13.

How'd That Happen?!

Failure in Game Spaces to Prepare Students for Future Learning

Alison Lee (Teachers College, Columbia University), Connie Liu (Teachers College, Columbia University), Mathurada Jullamon (Teachers College, Columbia University), & John Black (Teachers College, Columbia University)

Abstract

Educational games can be used effectively in the classroom, because exploration and responses to failure in game spaces can afford productive metacognition that better prepares students for future learning (PFL). In our study, we explored how the role of failure-induced metacognitive appraisal and strategy selection while playing a physics game better prepares students for learning from formal content after gameplay. Our results indicate that experiencing failure to prepare students for future learning can elicit more complex and robust mental representations of a complicated science system. However, experiencing failure unto itself isn't sufficient for improving general conceptual understanding – a good metacognitive response is required. Further investigation identified response to failure with info-seeking, then fixing one's answer was found significantly related to learning because it entails an appraisal of knowledge gaps, resolving such gaps through info-seeking, and apply newly acquired information to address prior misconceptions.

Theoretical Background

Games have been investigated for a variety of educational purposes, and despite mixed reporting, meta-analyses indicate that they can be good for learning under specific conditions (Wouters et al, 2013). Generally, games are purported to be good for learning because of (but not exclusive to) the following reasons: games can be highly motivating and intrinsically interesting to students because of mechanics such as narratives, reward systems, agency, and competition (Malone, 1981; Yee, 2006); games can provide the optimization of difficulty and scaffolding that naturally promotes persistence (Gee, 2005; Juul, 2013); games can provide clear-cut goal states, feedback, and opportunities for revision/multiple attempts (Shute et al, 2009); and games provide a "low-stakes" environment through which players can explore and fail without heavy repercussions (Juul, 2013). Furthermore, games and simulations can provide constrained environments that highlight key concepts and interactions within a system, which can be particularly useful for science content that can be abstract and difficult to understand (Honey & Hilton, 2011). However, games alone are not a panacea or replacement for explicit instruction. One of the most effective ways of using games for education identified by Wouters et al.'s (2013) meta-analysis is combining serious games with instructional content, because game experience provide students with "well-structured prior knowledge" that they can then access and build on during later learning. It is in this premise – that games can provide grounded exploration and interactions with abstract content that can inform later formal learning – that is the subject of investigation in this paper.

Transfer, or the ability to take what is learned or experienced in one context to use in another, is a fundamental goal of learning. However, a century of investigations and epistemological debate has yielded ambivalent conclusions about precisely what constitutes and yields transfer of knowledge (Barnett & Ceci, 2002; Chen & Klahr, 2008; Detterman & Sternberg, 1993; Singley & Anderson, 1989; Thorndike & Woodworth, 1901). Recently, investigations in the field of the learning sciences have led to novel approaches in the instruction, definition, and utility of transfer for learning and skill development. In the Preparation for Future Learning paradigm, transfer is defined as the use of prior experiences to inform and improve later learning (Bransford & Schwartz, 1999). The idea is that oftentimes we use prior knowledge to notice and frame new information, and that these "knowing with" kinds of prior experiences can greatly shape and improve understanding of the new context. Having relevant, familiar prior experiences can prepare you to ask the right questions, notice the important components, and therefore lead to deeper, more robust conceptual models of the learned material. While PFL may seem to be a common-sense approach to education, PFL as a framework for pedagogy is a relatively new approach.

PFL studies highlight the utility of activities where students explore and grapple with relevant content. Oftentimes, these activities are explicitly designed such that students are thrust into a problem-solving environment that compels them to wrestle with underlying principles of the concepts. Schwartz et al.'s (2011) work on inventing with contrasting cases as PFL demonstrated that even when students were not always ultimately successful in their inventions prior to learning, their invention experiences before instruction led to better transfer outcomes on two dimensions: better conceptual understanding of ratio structures in physics, and better application of this ratio structure to other domains. This finding demonstrates that students who explored the underlying principles of ratios through invention learned, abstracted, and applied these concepts better than those who took the traditional route of learning first and then practicing ("tell and practice"). Their conclusions suggest that a critical mechanism of this PFL activity is fostering an "appreciation of the deep structure" of the concept such that students readily called upon their experiences with this deep structure when learning about the formal concept later on. As such, we can look at transfer as a process, where accessing prior experiences and information is a skill to be cultivated for more effective learning, rather than only looking at transfer as an indicator that learning has occurred. However, they do not discuss what specific mechanisms of invention-with-contrasting-cases led to greater noticing of the deep structure. Thus, there are some questions not yet answered: what is the role of iteration, failure, strategy, and realizations about insufficient solutions in noticing these deep structures?

Manu Kapur's (2008) work with PFL attempted to address some of these questions by isolating failure as a vital component of preparing students for future learning. Kapur used the PFL framework to design an intervention using either well-structured (and scaffolded) problems or ill-structured problems prior to formal learning, His work revealed that despite students in the ill-structured condition struggling with defining, analyzing, and solving their problems (in other words, failing to generate explicit understanding of the concepts or effective solutions), these experiences were more conducive to learning later on. This phenomenon, which he coined "productive failure", demonstrated that success in the traditional sense (that is, success in clearly defining concepts and generating effective solutions) may not necessarily lead to greater learning; in fact, designing problem solving tasks that scaffolds and directs learning towards "success" may unwittingly undermine the effortful cognition that could benefit formal learning later on. Instead, environments and tasks that permits for students to initially fail and grapple with concepts rather than "succeed" can be more beneficial to future learning. However, Kapur also did not elucidate what specifically in the productive failure space led to greater learning. Does failure in itself call attention to deep features? Or does failure afford opportunities to engage in cognition that then leads to deep feature noticing? How much failure is sufficient for PFL? What makes the failure productive? Are there ways of designing tasks that promote productive failure, rather than just plain failure (that is, if there is even in fact a difference between regular failure and productive failure)?

Loibl & Rummel (2014) addressed some of these questions by asserting that productive failure improves learning by calling students' attention to the gaps in understanding when they confront a failure. In their work, they demonstrated that attempting to solve problems before formal learning can lead to a global awareness of knowledge gaps – that is, acknowledging that some component of their understanding is incomplete without specification. This awareness is a kind of global cognitive appraisal that arises from students' inability to solve the problems (a failure),that are then fully specified and addressed in teacher instruction. However, while they discuss global knowledge gaps (global metacognitive awareness) as a mediator for failure to positively impact learning, they do not explicitly discuss moment-to-moment metacognitive behaviors that happen in response to the failure that can also be productive for learning.

Much of the work surrounding failure in educational research entails demonstrating the importance of resilience in the face of failure, with advocates from pop culture like JK Rowling and Steve Jobs, to the seminal works on mindset by Carol Dweck (2006) and grit by Angela Duckworth (2007), who largely emphasize affective resilience in response to failure (that is, how we can encourage people to persist in the face of failure, given that failure is often the pathway to success). However, little has been done to identify when and how failure can be useful – that is, what are the kinds of "necessary and sufficient" conditions of the task, the learner, and the instructional method such that failure is productive? For example, are there certain kinds of reflection that should be happening in the failure space, or does productive failure entail the capacity to select the appropriate consequent actions in response to failure? What about the role of acknowledging and pinpointing what caused the failure, presumably a vital part of noticing the deep features of a concept?

These components are encapsulated by metacognition, or the ability to judge and monitor one's own states of knowing, and employ strategies to improve understanding. Metacognition is a critical part of learning because it permits learners to identify, more deeply understand and effectively address gaps in knowing (Flavell, 1979). Metacognitive processes are commonly discussed as a critical component of teaching students to transfer because the act of self-monitoring helps facilitate the recognition of when the information or strategy might be relevant in other contexts (Perkins & Salomon, 1992; Belmont, Butterfield, & Ferretti, 1982; Adey & Shayer, 1993). Each moment of failure affords an opportunity to make a metacognitive judgment about what knowledge component is lacking. In addition, environments like games that afford opportunities to address those gaps (through revision, hint-seeking, and scaffolds) can also mediate deeper conceptual understanding later on. We argue that students who act in more reflective ways during the PFL activity, whether those metacognitive behaviors are enacted by natural predilection or provoked by the environment, would attend more carefully to deep features and therefore will be more prepared to learn from future learning activities. Furthermore, metacognition is especially valuable in failure spaces, because the most "productive" affordance of failure is to address head-on what those gaps between expected and actual outcomes are, and what actions should be taken to resolve them. Presumably, what makes productive failure good for preparing students for future learning is contingent on students' abilities to reflect on their incorrect solutions, address gaps in knowledge, and select strategies and actions in response to these appraisals. It is through these metacognitive mechanisms that cue students to identify and engage with deep features of the concepts. However, getting students to engage in these metacognitive behaviors is an arduous task – students are often intimidated by failure, particularly in school tasks where failure often involves high-stakes consequences, such as failing a quiz or getting a low score on your homework. Furthermore, these common school tasks do not often permit or encourage efforts to respond to those failures – that is, they don't provide the tools, encouragement or opportunities for students to review their incorrect solutions, appraise where knowledge gaps occur, seek to close such gaps, and fix their solutions. What curricular tools might provide low-stakes, engaging problem-solving environments that encourage student iteration, permit for metacognitive behaviors, and allow for exploration of academic content in meaningful, goal directed ways? Games fit all of these criteria.

Failure is a critical component of games, where the process of failing (a level, a fight, a boss, a puzzle) is inherent in the game design in order for it to be compelling and entertaining. People appear to be incredibly productive when encountering failure in games, where they use the failure experience to inform future decision-making and understanding of the problem space. These kinds of metacognitive behaviors – reflecting, judging the goodness of one's performance, coordinating strategies, planning next actions to address what went wrong previously – are ones we strive for students to employ, but are enacted so naturally in game environments. Furthermore, game spaces seem to promote resilient behaviors in the face of failure – perhaps because the failures do not have high stakes (outside of the game), and therefore does not negatively impact motivation. On the contrary, despite deliberate designs for inducing failure, games seem to encourage engagement and persistence, even (and perhaps especially) when the player is frustrated and confused (Juul, 2013; Csikszentmihalyi & Larsen, 1980). Games are a natural fit as a PFL activity, as they provide a space for exploring real systems, especially those that would otherwise be impossible to interact with in real life, and are goal directed and constrained to cue learners to key concepts and system structures (Malone, 1981; Garris, Ahlers, & Driskell, 2002; Black, Khan, & Huang, 2014; Reese, 2007). Situating exploration, problem-solving, and systems manipulation in a game can be a powerful method for generating intuitions about a particular concept or system (Garris et al., 2002; Honey & Hilton, 2011). These grounding experiences can prepare students to better learn from formal content later on (Hammer & Black, 2009). Therefore, the game space is essential for us to investigate what cognitive mechanisms are at play in failure that are good for future learning, while alleviating the concerns about motivation and the high-stakes nature of failure in school tasks.

Given these threads of research on games, preparation for future learning, productive failure, and metacognition, we seek to investigate what role metacognitive responses to failure in a physics game play in preparing students for future learning. Thus, the broad scope of this research is to ask the following questions:

Q1: Does the affordance for responding to failure elicit deeper conceptual understanding?

Q2: Are there particular metacognitive responses to failure that are better for learning?

Study Design

36 adult subjects were recruited from a NYC university. 83% were pursuing a graduate degree, and all of them held at least a bachelor's degree. All subjects reported reported low prior knowledge of

the concepts covered in the study. Subjects learned about basic principles of direct current circuits by playing a game called Electropocalypse (the PFL activity) and by watching Khan Academy videos about direct current circuits. Electropocalypse, a puzzle game, takes players through a narrative where they act as electrical engineers who must reconfigure electrical circuits to meet level goals, each involving another physics principle. Subjects played a subset of levels (1-13) covering content like closed/open loops, short-circuiting, and resistors in series and parallel. All subjects played the game, watched four Khan Academy videos, responded to surveys, and completed three Open-Ended Worksheets (OEs) and a Post-Test. Subjects were randomly assigned to three study conditions: Tell-and-Practice (Tell), Full PFL (Full), and No Failure Response PFL (NFR). The Tell served as the control condition where students first watched the video, followed by the game (See Figure 1). In the Full condition, students played the game, followed by watching the videos. In the NFR condition, students followed the same order as the Full condition, but played a version of the game where if they submitted an incorrect solution (failure), they were not permitted to view or fix their solution, but instead received an explanation and screenshot of the correct solution (see Figure 2). This NFR condition was created so that we could isolate the effect of metacognitive responses to failure without manipulating the naturally-occurring amount of failure subjects experienced when playing each level for the first time. Furthermore, this condition mimics the structure of many common classroom activities, where students complete a set of problems (like homework), receive their grade indicating whether their solution was correct or not, but do not have the opportunity to fix or engage in metacognitive behaviors in response to those solutions.

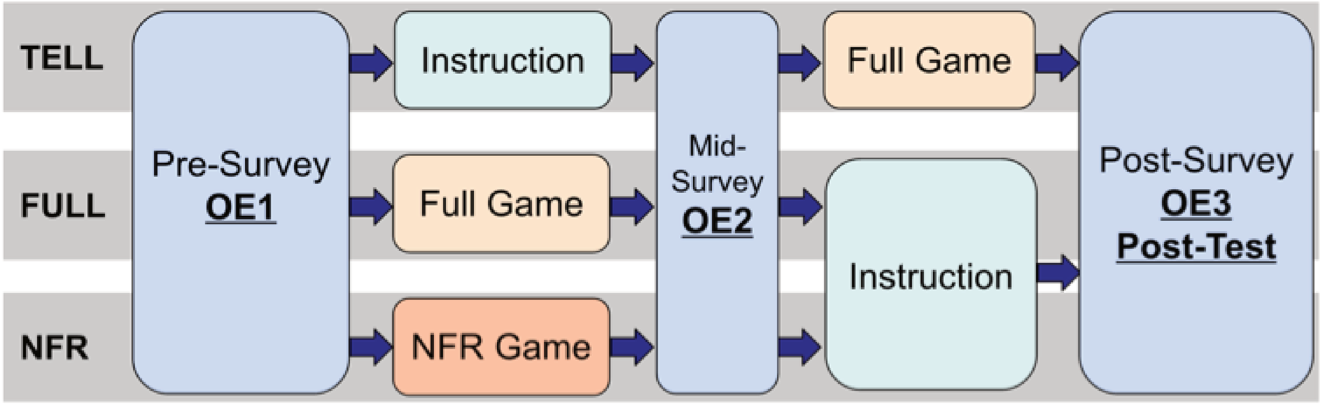

Figure 1: Study Design

Figure 2: Normal Feedback Screen (left) vs. NFR Feedback Screen (right)

Learning and behavioral measures evaluated each subject's learning of physics principles, complexity of his/her her conceptual model, and his/her metacognitive response to failure. Learning measures were assessed through three OE Worksheets and a Post-Test. The OEs contained the following prompt: "Draw and explain a parallel circuit. Be sure to label all relevant parts of the circuit system, explain what a parallel circuit is, and how it differs from a serial circuit." OE responses were coded on two dimensions: correctness and complexity. The OEs were given at three time points: TP1 (pre-measure), TP2 (after first activity), TP3 (post-measure). The Post-Test comprised of four sections: three multiple choice questions about Ohm's Law, three questions on reasoning about a diagram, two transfer questions that asked subjects to reason about water pipes based off of direct current circuit concepts and relating concepts/components of the water pipe system to direct current circuits (a common analogous reasoning prompt used to measure transfer), and two PFL questions on Voltmeters and Ammeters. Post-Test responses were also coded based on content, transfer, and complexity. Behavioral measures were assessed through log data from the game. The log data generated a list of subjects' actions over the duration of the gameplay. Analysis of the log data identified action sequences in response to failure, duration of time spent on a particular action or level, and how subjects navigates throughout the game.

Results

Learning Outcomes

There were no significant findings in Post-Test content (*p*=.227), transfer(*p*=.774), or complexity scores (*p*=.814). A repeated measures ANOVA on OE correctness scores also indicated that there were no significant differences in OE correctness scores by OE3 (*p*=.238). However, while learning of the physics concepts equally occurred across all of the conditions, subjects in the Full Condition produced more complex explanations of direct current circuits in parallel and series. A RM ANOVA on OE complexity scores revealed that the Full condition provided more robust explanations (F[4,64]=6.213, *p*<.001) of electrical circuits at TP3 than the other two groups (See Figure 3).

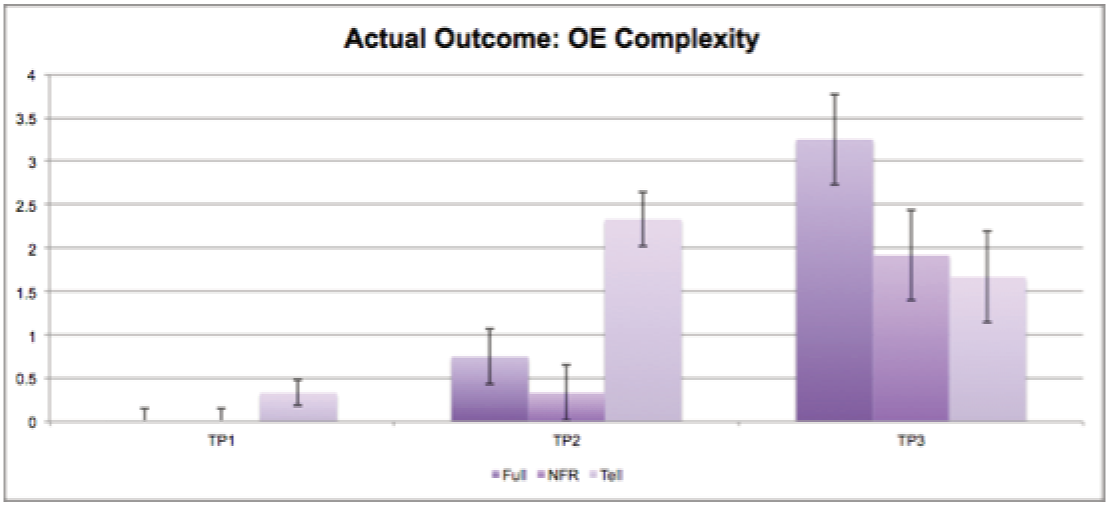

Figure 3. OE complexity score.

Behavioral Outcomes

Game log analysis also identified eight distinct responses to failure. Behavioral responses to failure included “Info-seeking, restarting the level”, “Info-seeking, fixing current solution”, “Fixing current solution”, “Quick Resubmit (did not change solution)”, “Restarted Level”, “Skipped backwards”, “Skipped forward”, and “Skipped to next level.” Of the eight responses to failure identified, “Info-seeking, then fixing your answer” was positively correlated to Post-Test complexity. (r[13]=.694, *p*=.009).

General Discussion

Our findings suggest that engaging with failure in game spaces before formal learning can elicit more nuanced mental representations of a complicated science system. While there were no condition differences in Post-Test and OE3 conceptual understanding measures, we found that students in the Full condition produced more complex and robust explanations of parallel and serial circuits in the last open-ended worksheet. This suggests that while the various methods of using games for learning across these three conditions can produce benefits to conceptual understanding, the affordance of experiencing failure can better prepare you to learn from formal content later on, thereby producing more nuanced and rich conceptual models. However, our expected finding that the Full condition would also perform better on the Post-Test analogous reasoning transfer and PFL measures was not confirmed, nor were the complexity of their responses on the post-test any better or worse than the other two conditions. One possible explanation for this is that our particular population – adult graduate students at a reputable institution of higher learning – already possessed the intuitive and grounded prior knowledge that would be required to make the analogy between electrical circuit and water pipe systems, thus negating any benefits that using games as preparation for future learning would otherwise yield. Another possibility is that our study did not provide enough of a treatment to have a significant effect on these other dimensions – after all, the game session only lasted 45 minutes and occurred only once. This is supported by Wouters et al.’s (2013) meta-analysis, which suggests that games are better for learning when there are multiple sessions of gameplay. Another possibility is that the measures of transfer used – the analogous reasoning and PFL questions – were not sensitive enough to detect significant differences between groups, or that they were not the right kind of transfer assessment to use for this kind of learning.

These results suggest that simply engaging in failure unto itself is insufficient for improving general conceptual understanding more than just playing the game with learning, even though failure can benefit the complexity of one’s conceptual model. However, our results did demonstrate that there are effortful metacognitive behaviors that are related to better learning. We found that of all the responses that one can take in response to failure in our game environment, the response of “info-seeking, then fixing one’s answer” was significantly positively related to learning. When considered within the framework of metacognition, this is beneficial because it required subjects to appraise and become aware of knowledge gaps, resolve identified gaps through info-seeking, and then apply the newly acquired information to adjust prior misconceptions. All three of these components – the appraisal (or the awareness of knowledge gaps), the resolution (or “filling-in”), and the application – are equally critical to learning. This is in contrast to Loibl & Rummel’s (2014) conclusion that the awareness of knowledge gaps alone account for the benefits of productive failure for learning – we see that a general awareness is not as important here (as evidenced by our nonsignificant differences in Post-Test conceptual questions) so much as the specified appraisals of what one does not know *in the moment of failure*. Furthermore, we see the importance of the "application" component when contrasting "info-seeking, fixing" with "info-seeking, restarting" – if resolving knowledge gaps alone (through info-seeking) were sufficient for productive failure, then we should have seen that both of these strategies would be statistically significantly related to learning. However, seeing that fixing one's solution after info-seeking may have been the key component suggests that the metacognitive actions taken after metacognitive judgments made are just as important as the judgments themselves – knowing that this newly acquired information must be used to address previous failures is a vital part of the learning process. This may be related to the "appreciation for the deep structures" Schwartz et al. (2011) referred to in their own PFL activity – that is, info-seeking and fixing one's solution may have led to greater intuitions about the underlying concepts and structures of parallel and serial circuits that then led to more complex conceptual models. An alternative explanation is that these other behaviors may also be significantly related to learning, but that our sample size for these correlations are too small and therefore not detectable.

To conclude, using failure spaces in games for learning can be an effective way of improving the complexity of students' conceptual model, but have tenuous impact on more general conceptual learning and transfer. Furthermore, the most effective metacognitive response to failure in our game was to fill in knowledge gaps through information seeking, then applying this new information towards fixing one's prior incorrect solution. However, we are limited in our conclusions about the degree to which games in general can be an effective PFL activity, because our population may have already possessed the appropriate prior intuitions about the system, or because our transfer measures were insufficient, or our treatment duration was insufficient. To address these limitations, future studies should examine whether this game can be used as an effective PFL activity for younger students with little or no exposure to the content, using alternative transfer measures such as delayed assessments, and using multiple game sessions for the treatment. Above all, future studies should compare whether the addition of metacognitive prompts after failure that provoke students to appraise, info-seek, and apply information to fix solutions in the game can produce better learning outcomes in both conceptual learning and conceptual complexity.

References

Adey, P., & Shayer, M. (1993). An exploration of long-term far-transfer effects following an extended intervention program in the high school science curriculum. *Cognition and instruction*, *11*(1), 1-29.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin*, *128*, 612–637.

Belmont, J. M., Butterfield, E. C., & Ferretti, R. P. (1982). To secure transfer of training instruct self-management skills. *How and how much can intelligence be increased*, 147-154

Black, J.B., Khan, S.A., & Huang, S.C.D. (2014). Video games as grounding experiences for learning.

In F. C. Blumberg (Ed.) *Learning by playing: Frontiers of video gaming in education*. New York: Oxford University Press.

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of research in education*, 61-100.

Chen, Z., & Klahr, D. (2008). Bridging the gap: Remote transfer of problem-solving and scientific reasoning strategies in children. In R. Kail (Ed.), Advances in child development and behavior (Vol. 36, pp. 419–470). Burlington, MA: Academic Press.

Csikszentmihalyi, M., & Larson, R. (1980). Intrinsic rewards in school crime. In M. Verble (Ed.), *in Discipline*, Omaha, NE: University of Mid-America.

Detterman, D. K., & Sternberg, R. J. (1993). Transfer on trial: Intelligence, cognition, and instruction. Norwood, NJ: Ablex.

Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York: Random House.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: perseverance and passion for long-term goals. *Journal of personality and social psychology*, *92*(6), 1087.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American psychologist*, *34*(10), 906.

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

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

Hammer, J., & Black, J. (2009). Games and (preparation for future) learning. *Educational Technology Magazine: The Magazine for Managers of Change in Education*, *49*, 29-34.

Honey, M. A., & Hilton, M. (Eds.). (2011). *Learning science through computer games and simulations*. Washington, D.C: The National Academies Press.

Juul, J. (2013). *The art of failure: An essay on the pain of playing video games*. Mit Press.

Kapur, M. (2008). Productive failure. *Cognition and Instruction*, *26*(3), 379–424.

Loibl, K., & Rummel, N. (2014). Knowing what you don't know makes failure productive. *Learning and Instruction*, *34*, 74-85.

Malone, T. W. (1981). What makes things fun to learn? A study of intrinsically motivating computer games. *Pipeline*, *6*(2), 50.

Perkins, D. N., & Salomon, G. (1992). Transfer of learning. *International encyclopedia of education*, *2*.

Reese, S. D. (2007). The Framing Project: A Bridging Model for Media Research Revisited. *Journal of Communication*, *57*, 148–154.

Schwartz, D. L., Chase, C. C., Oppezzo, M. A., & Chin, D. B. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. *Journal of Educational Psychology*, *103*(4), 759.

Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for learning: The hidden efficiency of original student production in statistics instruction. *Cognition & Instruction*, *22*, 129-184.

Shute, V. J., Ventura, M., Bauer, M., & Zapata-Rivera, D. (2009). Melding the power of serious games and embedded assessment to monitor and foster learning. *Serious games: Mechanisms and effects*, *2*, 295-321.

Singley, M. K., & Anderson, J. R. (1989). The transfer of cognitive skill. Cambridge, MA: Harvard University Press.

Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions: Functions involving attention, observation and discrimination. Psychological Review, 8, 553–564.

Wouters, P., 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*(2), 249.

Yee, N. (2006). Motivations for play in online games. *CyberPsychology & behavior*, *9*(6), 772-775.