Assessing Problem-Solving Skills in Game-Based Immersive Environments

44

Valerie J. Shute and Benjamin Emihovich

Contents

Abstract

According to Pearson's Law: "That which is measured improves." But some important constructs, like problem-solving skill, are generally viewed as very difficult to measure for a variety of reasons (e.g., lack of a clear and agreed-upon definition, psychological and/or statistical multidimensionality of the construct, subjectivity of scoring, and so on). Adding to the challenge of validly measuring problem-solving skill is that formal education settings tend to focus only on well-structured problems – those that have correct and incorrect answers. However, these problems tend to have little transfer to the real world. In short, there is a gap in the kinds of problems being assessed and taught in schools and those desired in workplace environments. In this chapter, we focus on how game-based immersive environments, such as well-designed video games, can be used to both measure and promote the development of problem-solving skills in formal education settings. In this chapter, we discuss the theoretical foundations of problem-solving

V. J. Shute (✉) · B. Emihovich
Florida State University, Tallahassee, FL, USA
e-mail: vshute@admin.fsu.edu

© Springer International Publishing AG, part of Springer Nature 2018

J. Voogt et al. (eds.), *Second Handbook of Information Technology in Primary and Secondary Education*, Springer International Handbooks of Education,
https://doi.org/10.1007/978-3-319-71054-9_40

and provide a worked example of assessing it from gameplay using a well-designed video game. We conclude the chapter with a discussion on future implications for using game-based environments to measure and promote problem-solving skills in education.

Keywords
Assessment · Game-based learning · Problem-solving skill · Stealth assessment · Video games

Introduction to the Problem

Over the past couple of decades, developed and developing countries are increasingly relying on a knowledge economy to help solve complex problems arising from globalization such as issues related to global warming including the depletion of the ozone layer and the scarcity of natural resources in the environment. Solving complex problems requires one to think critically, creatively, and systemically when planning solutions, thus involves cognitive processes aimed at achieving a goal when the solution is hidden from the problem solver (Mayer and Wittrock 2006). Educators consider the assessment and support of problem-solving skills a primary goal of teaching and learning (e.g., Jonassen 1997; Shute and Wang 2016).

Problem-solving is also becoming increasingly desirable in twenty-first century workplace environments (Partnership for 21st Century Learning 2016). One of the challenges in teaching problem-solving skills in formal education settings stems from limitations with providing authentic problem-solving contexts. That is, in formal learning settings, students are typically required to solve well-structured problems that have limited transfer after instruction (Jonassen 2000). A well-structured problem is one that can be solved with a particular sequence of steps that leads to a correct answer. Most classroom textbook problems – particularly in STEM areas – feature well-structured problems that have correct answers. However, an ill-structured problem possesses a claim and a justifying argument without a specific correct answer. An example of an ill-structured problem would be to figure out a viable solution to a problem like what can be done when the population of a community is growing but the current water supply will not support many new people.

Instruction designed on specific principles of learning related to solving a variety of problems – from well-structured to ill-structured – in an authentic context is important for initial and long-term learning as well as promoting problem-solving transfer to novel contexts (Van Eck et al. 2017). But again, there is a gap in the kinds of problems being taught in schools with those being required in the workplace. Recent survey data indicates that only 24% of employers who hired recent college graduates stated that the graduates were able to analyze and solve complex problems in their workplace environments (Hart Research Associates 2015).

Facilitating learning by using immersive environments, such as well-designed video games in educational settings, is one promising way to bridge the aforementioned gap between school-based versus workplace or real-world problem-solving skills. Immersion in this context refers to the subjective impression one experiences when interacting with a realistic, digitally enhanced environment like a video game (Lessiter et al. 2001). This feeling of immersion enhances learning by allowing multiple perspectives, offering players a situated learning experience and promoting transfer (Dede 2009). In general, immersive environments support the following: distribution of knowledge across a community through situated learning, acquisition of fluency in multiple media, and interactions with different human or computer characters with different skills and abilities within authentic and novel problem-solving scenarios (Dede 2005).

Multiple perspectives in immersive environments allow for egocentric and exocentric frames of reference. Egocentric frames of reference engender motivation through embodied learning, whereas exocentric frames of reference promote abstract symbolic insights that are gained at a distance from the context of the environment (Dede 2009). In addition, immersive learning environments help learners practice as well as understand how to apply facts, principles, and concepts in context rather than abstractly. Learning abstractly via memorizing certain facts as well as abstract principles and concepts is how knowledge is typically taught in school (Barab and Dede 2007). But there is an alternative (and considerably more engaging) pedagogical approach to support learning. According to the recent report entitled "Essential Facts About The Computer and Video Game Industry" published by the Entertainment Software Association (ESA 2016), video games are played by 65% of US households and consumers spent over $23 billion dollars in the video game industry in 2015. Given the success of the video game industry, education researchers are exploring how well-designed video games can be used to support learning and assessment, referred to as game-based learning (GBL).

Over the past decade, researchers have been examining and testing various ways to deeply embed valid assessment for learning directly into video games with a technology called stealth assessment (see Shute and Ventura 2013). Establishing validity is key; herein referring to the psychometric quality of the assessment relative to its accuracy in measuring what is intended to be measured (also known as construct or convergent validity). Stealth assessment is grounded by an assessment design framework called evidence-centered design (ECD; Mislevy et al. 2003). In general, the main purpose of any assessment is to collect information that will allow the assessor to make valid inferences about what people know, can do, and to what degree (collectively referred to as "competencies" in this chapter). ECD defines a framework that consists of several conceptual and computational models that work in concert. The framework requires an assessor to: (a) define the claims to be made about learners' competencies, (b) establish what constitutes valid evidence of a claim, and (c) determine the nature and form of tasks that will elicit that evidence.

In this chapter, we first review the literature on problem-solving skills. Next, we discuss current methods for assessing problem-solving skills and describe an approach to address some of the issues by presenting an example of stealth assessment in a research project. We then address the benefits and challenges of using video games in formal education settings. Finally, we conclude with ideas for next steps in this field of research.

Problem-Solving Skills

Early research on problem-solving included Thorndike's (1898) experiments on cats tasked with escaping a puzzle box. Thorndike determined that the cats were able to escape by mere associations and habits, or essentially by trial and error. The seminal work of Thorndike's research is regarded as significant in contributing to early behavioral learning theory in psychology, but many scholars regarded problem-solving as an unexplainable human behavior before Newell et al.'s (1958) theoretical work on the mechanisms of thinking and learning related to problem-solving (Simon and Newell 1971). The gap in research on problem-solving directed scholars to correct the antiquated narrative on the topic from unexplainable thinking to analyzing how human beings behave and what cognitive processes engage this critical aspect of thinking and learning.

In general, problem-solving can be defined as cognitive processing aimed at finding a solution to a goal when the solution is not known to the problem solver (Mayer 1992). According to Mayer and Wittrock's (2006) definition, problem-solving consists of four interrelated elements: (1) occurs within the problem solver's cognitive system; (2) requires the problem solver to theorize and manipulate information; (3) is goal oriented; and (4) relies upon the intellect and skills of the problem solver to establish the order in which each obstacle is handled before finding a solution.

While the elements of problem-solving seem straightforward, can problem-solving skills be fostered and improved? Polya (1945) argued that problem-solving is not an innate skill, but rather something that can, in fact, be developed, "Solving problems is a practical skill, let us say, like swimming. . . Trying to solve problems, you have to observe and imitate what other people do when solving problems; and, finally, you learn to solve problems by doing them." (p. 5). So students are not born with problem-solving skills. Instead, these skills are cultivated when students have opportunities to solve problems proportionate to their knowledge. Polya (1945) went on to identify a four-step problem-solving technique: understand the problem, devise a plan, carry out the plan, and review/extend. Later, Bransford and Stein (1984) integrated the collection of problem-solving research and proposed the IDEAL problem solver. Each letter of IDEAL stands for an important part of the problem-solving process: identify problems and opportunities; define alternative goals; explore possible strategies; anticipate outcomes and act on strategies; and then look back and learn. Two years later, Gick (1986) presented a simplified model of the problem-solving process including the construction of a representation, the search for a solution, the implementation of the solution, and monitoring the outcome of the solution. Around that time, the focus of research had been shifted toward a knowledge-based representation such as schemata.

How domain-general versus domain-specific are problem-solving skills? Unlike the aforementioned work on identifying general problem-solving models, the study of problem representation recognizes the importance of domain-specific knowledge in solving problems. The representation of problems refers to how a problem solver perceives and understands a phenomenon within the problem-solving environment. Jonassen concluded that "problem solving is not a uniform activity" (2000, p. 65). Instead, problems vary in terms of their structuredness, complexity, and the requirement for prior knowledge (Jonassen 2003). For example, the types of problems associated with driving a truck are quite different from those needed when negotiating a business deal. Thus, how well a person can solve a problem depends on the accuracy of the person's representation of the problem scenario.

Jonassen (1997) noted that people confront various types of problems that differ according to the problem scenario, the cognitive processes required to find a solution, and the problem structure. As mentioned earlier, the structuredness of a problem can range from very well-structured (i.e., facts, rules, and principles) to very ill-structured (i.e., poorly defined, inconclusive, or conflicting solution paths). Providing students with repeated practice on solving various ill-structured problems can be beneficial in formal education settings because students of all ages experience ill-structured problems that arise naturally from daily interactions with other people and the environment. Trying to solve these types of problems requires students to define (and sometimes redefine) the problem and develop a solution pathway based on the skills needed to reach the solution (Jonassen 2002).

Recent research indicates problem-solving skills involve two facets: rule identification and rule application (Schweizer et al. 2013). "Rules" in problem-solving refers to the principles that govern the procedures, the conduct, or the actions in a problem-solving context. Rule identification refers to the ability to acquire knowledge of the problem-solving environment, and rule application is the ability to control the environment by applying the knowledge acquired. Typically, it is difficult to directly collect data on students' rule identification ability. However, since rule application is the outward expression of one's rule identification, the measurement of rule application may be used to reflect students' ability to identify rules.

Complex problems usually combine a mix of basic rules with rules that require cognitive flexibility – the ability to adjust prior thoughts or beliefs and explore alternative strategies in response to changes in the environment (Miyake et al. 2000). Cognitive flexibility is the opposite of functional fixedness, which is the difficulty a person experiences when he or she is required to use objects (or strategies) in uncommon ways (Duncker 1945). Many researchers have targeted functional fixedness as the major obstacle to successful problem-solving (e.g., Anderson 1980). Moreover, problem-solving skills play an important role in everyday life and in many professions. Gagné (1980) believed that the central point of education is to teach people to become better problem solvers, and as noted earlier, problem-solving is recognized as one of the most important skills demanded in the job market (Partnership for 21st Century Skills 2016). Therefore, research on problem-solving skills and how to improve them are essential to prepare students for upcoming challenges in life and work. In the next two sections, we discuss challenges associated with assessing and supporting problem-solving skills in education followed by a concrete example of developing a stealth assessment for problem-solving skills embedded within a well-designed game.

Challenges with Assessing Problem-Solving Skills

Key measurement challenges for complex constructs like problem-solving skills include (a) lack of a clear and consensual definition and/or operationalization of the construct, (b) theoretical multidimensionality of the construct where certain dimensions may have internal as well as external sources, (c) difficulty disambiguating the generality of the construct (e.g., is there a single "problem-solving" variable or is it solely dependent on the context), and (d) reliance on outdated multiple-choice and self-report measures, the former narrowly focused and the latter flawed. Self-report measures in particular are problematic as they are subject to "social desirability effects" that can lead to false reports about behaviors, attitudes, and beliefs (see Paulhus 1991). In addition, test takers may interpret specific self-report items differently (e.g., what it means to enjoy solving "hard problems") leading to unreliability and lower validity (Lanyon and Goodstein 1997).

Innovative and valid performance-based assessments of problem-solving skills are needed that assess how students apply this skill in the solution of complex, real-world problems. One way to approach this problem is to use video games or similarly immersive environments, to simulate a variety of problems for performance-based assessment (DiCerbo and Behrens 2012; Gobert et al. 2013; Quellmalz et al. 2012).

Early approaches to performance-based assessment in GBL included using serious games as tools to improve skills as seen in *Marine DOOM*, a modified version of the game *DOOM*, designed to improve military thinking and decision-making skills (Krulak 1997). New developments in GBL include the use of virtual reality (VR) technologies to promote collaboration and deep learning across a range of educational areas and for a broad swath of students. And while some researchers have attempted to design specific games to promote problem-solving skill (e.g., Van Eck et al. 2009), there are various commercial video games on the market that may be used for this research, such as *Portal 2* (Shute et al. 2015a) and *Plants vs. Zombies 2* (Shute et al. 2016). Such game-based environments can provide meaningful assessment by supplying students with scenarios that require the application of different facets of problem-solving skill. Next we describe a way to measure problem-solving skills continually and unobtrusively.

Stealth Assessment of Problem-Solving Skills

Stealth assessment (Shute 2011) provides a way to embed valid assessments directly into immersive environments like video games, extending the evidence-centered assessment design framework by delineating specific gameplay behaviors (specified in the evidence model) and statistically linking them to competency model variables (Shute and Ventura 2013). Stealth assessment thus complements the use of well-designed games to measure and improve skills by eliciting and aggregating student data during gameplay that can be used to provide real-time inferences about knowledge and skill states that can inform policies relevant to a range of stakeholders (Cukier and Mayer-Schoenberger 2013).

The results of this analysis are data (e.g., scores) that are passed to the competency model, which statistically updates the claims about relevant competencies in the student model. The ECD approach combined with stealth assessment provides a framework for developing assessment tasks that are explicitly linked to claims about personal competencies via an evidentiary chain (i.e., valid arguments that serve to connect task performance to competency estimates) and are thus valid for their intended purposes. The estimates of competency levels can also be used diagnostically and formatively to provide adaptively selected levels, feedback, and other forms of learning support to students as they continue to engage in gameplay.

To illustrate, Shute et al. (2016) developed a stealth assessment of problem-solving skill and embedded it in a game called Use Your Brainz (a slightly modified version of Plants vs. Zombies 2). The game was then tested with a sample of middle-school students. The research team began by developing a problem-solving competency model based on a review of the relevant literature and also reviewed the Common Core State Standards (CCSS) related to problem-solving. They came up with a four-facet competency model, which included: (a) understanding givens and constraints, (b) planning a solution pathway, (c) using tools effectively/efficiently when implementing solutions, and (d) monitoring and evaluating progress. The first facet maps to “rule acquisition,” and the remaining facets map to “rule application.”

After playing the game repeatedly and watching expert solutions on YouTube, the researchers delineated 32 observable indicators that were associated with the four facets (i.e., listed the specific actions a person would do while engaged in gameplay that would provide evidence – positive and negative – towards one or more of the facets). For example, sunflowers produce sun power, which is the sole source of power that players may use to grow plants. At the beginning of a level, typically there are no sunflowers on the battlefield. To supply power to grow plants, players must plant sunflowers at the beginning of each level before zombies start to appear in waves. The scoring rule for this particular indicator was: “If a player plants more than three sunflowers before the second wave of zombies arrives, the student understands relevant time and resource constraints of the level.” Table 1 displays a sample of indicators for each of the four problem-solving facets.

Table 1 Examples of indicators for each problem-solving facet. (From Shute et al. 2017)

Facet	Example indicators
Analyzing givens & constraints	Plants >3 Sunflowers before the second wave of zombies arrives Selects plants off the conveyor belt before it becomes full
Planning a solution pathway	Places sun producers in the back/left, offensive plants in the middle, and defensive plants up front/right Plants Twin Sunflowers or uses plant food on (Twin) Sunflowers in levels that require the production of X amount of sun
Using tools effectively and efficiently	Uses plant food when there are >5 zombies in the yard or zombies are getting close to the house (within 2 squares) Damages >3 zombies when firing a Coconut Cannon
Monitoring and evaluating progress	Shovels Sunflowers in the back and replaces them with offensive plants when the ratio of zombies to plants exceeds 2:1

Next, the researchers created a Q-matrix (Almond 2010) laying out all of the indicators in rows and the four facets in columns. A "1" was included in the crossed cell if the indicator was relevant to the facet and "0" if the facet did not apply to the indicator. They then went through each indicator and discussed how to classify it into discrete scoring categories such as "yes/no" or "very good/good/ok/poor." The overall scoring rules were based on a tally of relevant instances of observables. Using the aforementioned sunflower indicator, if a player successfully planted more than three sunflowers before the second wave of zombies arrived on the scene, the log file would automatically record the action and categorize it as a "yes" status of the indicator.

Once all of the indicators were categorized into various states, statistical relationships were established between each indicator and the associated levels of the competency model variables. The researchers used Bayesian networks (BNs) to accumulate incoming data from gameplay and update beliefs in the competency model. A BN is a probabilistic graphical model that represents a set of random variables and their conditional dependencies via a directed acyclic graph. The relationship between each indicator and its associated competency model variable was thus expressed within conditional probability tables stored in each BN. There were a total of 43 BNs developed for this project, one for each level. Each of the 43 levels in the game had its own BN, reflecting the specific indicators of the level, as well as its particular difficulty and discrimination estimates. The statistical relationships carried in the BNs and the scoring rules described earlier formed the evidence model. For specific examples of scoring rules and BNs, see Shute et al. (2017).

To validate the stealth assessment, data were collected from middle school students who played the game-based assessment for 3 h (i.e., 1 h per day across 3 consecutive days) and then completed two external problem-solving measures (i.e., Raven's Progressive Matrices and MicroDYN). Results indicated that the game-based stealth assessment estimates of problem-solving skill were significantly correlated with both Raven's ($r = 0.40$, $p < 0.01$) and MicroDYN ($r = 0.41$; $p < 0.01$) which established the construct validity of the stealth assessment.

Next steps include running a larger validation study and developing tools to help educators interpret the results of the assessment, which will subsequently support the development of problem-solving skills. In the next section, we explore the benefits and challenges associated with using video games in a classroom environment.

Video Games in Formal Education Settings

Given the gap in problem-solving skills being taught in schools and those required in workplace environments, we suggested the integration of video games into formal learning curricula with the goal of improving problem-solving skills, and to prepare students with the skills necessary to succeed in a globally interconnected society. Similarly, Schrader et al. (2006) argued that gameplay experiences within video games can promote collaboration and problem-solving skills for students through immersive game environments. However, while many students and teachers clearly enjoy video gameplay, the majority do not fully understand how video games can be used to support learning outcomes (Selfe and Hawisher 2004).

Well-designed video games can support higher-level learning via offering players practice across multiple and varied problem-solving scenarios (Gee 2007, 2008; Van Eck 2006). Repeated practice with novel problems can improve problem-solving skills within a domain (Gagné et al. 2005). In K-12 settings, video games have been shown to improve creativity, recall of facts, and problem-solving skills when compared to computer-assisted instruction (Chuang and Chen 2009; Jackson et al. 2012). Overall, there seems to be consensus that video games can support changes in affect, behavior, knowledge, and skill acquisition for learners (Clark et al. 2016; Ifenthaler et al. 2012; Shute et al. 2015b; Sitzmann 2011; Wouters et al. 2013).

While well-designed video games can help improve cognitive competencies, there is also a need to bridge the gap between in-game knowledge and school knowledge (i.e., game-based learning and formal learning). One possibility involves the design of scaffolds that are both engaging and meaningful to instruction, such as presenting concepts that are directly tied to in-game performance and readily identifiable during gameplay (Barzilai and Blau 2014). As an example, avatars in *World of Warcraft* (i.e., the digital embodiment of a player in the game) communicate with each other to solve problems such as how to use given tools and resources to travel from one zone to another. Using tools and resources efficiently is a facet of problem-solving skill (Shute and Wang 2016). Introducing a problem-based scenario in *World of Warcraft* that requires students to learn how to reduce travel time from one zone to the next using resources (e.g., map and transportation hub) efficiently is one way to teach problem-solving skills in formal education settings.

In addition, reducing travel time is a relevant concept that can transfer from the game to the natural world. That is, designing scaffolding can challenge students to think critically during gameplay and provide them with meaningful connections between gameplay and educational concepts (Gee and Hayes 2011). Instruction that weaves gameplay concepts with educational content by using scaffolds can inform stakeholders and policy makers that video games are effective learning tools in formal education settings. Well-designed multiplayer video games like *World of Warcraft* also provide opportunities for collaborative learning and problem-solving that has relevant implications in workplace environments.

Certain well-designed video games that emphasize multiplayer interactivity during gameplay can also help players develop skills that are desirable in twenty-first century workplace environments. As an example, Symantec's former Chief Operating Officer (COO), Stephen Gillet, included *World of Warcraft* as one of his achievements on his résumé to highlight his experience in organizing large groups of individuals to solve difficult problems, demonstrate confidence in solving these problems through natural and immersive video game environments (i.e., analyzing different strategies and gameplay tactics with peers in-game and through social media platforms like Twitter), and distribute resources (i.e., gold, crafting components, weapons, and armor) to in-game friends (Pagliery 2014). These gameplay practices may lead to improvement in planning and executing strategies that can be applied to similar workforce-related collaborative projects in the natural world. We now conclude with our thoughts on future directions of game-based learning and assessment in academic research and practice.

Conclusions and Future Implications for Game-Based Assessment

The demand for skilled workers who can solve complex problems in twenty-first century workplace environments requires a shift in how to measure and support problem-solving skills in formal education settings. In this chapter, we defined problem-solving skills and explored the theoretical foundations for GBL to explain how interactions within immersive environments can promote the development of problem-solving skills. We outlined some of the challenges associated with assessing problem-solving skills in formal education settings (e.g., the prevalence of flawed assessment methods such as multiple-choice type of items and self-report surveys) and highlighted the importance of using performance-based assessments in education. This suggestion coincides with the growing labor demand for multi-skilled workers in a constantly changing global economic landscape.

In addition, we proposed a possible solution to reduce the gap between school-based and real-world problem-solving skills via integrating immersive environments like video games into formal education settings to help students improve their problem-solving skills by preparing them with repeated problem-solving scenarios through immersive gameplay. To illustrate, we included an example of a problem-solving stealth assessment (see Shute et al. 2016). Furthermore, recent research on the effects of immersive gameplay on cognitive and noncognitive skills demonstrates that learning does occur – for problem-solving skills, persistence, visual-spatial skills, and attention (e.g., Green and Bavelier 2012; Rowe et al. 2011; Shute et al. 2015a; Ventura et al. 2013). Some researchers, however, have reported null effects of immersive gameplay on cognitive skill acquisition (Ackerman et al. 2010; Baniqued et al. 2013; Boot et al. 2008) so clearly, more research is needed.

One future GBL research thread may aim to bridge the gap between the worlds of commercial versus educational video games. Towards that end, we need a systematic approach to designing game-based assessment that produces engaging video games and valid assessments. This will require an interdisciplinary team (e.g., experts in assessment, game design, content, instructional design, psychometrics, and learning sciences). Sound research methods are needed to explore how well-designed video games can be used as reliable and valid assessments of multiple and diverse competencies. From an assessment standpoint, research is needed to determine the value added of game-based assessment over traditional assessment relative to the quality of the assessment (i.e., validity, reliability, and generalizability) and the impact on learning, engagement, and transfer. This research should also include analyses on how game-based assessment data predict important external criteria (e.g., high school or college graduation, state test scores, and so on) relative to traditional assessments.

The ECD framework used in stealth assessment provides a systematic approach to assessment design and also provides a transparent way to reason about student performance. And while there are other ways to develop assessments, they often lack transparency as well as specification about the competencies and tend to focus on tasks that are too simple or inauthentic. This can result in creating assessments that measure unintended competencies and thus damage the reliability and validity of the assessment. In addition, using an ECD approach for the design of game-based assessments can provide a clear research and communication framework for assessment/learning experts and game designers who want to design and develop new educational and engaging game-based assessments that accurately measure important traditional and new competencies.

One final hurdle to surmount to weave such immersive video games into classrooms concerns the need for these games to provide easy-to-interpret reports to stakeholders – such as teachers, students, and parents. There are many ways an assessment can accumulate data (e.g., tallying frequency counts of events, automatically logging player information then mining the log file to make estimates of competency states via Bayesian networks). Regardless of the methodology for accumulating evidence, the ensuing estimates can be reported at various grain sizes (e.g., general problem-solving skill or specific facets) for diagnostic purposes (see Almond et al. 2009). And with regard to validity issues of game-based assessments, when embedded assessments are not only ongoing but also invisible, this can remove test anxiety, again leading to a more valid assessment.

In conclusion, we believe that conducting research with the goal of improving twenty-first century competencies like problem-solving skills for learners through immersive video gameplay environments is relevant and important to pursue given the emerging challenges of a globalized economy (i.e., renewable energy, pollution, climate change) and consequent problems to solve.

References

Ackerman, P. L., Kanfer, R., & Calderwood, C. (2010). Use it or lose it? Wii brain exercise practice and reading for domain knowledge. *Psychology and Aging, 25*(4), 753–766.

Almond, R. G. (2010). Using evidence centered design to think about assessments. In V. J. Shute & B. J. Becker (Eds.), *Innovative assessment for the 21st century: Supporting educational needs* (pp. 75–100). New York: Springer.

Almond, R. G., Shute, V. J., Underwood, J. S., & Zapata-Rivera, D. (2009). Bayesian networks: A teacher's view. *International Journal of Approximate Reasoning, 50*, 450–460.

Anderson, J. R. (1980). *Cognitive psychology and its implications*. New York: Freeman.

Baniqued, P. L., Kranz, M. B., Voss, M. W., Lee, H., Cosman, J. D., Severson, J., & Kramer, A. F. (2013). Cognitive training with casual video games: Points to consider. *Frontiers in Psychology, 4*, 1010.

Barab, S., & Dede, C. (2007). Games and immersive participatory simulations for science education: An emerging type of curricula. *Journal of Science Education and Technology, 16*(1), 1–3.

Barzilai, S., & Blau, I. (2014). Scaffolding game-based learning: Impact on learning achievements, perceived learning, and game experiences. *Computers & Education, 70*, 65–79.

Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M., & Gratton, G. (2008). The effects of video game playing on attention, memory, and executive control. *Acta Psychologica, 129*(3), 387–398.

Bransford, J. D., & Stein, B. S. (1984). *The IDEAL problem solver. A guide for improving thinking, learning, and creativity*. New York: Freeman.

Chuang, T.-Y., & Chen, W.-F. (2009). Effect of computer-based video games on children: An experimental study. *Educational Technology & Society, 12*(2), 1–10.

Clark, D. B., Tanner-Smith, E., & Killingsworth, S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research, 86*(1), 79–122. https://doi.org/10.3102/0034654315582065.

Cukier, K. N., & Mayer-Schoenberger, V. (2013). The rise of Big Data: How it's changing the way we think about the world. *Foreign Affairs, 92*(3). Retrieved 19 Aug 2017 from http://www.foreignaffairs.com/articles/139104/kenneth-neil-cukier-and-viktor-mayerschoenberger/the-rise-of-big-data.

Dede, C. (2005). Planning for neomillennial learning styles. *EDUCAUSE Quarterly, 28*(1), 7–12.

Dede, C. (2009). Immersive interfaces for engagement and learning. *Science, 323*(66), 66–69.

DiCerbo, K. E., & Behrens, J. T. (2012). Implications of the digital ocean on current and future assessment. In R. Lissitz & H. Jiao (Eds.), *Computers and their impact on state assessment: Recent history and predictions for the future* (pp. 273–306). Charlotte: Information Age Publishing.

Duncker, K. (1945). The structure and dynamics of problem-solving processes. *Psychological Monographs, 58*(5), 1–112.

Entertainment Software Association. (2016). *2016 Essential facts about the computer and video game industry*. The Entertainment Software Association. Retrieved from http://www.theesa.com/wp-content/uploads/2016/04/Essential-Facts-2016.pdf.

Gagné, R. M. (1980). Learnable aspects of problem solving. *Educational Psychologist, 15*(2), 84–92.

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

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

Gee, J. P. (2008). Learning and games. In K. Salen (Ed.), *The ecology of games: Connecting youth, games, and learning* (pp. 21–40). Cambridge, MA: MIT Press.

Gee, J. P., & Hayes, E. (2011). Nurturing affinity spaces and game-based learning. In C. Steinkuehler, K. Squire, & S. Barab (Eds.), *Games, learning, and society: Learning and meaning in the digital age* (pp. 129–153). New York: Cambridge University Press.

Gick, M. L. (1986). Problem-solving strategies. *Educational Psychologist, 21*(1–2), 99–120.

Gobert, J., Sao Pedro, M., Raziuddin, J., & Baker, R. (2013). From log files to assessment metrics for science inquiry using educational data mining. *Journal of the Learning Sciences, 22*(4), 521–563.

Green, C. S., & Bavelier, D. (2012). Learning, attentional control, and action video games. *Current Biology, 22*(6), 197–206.

Hart Research Associates. (2015). *Falling short? College learning and career success*. Washington, DC: Association of American Colleges and Universities.

Ifenthaler, D., Eseryel, D., & Ge, X. (2012). Assessment for game-based learning. In D. Ifenthaler, D. Eseryel, & X. Ge (Eds.), *Assessment in game-based learning. Foundations, innovations, and perspectives* (pp. 3–10). New York: Springer.

Jackson, L. A., Witt, E. A., Games, A. I., Fitzgerald, H. E., von Eye, A., & Zhao, Y. (2012). Information technology use and creativity: Findings from the Children and Technology Project. *Computers in Human Behavior, 28*, 370–376. https://doi.org/10.1016/j.chb.2011.10.006.

Jonassen, D. H. (1997). Instructional design models for well-structured and III-structured problem-solving learning outcomes. *Educational Technology Research and Development, 45*(1), 65–94.

Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educational Technology Research and Development, 48*(4), 63–85.

Jonassen, D. H. (2002). Engaging and supporting problem solving in online learning. *Quarterly Review of Distance Education, 3*(1), 1–13.

Jonassen, D. (2003). Using cognitive tools to represent problems. *Journal of Research on Technology in Education, 35*(3), 362–381.
Krulak, C. (1997). *Military thinking and decision-making exercises* (No. 1500.55). Washington, DC. Retrieved 19 Aug 2017 from http://www.marines.mil/Portals/59/Publications/MCO1500.55.pdf.
Lanyon, R. I., & Goodstein, L. D. (1997). *Personality assessment* (3rd ed.). New York: Wiley.
Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-sense of presence inventory. *Presence, 10*(3), 282–297.
Mayer, R. E. (1992). *Thinking, problem solving, cognition* (2nd ed.). New York: Freeman.
Mayer, R., & Wittrock, M. (2006). Problem solving. In P. Alexander & P. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 287–303). Mahwah: Erlbaum.
Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). On the structure of educational assessment. *Measurement: Interdisciplinary Research and Perspective, 1*(1), 3–62.
Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49–100.
Newell, A., Shaw, J. C., & Simon, H. A. (1958). Elements of a theory of human problem solving. *Psychological Review, 65*(3), 151–166.
Pagliery, J. (2014). *Why I put World of Warcraft on my resume*. Retrieved 26 Nov 2016 from http://money.cnn.com/2014/06/19/technology/world-of-warcraft-resume/.
Partnership for 21st Century Learning. (2016). Retrieved 21 Nov 2016 from http://www.p21.org/storage/documents/docs/P21_framework_0116.pdf.
Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes: Measures of social psychological attitudes* (Vol. 1, pp. 17–59). San Diego: Academic.
Polya, G. (1945). *How to solve it: A new aspect of mathematical method*. Princeton: Princeton University Press.
Quellmalz, E. S., Timms, M. J., Silberglitt, M. D., & Buckley, B. C. (2012). Science assessments for all: Integrating science simulations into balanced state science assessment systems. *Journal of Research in Science Teaching, 49*(3), 363–393.
Rowe, J. P., Shores, L. R., Mott, B. W., & Lester, J. C. (2011). Integrating learning, problem solving, and engagement in narrative-centered learning environments. *International Journal of Artificial Intelligence in Education, 21*(1), 115–133.
Schrader, P., Zheng, D., & Young, M. (2006). Teachers' perceptions of video games: MMOGs and the future of preservice teacher education. *Innovate, 2*(3). Retrieved April 17, 2018 from https://nsuworks.nova.edu/innovate/vol2/iss3/5/
Schweizer, F., Wüstenberg, S., & Greiff, S. (2013). Validity of the MicroDYN approach: Complex problem solving predicts school grades beyond working memory capacity. *Learning and Individual Differences, 24*, 42–52.
Selfe, C. L., & Hawisher, G. E. (2004). *Literate lives in the information age: Narratives on literacy from the United States*. Mahwah: Erlbaum.
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: Information Age Publishers.
Shute, V. J., & Ventura, M. (2013). *Measuring and supporting learning in games: Stealth assessment*. Cambridge, MA: MIT Press.
Shute, V. J., & Wang, L. (2016). Assessing and supporting hard-to-measure constructs. In A. A. Rupp & J. P. Leighton (Eds.), *The handbook of cognition and assessment: Frameworks, methodologies, and application* (pp. 535–562). Hoboken: Wiley.
Shute, V. J., Ventura, M., & Ke, F. (2015a). The power of play: The effects of Portal 2 and Lumosity on cognitive and noncognitive skills. *Computers & Education, 80*, 58–67.
Shute, V. J., D'Mello, S. K., Baker, R., Bosch, N., Ocumpaugh, J., Ventura, M., & Almeda, V. (2015b). Modeling how incoming knowledge, persistence, affective states, and in-game

progress influence student learning from an educational game. *Computers & Education, 86*, 224–235.

Shute, V. J., Wang, L., Greiff, S., Zhao, W., & Moore, G. R. (2016). Measuring problem solving skills via stealth assessment in an engaging video game. *Computers in Human Behavior, 63*, 106–117.

Shute, V. J., Ke, F., & Wang, L. (2017). Assessment and adaptation in games. In P. Wouters & H. van Oostendorp (Eds.), *Instructional techniques to facilitate learning and motivation of serious games* (pp. 59–78). New York: Springer.

Simon, H. A., & Newell, A. (1971). Human problem solving: The state of the theory in 1970. *American Psychologist, 26*(2), 145–159.

Sitzmann, T. (2011). A meta-analysis of self-regulated learning in work-related training and educational attainment: What we know and where we need to go. *Psychological Bulletin, 137*, 421–442.

Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Review, Monograph Supplement, 2*(4), 1–8.

Van Eck, R. (2006). Digital game-based learning: It's not just the digital natives who are restless. *EDUCAUSE review, 41*(2), 1–16.

Van Eck, R., Hung, W., Bowman, R., & Love, S. (2009). 21st century game design: A model and prototype for promoting scientific problem solving. In *Proceedings of the twelfth IASTED international conference on computers and advanced technology in education: Globalization of education through advanced technology*. Calgary: ACTA Press.

Van Eck, R. N., Shute, V. J., & Rieber, L. P. (2017). Leveling up: Game design research and practice for instructional designers. In R. Reiser & J. Dempsey (Eds.), *Trends and issues in instructional design and technology* (4rd ed., pp. 227–285). Upper Saddle River: Pearson Education.

Ventura, M., Shute, V. J., Wright, T., & Zhao, W. (2013). An investigation of the validity of the virtual spatial navigation assessment. *Frontiers in Psychology, 4*, 1–7.

Wouters, P. J. M., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology, 105*, 249–265.