non-permissible) actions. Further, the rules may be imaginative in that they need not relate to real-world events' (Gredler, 1996, p. 523).

This definition is in contrast to that of a simulation, which Gredler defined as 'a dynamic set of relationships among several variables that (1) change over time and (2) reflect authentic causal processes' (1996, p. 523). In addition, she described games as having a goal of winning whereas simulations have a goal of discovering causal relationships. Gredler also defined a mixed metaphor referred to as *simulation games* or *gaming simulations*, which is a blend of the features of the two interactive media, games and simulations.

In terms of goals, however, games and simulations differ. With games, once a goal is achieved, it cannot be repeated without intentionally reverting to a prior game state or restarting the game, that is, the flow of a game must be purposefully interrupted. In a game comprising multiple goals, achievement of one goal results in commencement of work toward the next goal or set of goals. Therefore, with games, the goal structure is linear. In contrast, simulations have non-linear goal structures. With simulations, the goal is to achieve a desired output state or simply to examine output states, based on the manipulation of input variables. Once the goal is achieved, the player can continually make modifications to the input variables, examining their effect on the output. Therefore, because this process can be repeated as often as desired, the goal structure of a simulation is non-linear.

While this article uses Gredler's (1996) definitions, it should be noted that other game researchers do not agree with her definitions. Table 1 lists a summary of the characteristics that various researchers have attributed to games and to simulations. We have included our extension of Gredler's reference to linearity and non-linearity by expressing linearity in terms of gameplay and goals separately. Gredler in defining games as linear and simulations as non-linear is referring to goal structure. In terms of play (i.e. interaction), both media are typically non-linear. A column for simulation games was not included in the table since, as its name implies, simulation games would have combinations of game and simulation characteristics.

Literature review results

A specific review to explore the classification of game literature using the frameworks of Kirkpatrick (1994) and the CRESST model of learning (Baker & Mayer, 1999) was conducted. Our literature review (covering the last 15 years) using information systems (i.e. PsycINFO, EducationAbs, SocialSciAbs) and the search terms *games, computer game, PC game, computer video game, video game,*

Table 1. Characteristics of games and simulations

Characteristic	*Game*	*Simulation*
Combination of one's actions plus at least one other's actions	Yes (via human or computer)	Yes
Rules	Defined by game designer/developer	Defined by system being replicated
Goals	To win	To discover cause–effect relationships
Requires strategies to achieve goals	Yes	Yes
Includes competition	Against computer or other players	No
Includes chance	Yes	Yes
Has consequences	Yes (e.g. win/lose)	Yes
System size	Whole	Whole or part
Reality or fantasy	Both	Both
Situation specific	Yes	Yes
Represents a prohibitive environment (due to cost, danger or logistics)	Yes	Yes
Represents authentic cause–effect relationships	No	Yes
Requires user to reach own conclusion	Yes	Yes
May not have definite end point	No	Yes
Contains constraints, privileges and penalties (e.g. earn extra moves, lose turn)	Yes	No
Linear goal structure	Yes	No
Linear intervention	No	No
Is intended to be playful	Yes	No

cooperation game and *multi-player game* resulted in only 18 articles with either qualitative or quantitative information on the effectiveness of games with adults as participants. We selected only journal articles for review. Thus, research based on dissertations or technical reports was not examined as these are not peer-reviewed. We also conducted by hand a journal search for 2004/5 and found one additional journal article that met our criteria. Thus there were a total of 19 journal articles. We then examined the viability of contextualizing games research using two frameworks: the Kirkpatrick (1994) levels and the CRESST model of learning (Baker & Mayer, 1999).

In the past 15 years, several thousand articles pertaining to games have been published. However, only 19 studies met our standards for empirical research (i.e. either qualitative or quantitative data available) and are reviewed in this article. As cited earlier, findings regarding the educational benefits of games are mixed, and it is hypothesized that the positive findings can be attributed to instructional design and not to games *per se*. Also discussed earlier was the issue that many studies claiming positive outcomes appear to be making unsupported claims for the media. These issues appear to be echoed in the studies we reviewed. For example, Mayer *et al.* (2002) examined performance outcomes using retention and transfer tests, and Carr and Groves (1998) examined performance outcomes using self-report surveys. Mayer *et al.* (2002) offered strong statistical support for their findings, using retention and transfer tests, whereas Carr and Groves used only participants' self-reports as evidence of learning effectiveness. In Carr and Groves's study, participants reported their belief that they learned something from the experience. No cognitive performance was actually measured, yet Carr and Groves suggested that their simulation game was a useful educational tool, and that use of the tool provided a valuable learning experience.

Table 2 lists the media and measures utilized by the 19 studies to assess learning. We have also categorized the measures in Table 2 as to Kirkpatrick's (1994) levels and also type of learning in terms of Baker and Mayer's (1999) framework. As may be seen in Table 2, of the 19 studies, five used a single measure to examine outcomes. For example, Carr and Groves (1998) used self-report via a survey instrument, and another study (Rosenorn & Kofoed, 1998) used observation.

The remaining studies used multiple measurements. For example, Day *et al.* (2001) used both performance on the Space Fortress game and a knowledge map. Likewise, Gopher *et al.* (1994) used both performance on Space Fortress II and airplane flight performance.

In terms of Kirkpatrick's (1994) four levels of evaluation, in Table 2, five of the studies reviewed involved both Level 1 (reaction to training) and Level 2 (learning during training), 11 studies involved Level 2 only (learning), only two studies involved Level 3 (on-the-job changes due to training), and one study involved Level 4 (benefits to the employer, for example: cost effectiveness). This suggests that while the Kirkpatrick model is appropriate for evaluating training programs, it is also a useful model for evaluating games studies. These results are different from those of the American Society for Training and Development (Sugrue & Kim, 2004) regarding the percentage of organizations using evaluation methods, in which the vast majority of evaluation methods were Level 1 (74%) and very few were Levels 2 (31%), 3 (14%) or 4 (8%). The difference between our study and the general training literature is most probably due to journal bias for significant results and our use of the criterion that there must be either qualitative or quantitative data.

The studies were also evaluated against the CRESST model of learning (Baker & Mayer, 1999). We counted a category to be present if it was explicitly measured; for example, almost all studies that we categorized as 'problem solving' used performance on the game. Only a few explicitly measured content understanding; for example, Day *et al.* (2001) used a knowledge map as a measure of content understanding. Likewise, several studies investigated the role of collaboration (e.g. Arthur *et al.*, 1995) but did not explicitly measure it. In terms of the CRESST model, seven of the studies involved content understanding, one involved collaboration, 16

Table 2. Media, measures, Kirkpatrick (1994) levels and CRESST categories

Study	*Media*[a]	*Measures*	*Kirkpatrick levels*[b]	*CRESST categories*
Arthur *et al.* (1995)	Space Fortress (a)	Performance on game, visual attention	2	Problem solving
Carr and Groves (1998)	Business Simulation in Manufacturing Management (c)	Survey	1, 2	Content understanding, collaboration, problem solving
Day *et al.* (2001)	Space Fortress (a)	Performance on game and knowledge map	2	Content understanding, problem solving
Galimberti *et al.* (2001)	3D-Maze (a)	Observation and time to complete game	2	Problem solving
Gopher *et al.* (1994)	Space Fortress II (a)	Performance on game and flight performance	2, 3	Problem solving
Green and Bavelier (2003)	Medal of Honor (c)	Visual attention	2	Problem solving
Green and Flowers (2003)	Video catching task (c)	Performance on game, exit questionnaire	2	Problem solving
Mayer *et al.* (2002)	Profile Game (c)	Performance on retention and transfer tests	2	Content understanding, problem solving
Moreno and Mayer (2000)	Design-a-Plant (b)	Performance on retention and transfer tests, plus survey	1, 2	Content understanding, problem solving
Morris *et al.* (2004)	Delta Force (c)	Performance on game, stress questionnaire, observation of military tactics used	2	Content understanding, problem solving
Parchman *et al.* (2000)	Adventure Game (a)	Retention test, transfer test, motivation questionnaire	2	Content understanding
Porter *et al.* (1990–1)	Whale Game (a)	Performance on game, satisfaction survey	1, 2	Problem solving
Prislin *et al.* (1996)	Space Fortress (a)	Performance on game, observation, discussion behaviour	2	Communication, problem solving

(*continued*)

Table 2. (*Continued*)

Study	*Media*[a]	*Measures*	*Kirkpatrick levels*[b]	*CRESST categories*
Rhodenizer *et al*. (1998)	AIRTANDEM (b)	Performance on game, retention tests	2	Content understanding, problem solving
Ricci *et al*. (1996)	QuizShell (b)	Performance on pre-, post- and retention tests, and trainee reaction questionnaire	1, 2	Content understanding
Rosenorn and Kofoed (1998)	Experiment Arium (b)	Observation	1, 2	None: affective learning, e.g. self-respect
Shebilske *et al*. (1992)	Space Fortress (a)	Performance on game	2, 4	Problem solving
Shewokis (2003)	Winter Challenge	Performance on game	2	Problem solving
Tkacz (1998)	Maze game (c)	Performance on game, transfer test of position location	2, 3	Content understanding, problem solving

[a]Letters in parentheses indicate type of media: (a) = game; (b) = simulation; (c) = simulation game.
[b]1 = reaction at training event, 2 = learning at training event, 3 = behaviour change on the job or transfer, and 4 = institutional benefits.

involved problem solving (domain specific), with one also involving communication. None measured self-regulation (metacognition, motivation, or both). It should be noted that the vast majority of the studies could be considered problem solving. Furthermore, all of the game studies involved motor learning (e.g. use of keyboard or joystick). Finally, in Table 2, transfer (Kirkpatrick's [1994] Level 3) is defined as at the job site. This definition, rather than transfer to a new training situation, resulted in few Level 3 studies, although a few studies measured transfer in the training environment. These results suggest that while the CRESST model is appropriate for evaluating K-16 (i.e. kindergarten through completion of a four-year college) learning, it is also a useful model for evaluating games studies with adults. These results further indicate that constructs investigated by empirical studies on the use of games to train adults are similar to the families of constructs that define the CRESST model of learning.

Because one of the major claimed advantages of games is that they are motivational, the existing CRESST framework needs to be augmented with an affective or motivational view of learning to be more useful for evaluating games and simulations. The current CRESST framework deals with motivation only in that the self-regulation component is composed of motivation (effort, self-efficacy), and the team skills component includes interpersonal skills as well as leadership, decision making, communication, adaptability. However, with these exceptions, the focus of the CRESST framework clearly has a cognitive bias.

Recent research offers some suggestions of what motivational factors to include. We have focused on those that predict academic achievement. For example, Robbins *et al.* (2004) provided an interesting basis for such a framework in a meta-analysis of the skill factors that predict college outcomes. They characterized motives as goals (academic performance or mastery goals) and expectancies (self-efficacy and outcome expectations). Because of our interest in games, we also added the affective variable of play. We have also added the motivational constructs of effort and anxiety, which are shown in Figure 3. In addition, based on Robbins *et al.* and our own research, we have provided definitions of these motivational constructs. These are shown in Table 3 (adapted from Robbins *et al.*, 2004, p. 267). Other sources of research evidence for these motivational constructs can be found in Hidi and Harackiewicz (2000), Pekrun *et al.* (2002), Marsh *et al.* (2004), Seifert (2004) and Artelt (2005).

A closing comment: relationship of instructional design to effective games and learning outcomes

Our position is that games themselves are not sufficient for learning, but there are elements in games that can be activated within an instructional context that may enhance the learning process (Garris *et al.*, 2002). In other words, outcomes are affected by the instructional strategies employed (Wolfe, 1997). Leemkuil *et al.* (2003), too, commented that there is general consensus that learning with interactive environments such as games, simulations and adventures is not effective when no effective instructional measure or support is added.

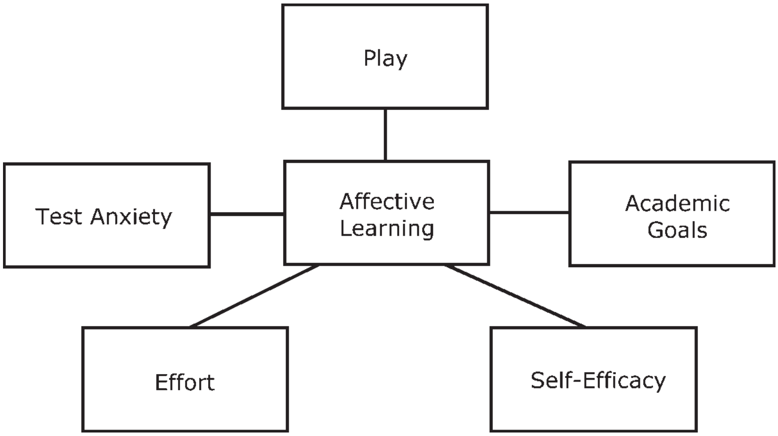

Figure 3. Affective/Motivation model of learning

Table 3. Motivation constructs and their definitions

Construct	*Definition*
Academic goals	One's persistence with and commitment to action, including general and specific goal-directed behaviour
Self-efficacy	Self-evaluation of one's ability and/or chances for success in the academic environment
Effort	O'Neil and Herl defined effort as the 'extent to which one works hard on a task' (1998, p. 1)
Play	Play is entertainment without fear of present or future consequences; it is fun (Resnick & Sherer, 1994)
Test anxiety	The test anxiety literature (Hembree, 1988) categorizes test anxiety as both worry and emotionality. Worry is the cognitive concern about performance (e.g. 'I wished I'd studied harder for this test'). Emotionality (or somatic anxiety) is a reaction to one's physiological responding (e.g. 'I feel my heart beating faster')

Source: Robbins *et al.* (2004), p. 267.

After reviewing a large number of studies on learning from simulations, de Jong and van Joolingen concluded, 'There is no clear and unequivocal outcome in favor of simulations. An explanation why simulation-based learning does not improve learning results can be found in the intrinsic problems that learners may have with discovery learning' (1998, p. 181). These problems are related to processes such as hypothesis generation, design of experiments, interpretation of data and regulation of learning. After analysing a large number of studies, de Jong and van Joolingen concluded that adding instructional support (i.e. scaffolding) to simulations might help to improve the situation. A similar conclusion can be applied to games. Gee (2003) asserted that 36 effective learning principles can be found in well-designed commercial video games.

Among those 36 principles, one states that learners will spend lots of time on task in a well-designed game, creating an environment that fosters practice. Another principle, the transfer principle, states that effective games provide learners with multiple opportunities to apply earlier learning to later problems, including problems that require adapting and transforming that earlier learning. A third principle, related to the concept of scaffolding, states that basic skills are not learned in isolation or out of context but, rather, as bottom-up learning by engaging more and more of those skills as the video game progresses (Gee, 2003).

According to Thiagarajan (1998), if not embedded with sound instructional design, games and simulations often end up truncated exercises frequently mislabelled as simulations. Gredler (1996) further commented that poorly developed exercises are not effective in achieving the objectives for which simulations are most appropriate—that of developing students' problem-solving skills. According to Lee (1999), for instructional prescription we need information dealing with instructional variables, such as instructional mode, instructional sequence, knowledge domain and learner characteristics. Further, Kirschner *et al.* (in press) provide a convincing argument that discovery, problem-based, experiential and enquiry-based techniques do not work instructionally. The basis of their argument is that such strategies require prior knowledge on the part of the student or trainee to be efficient and effective. Most individuals in an educational or training environment, however, are likely to have low prior knowledge. Since such techniques are the sole instructional strategy in the vast majority of games, it would be expected that most games would not lead to adult learning.

There appears to be consensus among a large number of researchers with regard to the negative, mixed or null findings of games research, suggesting that the cause might be a lack of sound instructional design embedded in the games (Gredler, 1996; Wolfe, 1997; de Jong & van Joolingen, 1998; Thiagarajan, 1998; Lee, 1999; Garris *et al.*, 2002; Leemkuil *et al.*, 2003; O'Neil & Fisher, 2004). However, as we embed instructional strategies in games, we must consider individual differences. In an examination of the role of training scenarios in video games, Oliver and Pelletier (2005) found that providing training in games can be effective for strategy development, but that players apply those strategies differentially, with some players more effective than others. These differences come not only from our knowledge, skills and abilities, they are socially motivated as well. According to Pelletier (2005), how we interact with games hinges on 'the network of social relations that always over-determine the way games assume meaning' (p. 57).

An important component of research on the effectiveness of educational games and simulations is the measurement and assessment of performance outcomes from the various instructional strategies embedded into the games or simulations that involve the learning outcome of problem solving. Problem solving is one of the cognitive demands in the CRESST model of learning. 'Problem solving is cognitive processing directed at achieving a goal when no solution method is obvious to the problem solver' (Mayer & Wittrock, 1996, p. 47). O'Neil's (1999) problem-solving model includes the following components: content understanding; problem-solving strategies—domain-independent (-general) and domain-dependent (-specific); and self-regulation, which comprises metacognition and motivation. Metacognition further comprises self-checking/monitoring and planning, and motivation comprises effort and self-efficacy.

Effective problem solving in games can place a large cognitive load on working memory. Cognitive load theory (Sweller, 1989, 1994) provides a theoretical framework for this assertion. This theory assumes that learning uses a very limited working (or short-term) memory and an unlimited long-term memory. If there are too many elements to be learned, then cognitive load will exceed the limits of working memory, and therefore there would be less learning. Thus, instructional strategies have been recommended to help reduce cognitive load, for example, scaffolding (Mayer *et al.*, 2002) and worked examples (Sweller, 1988, 2003; Sweller *et al.*, 1998). With respect to scaffolding, while there are a number of definitions (e.g. van Merrienboer *et al.*, 2002, 2003; Chalmers, 2003), what they all have in common is that scaffolding is an instructional method that provides support during learning by reducing cognitive load. Clark (2001) described instructional methods as external representations of the internal processes of selecting, organizing and integrating. These processes provide learning goals, monitoring procedures, feedback, selection methods, hints, prompts and various advance organizers (Jones *et al.*, 1995; Alessi, 2000; Clark, 2001; Leemkuil *et al.*, 2003). Each of these components either reflects a form of scaffolding or reflects a need for scaffolding.

One form of scaffolding is graphical scaffolding. A number of studies have reported the benefits of maps, a type of graphical scaffolding (Benbasat & Todd, 1993; Chou & Lin, 1998; Ruddle *et al.*, 1999; Chou *et al.*, 2000; Farrell & Moore, 2000). In virtual environments, navigation maps help the user to navigate, orient the user and facilitate an easier learning experience (Yair *et al.*, 2001). While navigation maps can reduce or distribute cognitive load (Cobb, 1997), they also have the potential to add load, ultimately counteracting their positive effects. Navigation maps can provide valuable cognitive support for navigating virtual environments, such as computer-based video games. This can be particularly useful when using the gaming environment to accomplish a complex problem-solving task. We currently have game research being conducted in our laboratory to investigate the use of both maps and worked examples to improve video game problem-solving performance.

Summary

In this article, we defined the terms *game* and *simulation* using the definitions of Gredler (1996) and reviewed the educational potential of games. The evidence of potential is striking, but the empirical evidence for effectiveness of games as learning environments is scant. The empirical research on computer-based games and training for adults was reviewed in the context of two theoretical frameworks: Kirkpatrick's four levels of evaluating training (Kirkpatrick, 1994) and the CRESST model of learning (Baker & Mayer, 1999). We believe that both frameworks are useful in the classification of learning outcomes. Finally we suggest the need for an augmented CRESST framework that includes affective learning.

Author note

Support for this article was provided by the Defense Advanced Research Projects Agency as administered by the Office of Naval Research under Award Number N00014-03-C-0357 and Award Number N00014-04-M-0088 to Advance Design Information. The findings and opinions expressed in this publication do not reflect the positions or policies of the Defense Advanced Research Projects Agency or the Office of Naval Research.

References

Adams, P. C. (1998) Teaching and learning with SimCity 2000 [electronic version], *Journal of Geography*, 97(2), 47–55.

Alessi, S. M. (2000) Simulation design for training and assessment, in: H. F. O'Neil, Jr & D. H. Andrews (Eds) *Aircrew training and assessment* (Mahwah, NJ, Lawrence Erlbaum Associates), 197–222.

Anderson, C. A. & Bushman, B. J. (2001) Effects of violent video games on aggressive behavior, aggressive cognition, aggressive affect, physiological arousal, and prosocial behavior: a meta-analytic review of the scientific literature, *Psychological Science*, 12, 353–359.

Artelt, C. (2005) Cross-cultural approaches to measuring motivation, *Educational Assessment*, 10, 231–255.

Arthur, W., Jr, Strong, M. H., Jordan, J. A., Williamson, J. E., Shebilske, W. L. & Regian, J. W. (1995) Visual attention: individual differences in training and predicting complex task performance, *Acta Psychologica*, 88, 3–23.

Baker, D., Prince, C., Shrestha, L., Oser, R. & Salas, E. (1993) Aviation computer games for crew resource management training, *The International Journal of Aviation Psychology*, 3, 143–156.

Baker, E. L., Aschbacher, P. & Bradley, C. (1985) *Evaluation of WEST: ICAI system* (report to Jet Propulsion Laboratory) (Los Angeles, CA, University of California, Center for the Study of Evaluation).

Baker, E. L. & Mayer, R. E. (1999) Computer-based assessment of problem solving, *Computers in Human Behavior*, 15, 269–282.

Baker, E. L. & O'Neil, H. F., Jr (2002) Measuring problem solving in computer environments: current and future states, *Computers in Human Behavior*, 18, 609–622.

BBN/DARWARS (2004) *Conference presentation folder provided at the 2004 Interservice/Industry Training, Simulation and Education Conference,* Orlando, FL, December.

Benbasat, I. & Todd, P. (1993) An experimental investigation of interface design alternatives: icon vs. text and direct manipulation vs. menus, *International Journal of Man-Machine Studies*, 38, 369–402.

Betz, J. A. (1995–6) Computer games: increase learning in an interactive multidisciplinary environment, *Journal of Educational Technology Systems*, 24, 195–205.

Brougere, G. (1999) Some elements relating to children's play and adult simulation/gaming, *Simulation & Gaming*, 30, 134–146.

Brozik, D. & Zapalska, A. (2002) The PORTFOLIO GAME: decision making in a dynamic environment, *Simulation & Gaming*, 33, 242–255.

Brunken, R., Plass, J. L. & Leutner, D. (2003) Direct measurement of cognitive load in multimedia learning, *Educational Psychologist*, 38(1), 53–61.

Carr, P. D. & Groves, G. (1998) The internet-based operations simulation game, in: J. A. Chambers (Ed.) *Selected papers for the 9th International Conference on College Teaching and Learning* (Jacksonville, Florida Community College at Jacksonville), 15–23.

Chalmers, P. A. (2003) The role of cognitive theory in human-computer interface, *Computers in Human Behavior*, 19, 593–607.

Chou, C. & Lin, H. (1998) The effect of navigation map types and cognitive styles on learners' performance in a computer-networked hypertext learning system [electronic version], *Journal of Educational Multimedia and Hypermedia*, 7, 151–176.

Chou, C., Lin, H. & Sun, C.-t. (2000) Navigation maps in hierarchical-structured hypertext courseware [electronic version], *International Journal of Instructional Media*, 27(2), 165–182.

Chuang, S. (2004) *The role of search strategies and feedback on a computer-based collaborative problem-solving task*. Unpublished doctoral dissertation, University of Southern California, Los Angeles.

Clark, R. E. (1999) The CANE model of motivation to learn and to work: a two-stage process of goal commitment and effort [electronic version], in: J. Lowyck (Ed.) *Trends in corporate training* (Leuven, University of Leuven Press).

Clark, R. E. (Ed.) (2001) *Learning from media: arguments, analysis, and evidence* (Greenwich, CT, Information Age Publishing).

Clark, R. E. (2003a) Fostering the work motivation of teams and individuals, *Performance Improvement*, 42(3), 21–29.

Clark, R. E. (2003b) Strategies based on effective feedback during learning, in: H. F. O'Neil (Ed.) *What works in distance education* (report to the Office of Naval Research by the National Center for Research on Evaluation, Standards, and Student Testing) (Los Angeles, University of Southern California, Rossier School of Education), February, 18–19.

Cobb, T. (1997) Cognitive efficiency: toward a revised theory of media, *Educational Technology Research and Development*, 45(4), 21–35.

Cross, T. L. (1993) AgVenture: a farming strategy computer game, *Journal of Natural Resources and Life Sciences Education*, 22, 103–107.

Csikszentmihalyi, M. (1975) *Beyond boredom and anxiety* (San Francisco, CA, Jossey Bass).

Csikszentmihalyi, M. (1990) *Flow: the psychology of optimal performance* (New York, Cambridge University Press).

Davis, S. & Wiedenbeck, S. (2001) The mediating effects of intrinsic motivation, ease of use and usefulness perceptions on performance in first-time and subsequent computer users, *Interacting with Computers*, 13, 549–580.

Day, E. A., Arthur, W., Jr & Gettman, D. (2001) Knowledge structures and the acquisition of a complex skill, *Journal of Applied Psychology*, 86, 1022–1033.

de Jong, T. & van Joolingen, W. R. (1998) Scientific discovery learning with computer simulations of conceptual domains, *Review of Educational Research*, 68, 179–202.

deCharms, R. (1986) *Personal causation* (New York, Academic Press).

Deci, E. L. (1975) *Intrinsic motivation* (New York, Plenum Press).

Dekkers, J. & Donatti, S. (1981) The interpretation of research studies on the use of simulation as an instructional strategy, *Journal of Educational Research*, 74(6), 64–79.

Donchin, E. (1989) The learning strategies project, *Acta Psychologica*, 71, 1–15.

Druckman, D. (1995) The educational effectiveness of interactive games, in: D. Crookall & K. Aria (Eds) *Simulation and gaming across disciplines and cultures: ISAGA at a watershed* (Thousand Oaks, CA, Sage), 178–187.

Dugdale, S. (1998) Mathematical problem-solving and computers: a study of learner-initiated application of technology in a general problem-solving context, *Journal of Research on Computing in Education*, 30, 239–253.

Farrell, I. H. & Moore, D. M. (2000) The effect of navigation tools on learners' achievement and attitude in a hypermedia environment, *Journal of Educational Technology Systems*, 29(2), 169–181.

Galimberti, C., Ignazi, S., Vercesi, P. & Riva, G. (2001) Communication and cooperation in networked environment: an experimental analysis, *CyberPsychology & Behavior*, 4(1), 131–146.

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

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

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

Gredler, M. E. (1996) Educational games and simulations: a technology in search of a research paradigm, in: D. H. Jonassen (Ed.) *Handbook of research for educational communications and technology* (New York, Simon & Schuster Macmillan), 521–540.

Green, C. S. & Bavelier, D. (2003) Action video game modifies visual selective attention, *Nature*, 423, 534–537.

Green, T. D. & Flowers, J. H. (2003) Comparison of implicit and explicit learning processes in a probabilistic task, *Perceptual and Motor Skills*, 97, 299–314.

Greenfield, P. M., deWinstanley, P., Kilpatrick, H. & Kaye, D. (1994) Action video games and informal education: effects on strategies for dividing visual attention, *Journal of Applied Developmental Psychology*, 15, 105–123.

Harter, S. (1978) Effectance motivation reconsidered: toward a developmental model, *Human Development*, 1, 34–64.

Hembree, R. (1988) Correlates, causes, effects, and treatment of test anxiety, *Review of Educational Research*, 58(1), 47–77.

Herl, H. E., O'Neil, H. F., Jr, Chung, G. & Schacter, J. (1999) Reliability and validity of a computer-based knowledge mapping system to measure content understanding, *Computers in Human Behavior*, 15, 315–333.

Hidi, S. & Harackiewicz, J. M. (2000) Motivating the academically unmotivated: a critical issue for the 21st century, *Review of Educational Research*, 70, 151–179.

Hsieh, I. (2001) *Types of feedback in a computer-based collaborative problem-solving Group Task*. Unpublished doctoral dissertation, University of Southern California, Los Angeles.

Hubbard, P. (1991) Evaluating computer games for language learning, *Simulation & Gaming*, 22, 220–223.

Jones, M. G., Farquhar, J. D. & Surry, D. W. (1995) Using metacognitive theories to design user interfaces for computer-based learning, *Educational Technology*, 35(4), 12–22.

Kalyuga, S., Chandler, P., Touvinen, J. & Sweller, J. (2001) When problem-solving is superior to worked examples, *Journal of Educational Psychology*, 93, 579–588.

Katz, S. & Lesgold, A. (1991) Modeling the student in SHERLOCK II, in: J. Kay & A. Quilici (Eds) Proceedings of the IJCAI-91 Workshop W.4: Agent modelling for intelligent interaction (Sydney, Australia), 93–127.

Khoo, G.-s. & Koh, T.-s. (1998) Using visualization and simulation tools in tertiary science education [electronic version], *The Journal of Computers in Mathematics and Science Teaching*, 17(1), 5–20.

King, K. W. & Morrison, M. (1998) A media buying simulation game using the internet, *Journalism and Mass Communication Educator*, 53(3), 28–36.

Kirkpatrick, D. L. (1994) *Evaluating training programs: the four levels* (San Francisco, CA, Berrett-Koehler).

Kirschner, P. A., Sweller, J. & Clark, R. E. (in press) Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching, *Educational Psychologist*.

Lee, J. (1999) Effectiveness of computer-based instructional simulation: a meta analysis, *International Journal of Instructional Media*, 26(1), 71–85.

Leemkuil, H., de Jong, T., de Hoog, R. & Christoph, N. (2003) KM Quest: a collaborative internet-based simulation game, *Simulation & Gaming*, 34, 89–111.

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

Malone, T. W. & Lepper, M. R. (1987) *Aptitude, learning and instruction III: cognitive and affective process analysis* (Hillsdale, NJ, Lawrence Erlbaum Associates).

Marsh, H. W., Hau, K-T., Artelt, C. & Baumert, J. (2004) *OECD's brief self-report measures of educational psychology's most useful constructs: cross-cultural, psychometric comparisons across 25 countries*. Unpublished manuscript, SELF Research Centre, University of West Sydney, Penrith South DC, New South Wales, Australia.

Mayer, R. E. (2004) Should there be a three-strikes rule against pure discovery learning?, *American Psychologist*, 59, 14–19.

Mayer, R. E., Dow, G. T. & Mayer, S. (2003) Multimedia learning in an interactive self-explaining environment: what works in the design of agent-based microworlds?, *Journal of Educational Psychology*, 95, 806–813.

Mayer, R. E., Mautone, P. & Prothero, W. (2002) Pictorial aids for learning by doing in a multimedia geology simulation game, *Journal of Educational Psychology*, 94, 171–185.

Mayer, R. E. & Moreno, R. (1998) A split-attention effect in multimedia learning: evidence for dual processing systems in working memory, *Journal of Educational Psychology*, 90, 312–320.

Mayer, R. E., Sobko, K. & Mautone, P. D. (2003) Social cues in multimedia learning: role of speaker's voice, *Journal of Educational Psychology*, 95, 419–425.

Mayer, R. E. & Wittrock, M. C. (1996) Problem-solving transfer, in: D. C. Berliner & R. C. Calfee (Eds) *Handbook of educational psychology* (New York, Simon & Schuster Macmillan), 47–62.

Mislevy, R. J. (in press) Substance and structure in assessment arguments, *Law, probability, and risk.* Available online at: http://www.education.umd.edu/EDMS/mislevy/papers/Argument.pdf

Mislevy, R. J., Steinberg, L. S. & Almond, R. G. (2003) On the structure of educational assessments, *Measurement: Interdisciplinary Research and Perspectives*, 1, 3–67.

Mislevy, R. J., Wilson, M. R., Ercikan, K. & Chudowsky, N. (2003) Psychometric principles in student assessment, in: T. Kellaghan & D. Stufflebeam (Eds) *International handbook of educational evaluation* (Dordrecht, Netherlands, Kluwer Academic Press), 489–531.

Moreno, R. & Mayer, R. E. (2000) Engaging students in active learning: the case for personalized multimedia messages, *Journal of Educational Psychology*, 92, 724–733.

Morris, C. S., Hancock, P. A. & Shirkey, E. C. (2004) Motivational effects of adding context relevant stress in PC-based game training, *Military Psychology*, 16, 135–147.

Oliver, M. & Pelletier, C. (2005) The things we learned on Liberty Island: designing games to help people become competent game players, proceedings of *DIGRA 2005 Conference: Changing Views–Worlds in Play.*

O'Neil, H. F., Jr (1999) Perspectives on computer-based performance assessment of problem solving: editor's introduction, *Computers in Human Behavior*, 15, 255–268.

O'Neil, H. F., Jr (2002) Perspective on computer-based assessment of problem solving [special issue], *Computers in Human Behavior*, 18, 605–607.

O'Neil, H. F., Jr, Baker, E. L. & Fisher, J. Y.-C. (2002) *A formative evaluation of ICT games* (report to the Institute for Creative Technologies, University of Southern California) (Los Angeles, CA, University of Southern California, Rossier School of Education).

O'Neil, H. F., Jr & Fisher, Y.-C. (2004) A technology to support leader development: computer games, in: D. V. Day, S. J. Zaccaro & S. M. Halpin (Eds) *Leader development for transforming organizations* (Mahwah, NJ, Lawrence Erlbaum Associates), 99–121.

O'Neil, H. F., Jr & Herl, H. E. (1998) Reliability and validity of a trait measure of self-regulation, paper presented at the annual meeting of the *American Educational Research Association,* San Diego, CA, April.

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

Pekrun, R., Goetz, T., Titz, W. & Perry, R. P. (2002) Academic emotions in students' self-regulated learning and achievement: a program of qualitative and quantitative research, *Educational Psychologist*, 37, 91–105.

Pelletier, C. (2005) The uses of literacy in studying computer games: comparing students' oral and visual representations of games, *English Teaching: Practice and Critique*, 4(1), 40–59.

Porter, D. B., Bird, M. E. & Wunder, A. (1990–1) Competition, cooperation, satisfaction, and the performance of complex tasks among Air Force cadets, *Current Psychology: Research & Reviews*, 9, 347–354.

Prislin, R., Jordan, J. A., Worchel, S., Semmer, F. T. & Shebilske, W. L. (1996) Effects of group discussion on acquisition of complex skills, *Human Factors*, 38, 404–416.

Quinn, C. N. (1991) Computers for cognitive research: a HyperCard adventure game, *Behavior Research Methods, Instruments, & Computers*, 23, 237–246.

Ramsberger, P. F., Hopwood, D., Hargan, C. S. & Underhill, W. G. (1983) *Evaluation of a spatial data management system for basic skills education. Final Phase I Report for Period 7 October 1980–30 April 1983* (HumRRO FR-PRD-83-23) (Alexandria, VA, Human Resources Research Organization).

Resnick, H. & Sherer, M. (1994) Computerized games in the human services–an introduction, in: H. Resnick (Ed.) *Electronic tools for social work practice and education* (Binghamton, NY, Haworth Press), 5–16.

Rhodenizer, L., Bowers, C. A. & Bergondy, M. (1998) Team practice schedules: what do we know?, *Perceptual and Motor Skills*, 87, 31–34.

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

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(2), 43–58.

Rieber, L. P. & Parmley, M. W. (1995) To teach or not to teach? Comparing the use of computer-based simulations in deductive versus inductive approaches to learning with adults in science, *Journal of Educational Computing Research*, 13, 359–374.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R. & Carlstrom (2004) Do psychosocial and study skill factors predict college outcomes? A meta-analysis, *Psychological Bulletin*, 130, 261–288.

Rosenorn, T. & Kofoed, L. B. (1998) Reflection in learning processes through simulation/gaming, *Simulation & Gaming*, 29, 432–440.

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

Ruddle, R. A., Payne, S. J. & Jones, D. M. (1999) The effects of maps on navigation and search strategies in very-large-scale virtual environments, *Journal of Experimental Psychology: Applied*, 5(1), 54–75.

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

Salomon, G. (1983) The differential investment of mental effort in learning from different sources, *Educational Psychology*, 18, 42–50.

Santos, J. (2002) Developing and implementing an internet-based financial system simulation game, *The Journal of Economic Education*, 33(1), 31–40.

Schacter, J., Herl, H. E., Chung, G., Dennis, R. A. & O'Neil, H. F., Jr (1999) Computer-based performance assessments: a solution to the narrow measurement and reporting of problem-solving, *Computers in Human Behavior*, 13, 403–418.

Schank, R. C. (1997) *Virtual learning: a revolutionary approach to build a highly skilled workforce* (New York, McGraw-Hill).

Seifert, T. L. (2004) Understanding student motivation, *Educational Research*, 46, 137–149.

Shebilske, W. L., Regian, W., Arthur, W., Jr & Jordan, J. A. (1992) A dyadic protocol for training complex skills, *Human Factors*, 34, 369–374.

Shewokis, P. A. (2004) Memory consolidation and contextual interference effects with computer games, *Perceptual and Motor Skills*, 97, 381–389.

Spiker, V. A. & Nullmeyer, R. T. (n.d.) *Benefits and limitations of simulation-based mission planning and rehearsal.* Unpublished manuscript.

Stolk, D., Alexandrian, D., Gros, B. & Paggio, R. (2001) Gaming and multimedia applications for environmental crisis management training, *Computers in Human Behavior*, 17, 627–642.

Sugrue, B. & Kim, K.-H. (2004) *State of the industry. ASTD's annual review of trends in workplace learning and performance. Executive summary* (Alexandria, VA, American Society for Training and Development).

Sweller, J. (1988) Cognitive load during problem solving: effects on learning, *Cognitive Science*, 12, 257–285.

Sweller, J. (1989) Cognitive technology: some procedures for facilitating learning and problem solving in mathematics and science, *Journal of Educational Psychology*, 81, 457–466.

Sweller, J. (1994) Cognitive load theory, learning difficulty and instructional design, *Learning and Instruction*, 4, 295–312.

Sweller, J. (2003) Evolution of human cognitive architecture, in: B. Ross (Ed.) *The psychology of learning and motivation*, (vol. 43), (San Diego, CA, Academic Press), 215–266.

Sweller, J., van Merrienboer, J. & Paas, F. (1998) Cognitive architecture and instructional design, *Educational Psychology Review*, 10, 251–269.

Thiagarajan, S. (1998) The myths and realities of simulations in performance technology, *Educational Technology*, 38(4), 35–41.

Thomas, P. & Macredie, R. (1994) Games and the design of human-computer interfaces, *Educational Technology*, 31(2), 134–142.

Tkacz, S. (1998) Learning map interpretation: skill acquisition and underlying abilities, *Journal of Environmental Psychology*, 18, 237–249.

Toulmin, S. E. (1958) *The uses of argument* (Cambridge, Cambridge University Press).

van Merrienboer, J. J. G., Clark, R. E. & de Croock, M. B. M. (2002) Blueprints for complex learning: the 4C/ID-model, *Educational Technology Research & Development*, 50(2), 39–64.

van Merrienboer, J. J. G., Kirschner, P. A. & Kester, L. (2003) Taking a load off a learner's mind: instructional design for complex learning, *Educational Psychologist*, 38(1), 5–13.

Wainess, R. & O'Neil, H. F., Jr (2003) *Feasibility study*. Unpublished manuscript, University of Southern California, Rossier School of Education.

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

Westbrook, J. I. & Braithwaite, J. (2001) The Health Care Game: an evaluation of a heuristic, web-based simulation, *Journal of Interactive Learning Research*, 12(1), 89–104.

White, B. Y. & Frederiksen, J. R. (1998) Inquiry, modeling, and metacognition: making science accessible to all students, *Cognition and Instruction*, 16, 3–118.

Wolfe, J. (1997) The effectiveness of business games in strategic management course work [electronic version], *Simulation & Gaming Special Issue: Teaching Strategic Management*, 28, 360–376.

Wolfe, J. & Roge, J. N. (1997) Computerized general management games as strategic management learning environments [electronic version], *Simulation & Gaming Special Issue: Teaching Strategic Management*, 28, 423–441.

Yair, Y., Mintz, R. & Litvak, S. (2001) 3D-virtual reality in science education: an implication for astronomy teaching, *Journal of Computers in Mathematics and Science Teaching*, 20, 293–305.

Yu, F.-Y. (2001) Competition within computer-assisted cooperative learning environments: cognitive, affective, and social outcomes, *Journal of Educational Computing Research*, 24, 99–117.