Can Video Gameplay Improve Undergraduates' Problem-Solving Skills?

Benjamin Emihovich, University of Michigan - Flint, Flint, USA

 https://orcid.org/0000-0002-9018-2850

Nelson Roque, Pennsylvania State University, State College, USA

 https://orcid.org/0000-0003-1184-202X

Justin Mason, University of Florida, Gainesville, USA

 https://orcid.org/0000-0001-7452-8578

ABSTRACT

In this study, the authors investigated if two distinct types of video gameplay improved undergraduates' problem-solving skills. Two groups of student participants were recruited to play either a roleplaying video game (World of Warcraft; experimental group) or a brain-training video game (CogniFit; control group). Participants were measured on their problem-solving skills before and after 20 hours of video gameplay. Two measures were used to assess problem-solving skills for this study, the Tower of Hanoi and The PISA Problem Solving Test. The Tower of Hanoi measured the rule application component of problem-solving skills and the PISA Problem Solving test measured transfer of problem-solving skills from video gameplay to novel scenarios on the test. No significant differences were found between the two groups on either problem-solving measure. Implications for future studies on game-based learning are discussed.

KEYWORDS

Assessment, Brain Training, Cognition, Gameplay, Problem-Solving Skills, Roleplaying, Rule Application, Transfer, Video Games

INTRODUCTION

Video games are played by more than half of the U.S population and the video game industry generated $36 billion in 2018 (ESA, 2018). Given the popularity and success of the video game industry, game-based scholars are exploring how well-designed video games can be used to improve a wide range of knowledge, skills, and abilities referred to as game-based learning (GBL). Proponents of GBL argue that well-designed video games are grounded by active participation and interaction as the focal point of the learner experience and can lead to changes in behavior and cognition (Ifenthaler, Eseryel, & Ge, 2012; Shute et al., 2019). Moreover, well-designed video games immerse players in environments that can provide a framework for learning experiences by promoting engagement and transfer from simulated worlds to the natural world (Dede, 2009).

Current American students are not receiving adequate exposure to authentic ill-structured problem-solving scenarios in their classrooms, and schools need to address the acquisition of problem-solving skills for students in the 21st century (Shute & Wang, 2016). American students trail their international counterparts in problem-solving skills on the Program for International Student Assessment (PISA) Problem Solving Test. Furthermore, American business leaders complain about recent college graduates' lack of problem-solving skills. Two surveys conducted by the Association of American Colleges and Universities of business leaders and students indicated that problem-solving skills are increasingly desirable for American employers, but only 38% of employers reported that recently hired American college graduates could analyze and solve complex problems while working (Hart Associates, 2018).

Researchers of video game studies find that gameplay can be positively associated with the improvement of problem-solving skills (Shute, Ventura, & Ke, 2015; Spires et al., 2011). However, current discourse in the field of gameplay and problem-solving skills centers primarily on descriptive research (Eseryel et al., 2014) which can be summarized based on the following premise: video games require players to solve problems, and over time, playing video games will lead to improved problem-solving skills (Hung & Van Eck, 2010). Descriptive research is important to argue that video games support problem-solving skills, but further empirical research is needed to demonstrate whether problem-solving skills are acquired through video gameplay. This research study addressed whether two distinct types of video gameplay empirically affects undergraduates' problem-solving skills.

Video Games and Problem-Solving Skills

According to Mayer and Wittrock's (2006) definition, problem solving includes four central characteristics: (1) occurs internally to the problem solver's cognitive system; (2) is a process that involves conceptualizing and manipulating knowledge; (3) is goal directed; and (4) is dependent on the knowledge and skills of the problem solver to establish the difficulty in which obstacles must be overcome to reach a solution. Unlike the well-structured problems that students face in formal learning settings, well-designed games provide students with challenging scenarios that promote problem-solving skills by requiring players to generate new knowledge from challenging scenarios within interactive environments, while also providing immersive gameplay that includes ongoing feedback for the players to hone their problem-solving skills over time (Van Eck, Shute, & Rieber, 2017). Rules govern video gameplay mechanics and one component of problem solving is the ability to apply existing rules in the problem space known as rule application (Shute et al., 2015). One example of a rule application is found in the well-researched problem-solving puzzle the Tower of Hanoi (Huyck & Kreivenas, 2018; Schiff & Vakil, 2015; TOH, 2019). The rule application component of problem-solving skill is one of the dependent variables in this study. Rule application refers to the problem-solver's representation of the problem space through direct action, which is critical to problem solving (Van Eck et al., 2017).

LITERATURE REVIEW

Video Gameplay and Transfer

Researchers contend that the hidden power of well-designed video games is their potential to address higher-level learning, like retention, transfer, and problem-solving skills (Gee, 2008; Shute & Wang, 2015). Retention is the ability to remember the presented information and correctly recall it when needed, while transfer is the ability to apply previously learned information in a novel situation (Stiller & Schworm, 2019). Possible outcomes of playing video games may include the improvement of collaborative problem-solving skills, confidence, and leadership skills that are transferable to the workforce environment. Recent research on video game training studies and transfer of cognitive and noncognitive skills indicates that gameplay is positively associated with the improvement of attention,

problem-solving skills, persistence (Green & Bavelier, 2012; Rowe et al., 2011; Shute et al., 2015; Ventura et al., 2013), executive functions (Oei & Patterson, 2014), and hypothesis testing strategies (Spires et al., 2011). However, other researchers have found null effects of video gameplay and transfer of cognitive skills (Ackerman, et al., 2010; Baniqued, Kranz, et al., 2013; Boot et al., 2008).

A recent meta-analysis of brain-training interventions found that brain-training interventions can improve performance on trained tasks but there were fewer examples of interventions indicating improved performance on closely related tasks, and minimal evidence that training enhances performance on daily cognitive abilities (Simons et al., 2016). Among those finding null effects, questions were raised about the methodological shortcomings of video game training and transfer studies that are common pitfalls in experimental trials. Some of the pitfalls included failing to report full methods used in a study and lack of an effective active control condition that can expect to see similar improvement in competencies as the experimental group (Baniqued et al., 2013; Boot, 2015; Boot, Blakely & Simons, 2011). Unless researchers define recruitment methods for participants and their gaming expertise (novice vs. expert), as well as compare active control groups with experimental groups receiving equal training games, then differential improvement is indeterminable (Boot et al., 2013; Shute et al., 2015). The recruitment approach is outlined in the *Method* section.

Motivation for Selection of Games

The video games selected for this research study were based on the problem-solving skills players exercise and acquire through gameplay that were aligned with the problem-solving skills assessed on the external measures, the PISA Problem Solving Test and the Tower of Hanoi (TOH). Well-designed video games include sound learning principles embedded within gameplay such as requiring players to solve complex problems which can then be applied to other learning contexts (Lieberman et al., 2014). In this study, the authors examined the effects of playing *World of Warcraft* (Activision Blizzard, 2019) and *CogniFit* (CogniFit, 2019) for twenty hours on undergraduates' problem-solving skills (rule application and problem-solving transfer). The inclusion of *CogniFit* addresses a main concern of game-based research which is the lack of an active control condition to determine differential improvement (Boot et al., 2013).

Problem-Solving and Video Gameplay Model

The authors have identified observable in-game behaviors (i.e., indicators) during gameplay that provide evidence for each of the problem-solving processes on the PISA Problem Solving Test. The process included playing each video game extensively, checking community forums for solutions to the most challenging problems for each game, and viewing experts' gameplay video channel streams on YouTube. After generating a list of credible indicators, those selected were based on the following criteria: (a) relevance to the PISA problem solving levels of proficiency and (b) verifiable through gameplay mechanics. Examples of indicators for the PISA problem-solving processes for each game are listed in Tables 1 and 2. The purpose of developing the problem-solving behavior model is to operationalize the indicators of gameplay that align with the cognitive processes being assessed on the PISA test (i.e., Exploring and Understanding, Representing and Formulating). The PISA Problem Solving Test contains questions representing six levels of proficiency: Level 1 is the most limited form of problem-solving ability such as rule application (solving problems with simple rules or constraints) and Level 6 is the complex form of problem-solving ability (executing strategies and developing mental models to solve problems). The PISA test will determine whether there is transfer of problem-solving skills from video gameplay to novel scenarios.

World of Warcraft

Massive multiplayer online role-playing games (MMORPGs) require players to manage resources, adapt playstyle to the environment, test new skills and abilities, identify and apply rules to solve problems as well as explore the story of the game through questing. MMORPGs like *Warcraft* provide

Table 1. Examples of indicators for each PISA problem-solving process in Warcraft

PISA Problem Solving Process	Examples of Indicators
Exploring and Understanding	Prioritize skills and spells that are purchased from vendors in the spell book and action bars; Complete the initial combat introductory quest; Interact with the flight path tool
Representing and Formulating	Use models and charts to assess class and role performance; Analyze pros and cons of equipping awarded weapons and armor in relation to performance
Planning and Executing	Rearrange spells and abilities on the action bar after combat testing (which spells or abilities should be used together and in combination with each other); After combat, prioritize quests and abilities with enemies that can be defeated alone or in groups
Monitoring and Reflecting	Adjust combat distance (short, medium, long) to enemies after testing skills and abilities; Explore the environment for progression; Reorder action bar as new skills are acquired; Use flight path tool to reduce travel time

Table 2. Examples of indicators for each PISA problem-solving process in CogniFit

Problem Solving Process	Examples of Indicators
Exploring and Understanding	Break bricks with the ball and paddle by pressing the space bar and mouse; Avoid letting the ball fall to the bottom of the screen; Use powerups to fire missiles, increase ball speed, or add extra balls
Representing and Formulating	Identify special blocks for bonuses; Test and use missiles to find optimal conditions; Select appropriate powerups based on gem locations on screen in relation to paddle and ball
Planning and Executing	Unlock new paddle and ball abilities after completing each level; Once the ball is released, plan a solution pathway to eliminate all bricks that can work for each level beginning by angling the paddle to direct the ball in the desired direction
Monitoring and Reflecting	Avoid traps and negative powerups; Use missiles under optimal conditions after testing; Save long paddle powerup as ball speed increases

gameplay experiences that are analogous to meaningful instruction by offering complex multifaceted problems that require model-based reasoning—understanding interrelated components of a system, and feedback mechanisms among the components to find the best solutions to problems that arise using available tools and resources in a given environment (Chinn & Malhotra, 2002; Steinkuehler & Chmiel, 2006). Therefore, if MMORPGs provide an authentic sense of inquiry into solving problems through gameplay, then it is worth testing whether these gameplay experiences transfer to novel problem-solving scenarios.

One specific example of transfer from gameplay in the MMORPG *Warcraft* to a natural context concerns the problem of reducing travel time. When players enter the game environment, they must account for extended travel time between different activities such as exploration, questing, and combat. To solve this problem, players are given a tool that can be accessed on their user interface by pressing (M) on their keyboard, which opens the map. Listed on the map are designated flight paths (FPs) that act as a taxi service for players. The image in Figure 1 indicates the various FPs a player has unlocked on their world map as well as those that have yet to be discovered (Activision Blizzard, 2019). The flight path is a handy tool because it connects the goal of completing quests as soon as possible to earn rewards with the knowledge that using flight paths greatly reduces travel time between quests. Greatly reducing travel time results in a more efficient way to complete many of the sub goals in the game, and as noted by Shute and Wang (2016) the use of tools and resources efficiently is an important part of problem solving during gameplay.

Figure 1. Player map listing flight path locations in World of Warcraft (2019)
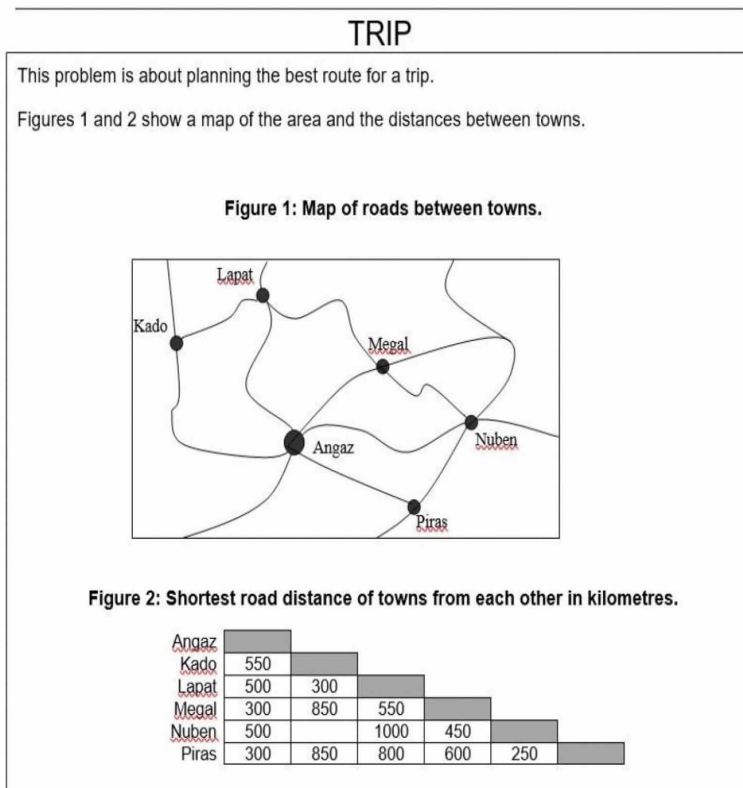

Now, consider one of the questions being assessed on an external measure in the study, the PISA Problem Solving Test. Individuals are given a map that shows the roads between each city, a partially filled-in key that shows distances between cities in kilometers, and the overall layout of the area. The purpose of this question is to assess how individuals calculate the shortest distance from one city to another. To solve the problem, individuals are required to calculate the distance between the two cities of Nuben and Kado using the resources available. This is the same kind of problem that *Warcraft* players experience during gameplay when travelling between locations to complete quests. Both problem scenarios share the same overlapping components, the ability of the problem solver to use given tools and resources efficiently to find the most direct route that reduces travel time between two separate locations. Figure 2 illustrates this problem scenario on the PISA test (OECD, 2003).

CogniFit

The brain training game *CogniFit* claims to have developed a patented system that measures, trains, and monitors cognitive skills like rule application, attention, memory, and visual perception and their relation to neurological pathologies. According to the *CogniFit* (2019) website the company states there are transfer effects from their mini games to problem solving in the natural world. The brain training game is selected as an active control condition based on this claim as well as repeated practice of rule application embedded into the gameplay experience.

One example of rule application in the brain training game *CogniFit* occurs in the mini-game Gem Breaker 3D. This mini-game requires players to direct a paddle back and forth across the screen to bounce a ball off the paddle that breaks the gem blocks without letting the ball touch the bottom of the screen. The initial tutorial informs players that improvement of their hand-eye coordination and processing speed skills are emphasized through gameplay with over 100 levels available to master. Feedback is provided to players with a score for each level showing where they can improve. Once all gem blocks are broken the level is completed and a new level begins. However, each player only

Figure 2. Problem scenario for planning the best route for a trip from PISA (2003)

TRIP

This problem is about planning the best route for a trip.

Figures 1 and 2 show a map of the area and the distances between towns.

Figure 1: Map of roads between towns.

Figure 2: Shortest road distance of towns from each other in kilometres.

	Angaz	Kado	Lapat	Megal	Nuben	Piras
Angaz						
Kado	550					
Lapat	500	300				
Megal	300	850	550			
Nuben	500		1000	450		
Piras	300	850	800	600	250	

has access to 4 balls for each level, and if they lose, the game reverts to the beginning. The tutorial shows players how to use the mouse to control the paddle back and forth across the screen while the spacebar launches the ball. Once a gem is broken there is a chance for a power-up to be gained such as shooting multiple balls, explosives, missiles, side quests or power-ups. Figure 3 illustrates the rules of the mini-game in Gem Breaker 3D (CogniFit, 2019).

Rule application occurs when playing the TOH and requires one to move an entire stack of disks (i.e., a number between 3 and 8) of varied sizes from one of three rods to another. While playing, players are constrained by the following rules: (1) only one disk can be moved at a time; (2) no disk can be placed on a smaller one; (3) only the uppermost disk can be moved on a stack. Rule application is demonstrated by the problem solver in the TOH by configuring the disks and the rods to reach a solution in the problem space. By configuring the disks onto the rods, each move of a disk indicates the problem solver attempting to creatively apply the rules, which is vital to problem solving (Shute et al., 2019). Figure 4 illustrates the problem space in an online version of the TOH (2019).

Rationale

Both video games require players to apply rules to solve problems and rule application is a component of problem solving (Van Eck et al., 2017). As an example, *Warcraft* players learn that they can only cast certain spells in combat while standing still or that eating and drinking food while sitting down hastens the regeneration of health. Similarly, when playing the mini-game Gem Breaker 3D in *CogniFit* players use a paddle and a ball to break bricks. One of the first

Figure 3. Rules for the mini-game Gem Breaker 3D listed in the initial tutorial (2019)

rules players encounter in the game is that they can only move the paddle left or right across the screen or that bonus bricks have special effects like increasing ball speed. The rules are more explicit in *CogniFit* than *Warcraft* so brain-training gameplay may promote better performance on solving the TOH. Each move with the paddle and ball is an example of applying the rules, and this is frequently done during gameplay in *CogniFit*.

However, *CogniFit* mini-games lack some of the salient gameplay features in *Warcraft* such as roleplaying gameplay, meaningful interactions with other players, and richly designed problem spaces that GBL scholars suggest are important to the transfer of problem-solving skills from video gameplay to novel contexts measured on the PISA Problem Solving Test. *Warcraft* gameplay provides players with repeated practice to solve authentic ill-structured problems in rich detailed problem-solving scenarios that may be better suited for transfer to novel scenarios on the test.

Research Questions

After describing the video gameplay conditions of *Warcraft* and *CogniFit* as well as reviewing the literature on problem-solving skills, the authors seek to answer the following research questions:

1. Is there a change, from pretest to posttest, on the rule-application component of problem solving, after 20 hours of video gameplay, on either a role playing or brain-training video game?
2. Does an immersive, collaborative role-playing video game promote transfer of problem-solving skills to novel scenarios better than a brain-training video game for undergraduates after 20 hours of video gameplay?

Figure 4. Problem space in an online version of the Tower of Hanoi puzzle with 5 disks (2019)
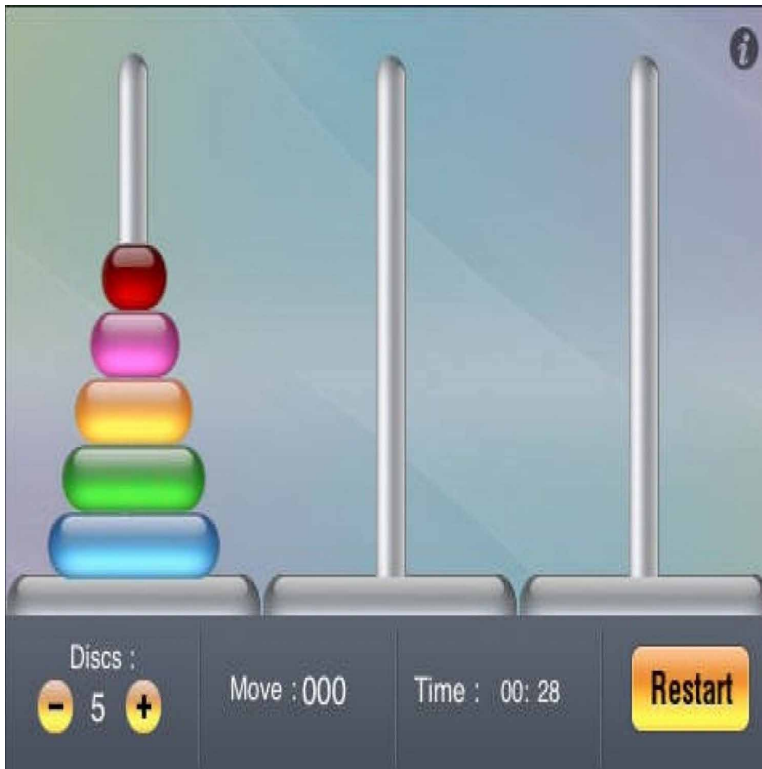
METHOD

Setting and Participants

For this study, 91 undergraduate student participants (M_{Age} = 19.32; SD_{Age} = 1.43) were recruited to participate in this study and completed the initial questionnaire for the study, assessing: age, gender, ethnicity, major, and video games played daily. Participants were not invited to participate if they were not students at the data-collecting institution, were not 18-23 years old, or if they reported playing 30 or more minutes of *Warcraft* or *CogniFit.* 56 participants were randomly assigned to either the experimental group *Warcraft* or the control group *CogniFit,* yet only 34 completed the study (n = 17 per group). Participant attrition for both groups were attributed to lack of time to complete the study or being too busy with schoolwork. Given the nature of our research questions assessing change as a function of training, subsequently presented analyses only include data from the 34 participants (17 males and 17 females) who completed the study (M_{Age} = 19.44; SD_{Age} = 1.41).

Measures

The independent variable in this research study is the video game with two levels: a roleplaying video game (*Warcraft*) and a brain-training video game (*CogniFit*). The video games provide players with repeated problem-solving scenarios requiring players to engage in problem-solving processes. The dependent variable measured for this study is problem-solving skill. One measure assessed the component of rule application of problem solving to solve a puzzle which is the TOH. The second measure assessed problem-solving in novel scenarios which is the PISA Problem Solving Test. Both

groups were assessed on the TOH and the PISA Problem Solving Test. The TOH was used to assess research question 1 and the PISA Problem Solving Test was used to assess research question 2.

The Tower of Hanoi

Recall, the TOH is a valid and reliable experimental paradigm that can be used to assess rule application, problem solving and transfer (Huyck & Kreivenas, 2018; Schiff & Vakil, 2015). Rule application is demonstrated by the problem solver in the TOH by configuring the disks and the rods to reach a solution in the problem space. By configuring the disks on to the rods, each move of a disk indicates the problem solver attempting to creatively apply the rules. Participants played the TOH on a computer from a free website online. The test score (i.e., lower scores are better) for completing the TOH can range anywhere from 31 (which is the minimal number of moves to execute) until it is solved.

PISA Problem Solving Test

The second external problem-solving measure in this study is the (2003) version of the PISA Problem Solving Test. The PISA Problem Solving Test (OECD, 2003) contains 10 novel problem-solving scenarios, and within each scenario there is a range of one to three different questions that must be solved. There are 19 total questions on the test across all scenarios that required students to solve problems. For this study, participants completed five novel problem-solving scenarios for the pretest and the remaining five novel problem-solving scenarios for the posttest. The levels of proficiency for each question are randomized across all problem-solving scenarios. Each problem-solving scenario is independent from one another and each of the 19 questions across all scenarios being assessed in this study are isomorphic from the questions that were implemented in 2003. The scoring for most questions was either correct or incorrect, with some questions allowing for partially correct answers. Participants that answered each question correctly were awarded one point, while partially correct answers awarded participants a half-point.

Procedure

Participants for this study were recruited via flyers posted publicly on campus and dormitory bulletin boards. Over the course of eight weeks, participants engaged in 10 gameplay sessions that lasted two hours each. Participants had the opportunity to complete these 10 sessions in two-hour time-blocks that were made available Monday through Friday for eight consecutive weeks. Participants completed the experiment in a classroom lab on campus at the university. In this experiment, student participants were randomly assigned to play one of two video games.

Participants in the experimental condition played the popular roleplaying video game *Warcraft* that promotes learning new terminologies, mastering interrelated skills and abilities, applying rules to solve problems, goal setting, and reflecting on progress. In addition, participants in the active control condition played the brain-training video game *CogniFit* (2019). The video game allows players to select various mini-games including Gem Breaker 3D that may enhance cognitive abilities including rule application, memory, and focus. Student participants in this study were guided by discovery learning and provided with in-game tutorials for each condition while learning to solve problems through active exploration, interacting with the game environment and self-direction (Westera, 2019).

At pre-test and post-test participants had 20 minutes to complete isomorphic versions of the TOH as many times as possible. All participants successfully completed the TOH once during the pretest and once during the posttest. At pre-test and post-test, participants also had 20 minutes to complete as many questions as possible on The PISA Problem Solving Test. The pretest required participants to answer nine questions and the posttest required participants to answer 10 questions from multiple problem-based scenarios. Each problem-based scenario was unique, and some examples included the following: (1) calculating the distance between two points given a map; (2) developing a decision tree diagram of a library loan system; and (3) calculating daily energy needs for an individual given a set menu.

Data Structure and Analyses

The full dataset used for all analyses to be presented, contained data from 34 participants. All participants attempted three parallel, computerized forms of the TOH at baseline and at the end of the intervention. Due to the nature of the task's programming, if participants did not complete a TOH task, the total number of moves attempted was not output to the data file. This will be expanded upon in the results section by utilizing three analyses which included an independent t-test comparing the mean number of incomplete TOH games between the groups, an independent t-test comparing the mean gain score of TOH between the groups, and a multiple linear regression predicting max gain score of TOH by group, by gain score count, and by group, gain score count, and PISA gain. All analyses in sections below were completed in R, version 3.4.3. Packages used for data analysis include: *dplyr*, for data wrangling (Wickham et al., 2019), and *ggplot2* for visualizations (Wickham, 2016), and *MASS* for stepwise regression analyses (Venables & Ripley, 2002).

RESULTS

Assessing Group Differences in Completion

Although groups differed on the overall number of incomplete TOH sessions at pre-testing ($N_{COGNITIVE}$ = 13; N_{GAMING} = 8), an independent t-test of the average number of incomplete games by group, was not significant (p > .05). Furthermore, an independent t-test revealed no group differences for the overall number of incomplete TOH sessions at post-testing ($N_{COGNITIVE}$ = 3; N_{GAMING} = 2; $p > .05$). A repeated-measures ANOVA revealed a significant time effect, $F(1,32) = 13.386$, p<.001. However, group, $F(1,32) = 1.609$, p=.214, nor group by time interaction were significant, $F(1,32)=.837$, p=.367. On average, participants completed an additional half TOH session (i.e., .47, SD = .53) after receiving either training package (M_{Pre} = .62, SD = .70; M_{Post} = .15, SD = .36). Table 3 shows the means and standard deviations for the pretest and posttest scores participants completed in the experimental (*Warcraft*) and control (*CogniFit*) groups. The mean scores in the table indicate how many moves on average each participant could successfully solve the puzzle per group. For this study, participants had 20 minutes to complete as many questions as possible for the pretest and 20 minutes to do the same for an isomorphic version of the posttest. Table 4 shows the means and standard deviations for the PISA pretest and posttest scores of participants in the experimental (*Warcraft*) and control (*CogniFit*) groups.

Table 3. Pretest and posttest scores by group on the Tower of Hanoi

	Pretest M (SD)	**Posttest M (SD)**
Warcraft	99.47 (47.73)	66.82 (27.16)
CogniFit	116.00 (83.59)	102.65 (60.80)

Table 4. Pretest and posttest scores by group on the PISA Problem Solving Test

	Pretest M (SD)	**Posttest M (SD)**
Warcraft	5.64 (1.35)	5.02 (1.45)
CogniFit	5.17 (1.80)	4.38 (1.67)

Quantifying Improvement in Performance

In order to quantify improvement after the intervention, gain scores were calculated by the following formula, for each instance of the TOH task encountered (i.e. three sessions):

Gain Score = Number of moves $_{POST}$ - Number of moves $_{PRE}$

Gain scores produced from this calculation can be interpreted as follows: negative gain scores indicating improvement (fewer total moves at post-testing), and positive gain scores indicating a decrement in performance (more total moves at post-testing). As a result of incomplete games not producing the number of moves, for some participants, no gain score calculation was possible. At pretesting, the cognitive training group had three missing gain scores for the second TOH and 10 for the third TOH whereas the game training group had one missing gain score for the second TOH and seven for the third TOH. To account for this, when calculating average gain scores for each participant, averages were weighted by the number of completed games (i.e. averaging by the number of incomplete sessions would result in an undefined calculation, as some participants completed all sessions). Table 5 shows the results of an unpaired t-test on the average weighted gain scores found no group differences in TOH gain scores ($p > .05$). Additionally, an unpaired t-test on the average PISA gain scores found no group differences gain scores ($p > .05$).

Sensitivity Analysis

Due to missing data issues discussed above, the final analysis involves a stepwise multiple linear regression (forward and backward; AIC used for final model variable selection conducted using R package MASS, function stepAIC; Venables & Ripley, 2002), predicting max gain score (max of all three potential gain scores) by group membership (WoW or Cognitive Training), total gain score count, and a gain score derived from pre and post measurements on the PISA task (2003). Based on the stepwise regression procedure analysis results in Table 6, the best fitting, significant, multiple regression model was found to be a model predicting max gain score from gain score count (no

Table 5. Problem solving performance compared across training groups

Variable	CogniFit	Warcraft	Test Statistics	*p*
Tower of Hanoi				
Pre-test incomplete sessions	.76 (.75)	.47 (.62)	*t*(32)=1.240	.224
Post-test incomplete sessions	.18 (.39)	.12 (.33)	*t*(32)=-.471	.641
Gain scores	-18.17 (45.74)	-14.15 (22.19)	*t*(32) = .326	.746
PISA Gain scores	-.79 (2.26)	-.62 (1.67)	*t*(32) = .258	.798

Table 6. Stepwise regression model path, analysis of deviance table and the row with the best fitting model, using AIC as criterion, is highlighted in gray

Step	Model	Deviance Residual	*df*	Residual Deviance	AIC
1	Max Gain Score ~ Gain Score Count + Group + PISA Gain Score		30	78913.60	271.491
2	Max Gain Score ~ Gain Score Count + Group	416.527	31	79330.12	269.670
3	Max Gain Score ~ Gain Score Count	2585.619	32	81915.74	268.761

predictor for group membership or PISA gain score; $F(1,32) = 14.41$; $p < .001$; $R^2 = .3104$; adjusted $R^2 = 0.2889$). Participants predicted max gain score is equal to $-111.70 + 48.87$ (Gain Count), where gain score is in the unit of number of moves. Max gain score increased by 48.87 for every one unit increase in gain score count (more gain scores, closer to 0; less improvement after the intervention). Gain score count was a significant predictor of max gain score ($t=3.796$; $p < 0.001$), indicating potential practice effects from repeated exposure to the task. Practice effects will be discussed in subsequent sections.

CONCLUSION

Evidence for Research Question 1

The initial hypothesis regarding the first question was that a brain-training game would help participants improve their rule application component of problem-solving skill better than a roleplaying game after 20 hours of gameplay for several reasons. One reason is that the rules are more explicit during brain-training gameplay and because of claims made by *CogniFit* that brain-training gameplay will improve its users' *brain fitness* or ability to rely on more than one problem-solving strategy. While both games require players to apply rules to solve problems, only *CogniFit* markets its product as a tool that can help users to solve problems in their daily lives (CogniFit, 2019). This claim also suggests that brain-training gameplay can help users transfer skills learned in-game to novel problem-solving scenarios in the natural world. However, the results indicated that there was no significant difference in gain scores (i.e., in Post - Pre Gain scores) in terms of TOH performance (t-test comparing gain scores: $p = .746$) between the two gaming conditions (i.e., *Warcraft* and *CogniFit*), though both groups improved from baseline to post-testing assessment, likely attributable to practice effects (see Figure 5). Overall, the results contradicted our initial hypothesis for Research Question 1; implications are discussed next.

Implications of Results for Research Question 1

Solving problems in an immersive game like *Warcraft* provided players with repeated practice of applying rules and using tools to find creative solutions to similar but varied problems. As players reflected on their choices, they learned how to use the tools by analyzing givens and constraints in unison to achieve maximum character performance and develop optimal solutions to general problems. *CogniFit* players did not experience immersive gameplay, but instead repeated problem-solving scenarios that were varied but required fewer tools and resources to be solved. Once *CogniFit* players knew how to use the paddle and the ball in unison, the only additional resources to use during gameplay were power-ups, bonus bricks, and traps. Roleplaying gameplay required players to solve problems using additional tools and resources efficiently which was a more complex task than using the ball and paddle during brain-training gameplay. Strategizing when and how to apply rules through varied but different problem scenarios with multiple tools and resources through immersive gameplay was beneficial for *Warcraft* participants. Moreover, players in *Warcraft* could receive feedback with help from other players learning when and how to apply tools and resources to solve problems. *CogniFit* players received feedback at the end of each level with an overall score and corrected mistakes through trial and error without additional support.

Evidence for Research Question 2

The initial hypothesis regarding the second question was that training on an immersive, collaborative roleplaying video game for 20 hours would engender transfer of problem-solving skills to novel problem-solving scenarios on the PISA Problem Solving Test better than a brain-training video game. One reason is that research on MMORPGs including *Warcraft* indicates that players co-constructed knowledge by challenging and supporting novel ideas to in-game problem-solving scenarios through

Figure 5. Average number of moves in the Tower of Hanoi task across (up to 3) sessions per person, per timepoint. The left panel represents scores for the CogniFit (COG) group, and the right panel represents scores for the Warcraft (WOW) group.
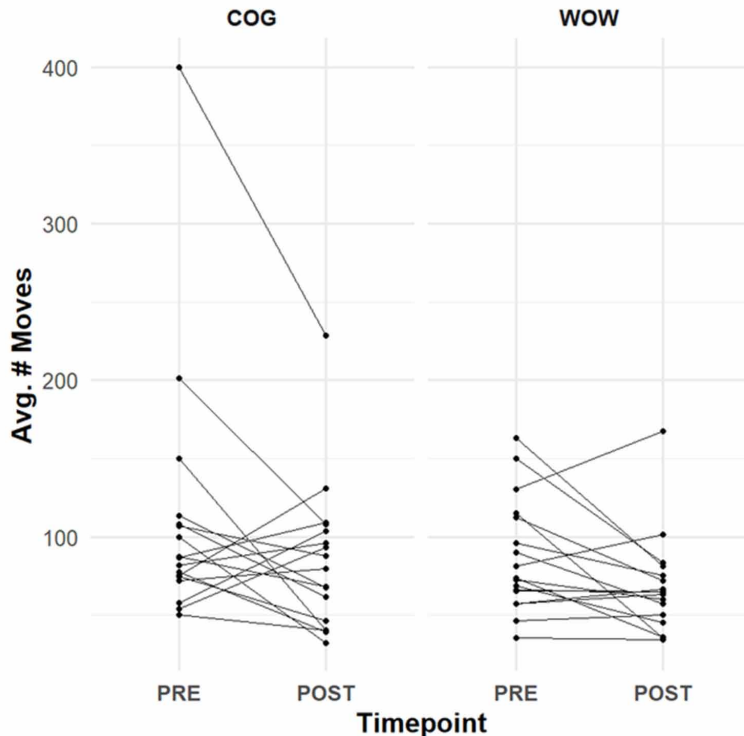
online discussion forums as well as discovering optimal solutions to in-game problems by combining multiple abilities and resources available to players (Chinn & Malhotra, 2002; Steinkuehler & Chmiel, 2006). Efficiently using tools and resources is a component of problem solving and is central to the roleplaying gameplay experience (Shute & Wang, 2016).

However, the results indicated that after 20 hours of gameplay of *Warcraft* or *CogniFit* there was no improved performance on the PISA (i.e., comparing PISA Gain Scores; p = .748). Overall, the mean scores for *Warcraft* participants were slightly better than *CogniFit* participants on the isomorphic versions of the PISA Problem Solving pretest and posttest - indicating baseline differences between the two groups in terms of performance. Overall, there were no significant differences found between roleplaying and brain-training gameplay on transfer of problem-solving skills (see Figure 6). The implications for the results from research question 2 are discussed next.

Implications of Results for Research Question 2

Given that both video game training and "brain-training" did not significantly improve problem-solving skills has several implications. The gameplay behaviors exhibited by players in each condition were aligned with the problem-solving processes on the PISA Problem Solving Test. However, possible reasons for lack of transfer in this study in addition to small sample size include (a) collaborative, immersive roleplaying gameplay may help promote problem-solving skills related to in-game problem solving scenarios but not necessarily to improved performance on external problem-solving assessments, and (b) problem-solving during *Warcraft* gameplay may be too domain specific to transfer to novel problem-solving scenarios on the PISA Problem Solving Test.

The misalignment between the problem-solving domains of *Warcraft* and the PISA Problem Solving Test could have hindered the possibility of finding a transfer effect. As an example, *Warcraft*

Figure 6. PISA Scores before and after the intervention. The left panel represents scores for the COG group, and the right panel represents scores for the WOW group.
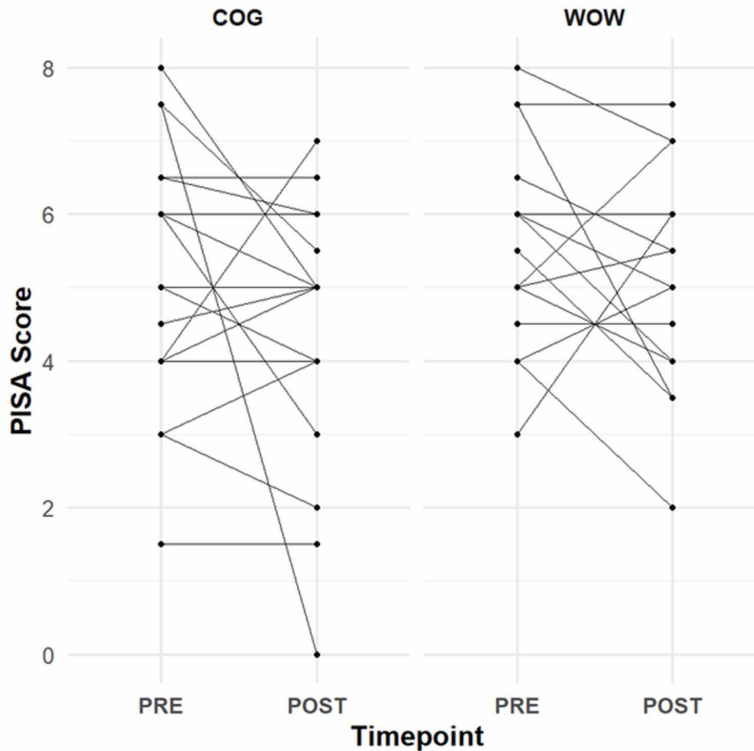
players must learn how to navigate an immersive environment, use complex tools efficiently and effectively to solve problems during gameplay and interact with both the environment and other characters to solve problems. However, solving problems on the PISA Problem Solving Test is not an immersive experience. It was also a solitary activity; participants did not collaborate or interact with each other while taking the test. The OECD designed the PISA Problem Solving Test to cover more general problem-solving skills to complement domain-specific skills (Greiff et al., 2014). Selecting a problem-solving assessment which is embedded within an immersive environment that requires players to engage in collaborative problem-solving processes (i.e. experienced in video gameplay) using tools and resources efficiently could have been a more viable assessment to measure transfer of problem-solving skills in this study. Further research is still warranted to determine if video gameplay can promote transfer of problem-solving skills to novel scenarios. The limitations of this research study are addressed in the next section.

Limitations

Given time and resource constraints, the sample size of this study is small and lacks statistical significance to make claims regarding the general population. With more available resources, recruitment would have likely continued for an additional semester to raise the sample size for the study. Students that did not complete the study cited *time constraints* as the main reason they were unable to fulfill the 20 hours of video gameplay requirement. The optimal time to run the study would have been during Fall and Spring semesters instead of Spring and Summer. In Fall and Spring, more students would have been available for recruitment as well as increased scheduling flexibility and time to complete the intervention during the academic year for the participants. Given that the

authors monitored participants during video gameplay in case any problems arose, there may have been expectancy effects that impacted participants. For example, participants' gameplay experiences may have been negatively or positively affected when being monitored. The potential for participants to alter their behavior simply because they are being studied is known as the Hawthorne Effect (Benedetti, Carlino & Piedimonte, 2016). In addition, the inclusion of a more immersive assessment that measures problem-solving skill transfer could have led to improved outcomes when compared to a more traditional assessment like the PISA Problem-Solving Test (2003).

Future Implications

The main goal of this study was to examine the impact of two distinct types of video gameplay; role playing (*Warcraft*) and brain-training (*CogniFit*) on problem-solving skills for undergraduates. Specifically, if video gameplay can improve the rule application component of problem solving and whether problem solving during gameplay transferred to novel problem-solving scenarios. This study addressed some of the methodological shortcomings found in previous video game training and transfer studies that failed to report recruitment methods, define study variables, and provide an active control group in which participants could expect receive equal improvement from competencies (Baniqued et al., 2013; Boot et al., 2013). As a result, possible placebo effects are likely mitigated in this experiment improving upon methodological pitfalls affecting other video game training studies (Anderson et al., 2010; Ferguson & Kilburn, 2009).

The results from this study suggest that neither a commercially available video game (*Warcraft*) or a commercially available "brain-training" package (*CogniFit*) resulted in improvements in the rule-based component of problem solving (as assessed by the TOH puzzle). Moreover, aside from a lack of improvement in the rule-based component, 20-hours of training did not promote transfer of problem-solving skills to novel scenarios (as assessed by the PISA Problem Solving Task), which is consistent with similar research findings on cognitive training and transfer (Souders et al., 2017). Sensitivity analyses conducted found evidence for practice effects in gain scores, illustrating that rather than improvement due to the training packages, improvement seems related to multiple, closely spaced assessments. Future research can complement this study by increasing the sample size and testing similar immersive well-designed video games on participant knowledge, skills, and abilities, in addition to directly cuing participants to be aware of the strategies (i.e., perceptual and cognitive strategies) they might carry with them from the digital world to the real-world.

ACKNOWLEDGMENT

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. doi:10.1037/a0019277 PMID:20822257

Activision Blizzard. (2019). *World of Warcraft* [Digital Video Game]. Irvine, CA: Blizzard Entertainment, Inc.

Anderson, C. A., Shibuya, A., Ihori, N., Swing, E. L., Bushman, B. J., Sakamoto, A., & Saleem, M. et al. (2010). Violent video game effects on aggression, empathy, and prosocial behavior in Eastern and Western countries. *Psychological Bulletin*, *136*(2), 151–173. doi:10.1037/a0018251 PMID:20192553

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. PMID:24432009

Baniqued, P. L., Lee, H., Voss, M. W., Basak, C., Cosman, J. D., DeSouza, S., & Kramer, A. F. et al. (2013). Selling points: What cognitive abilities are tapped by casual video games? *Acta Psychologica*, *142*(1), 74–86. doi:10.1016/j.actpsy.2012.11.009 PMID:23246789

Benedetti, F., Carlino, E., & Piedimonte, A. (2016). Increasing uncertainty in CNS clinical trials: The role of placebo, nocebo, and Hawthorne effects. *Lancet Neurology*, *15*(7), 736–747. doi:10.1016/S1474-4422(16)00066-1 PMID:27106073

Boot, W. R. (2015). Video games as tools to achieve insight into cognitive processes. *Frontiers in Psychology*, *6*, 3. doi:10.3389/fpsyg.2015.00003 PMID:25653634

Boot, W. R., Blakely, D. P., & Simons, D. J. (2011). Do action video games improve perception and cognition? *Frontiers in Psychology*, *2*, 226. doi:10.3389/fpsyg.2011.00226 PMID:21949513

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. doi:10.1016/j.actpsy.2008.09.005 PMID:18929349

Boot, W. R., Simons, D. J., Stothart, C., & Stutts, C. (2013). The pervasive problem with placebos in psychology: Why active control groups are not sufficient to rule out placebo effects. *Perspectives on Psychological Science*, *8*(4), 445–454. doi:10.1177/1745691613491271 PMID:26173122

Chinn, C. A., & Malhotra, B. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education*, *86*(2), 175–218. doi:10.1002/sce.10001

CogniFit. (2019). *CogniFit* [Digital Video Game]. CogniFit.

Dede, C. (2009). Immersive interfaces for engagement and learning. *Science*, *323*(66), 66–69. doi:10.1126/science.1167311 PMID:19119219

Entertainment Software Association. (2018). *2018 Essential facts about the computer and video game industry*. The Entertainment Software Association. Retrieved from https://www.theesa.com/wp-content/uploads/2019/03/ESA_EssentialFacts_2018.pdf

Eseryel, D., Law, V., Ifenthaler, D., Ge, X., & Miller, R. (2014). An investigation of the interrelationships between motivation, engagement, and complex problem solving in game-based learning. *Journal of Educational Technology & Society*, *17*(1), 42–53.

Ferguson, C. J., & Kilburn, J. (2009). The public health risks of media violence: A meta-analytic review. *The Journal of Pediatrics*, *154*(5), 759–763. doi:10.1016/j.jpeds.2008.11.033 PMID:19230901

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.

Green, C. S., & Bavelier, D. (2012). Learning, attentional control, and action video games. *Current Biology*, *22*(6), 197–206. doi:10.1016/j.cub.2012.02.012 PMID:22440805

Greiff, S., Wüstenberg, S., Csapo, B., Demetriou, A., Hautamäki, J., Graesser, A. C., & Martin, R. (2014). Domain-general problem solving skills and education in the 21st century. *Educational Research Review*, *13*, 74–83. doi:10.1016/j.edurev.2014.10.002

Hart Research Associates. (2018). *Falling short? College learning and career success*. Washington, DC: Association of American Colleges and Universities.

Hung, W., & Van Eck, R. (2010). Aligning problem solving and gameplay: A model for future research and design. In R. van Eck (Ed.), *Interdisciplinary models and tools for serious games: Emerging concepts and future directions* (pp. 227–263). Hershey, PA: IGI Global. doi:10.4018/978-1-61520-719-0.ch010

Huyck, C. R., & Kreivenas, D. (2018). Implementing Rules with Artificial Neurons. In M. Bramer & M. Petridis (Eds.), Lecture Notes in Computer Science: Vol. 11311. *Artificial Intelligence XXXV. SGAI 2018* (pp. 21–33). Cham: Springer. doi:10.1007/978-3-030-04191-5_2

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, NY: Springer. doi:10.1007/978-1-4614-3546-4_1

Lieberman, D. A., Biely, E., Thai, C. L., & Peinado, S. (2014). Transfer of Learning from Video Game Play to the Classroom. In *Learning by Playing* (pp. 189–204). Academic Press. doi:10.1093/acprof:oso bl/9780199896646.003.0013

Mayer, R., & Wittrock, M. (2006). Problem solving. In P. Alexander & P. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 287–303). Mahwah, NJ: Erlbaum Publishers.

OECD. (2003). *PISA 2003 Technical Report*. Paris: OECD Publishing.

Oei, A. C., & Patterson, M. D. (2014). Playing a puzzle video game with changing requirements improves executive functions. *Computers in Human Behavior*, *37*, 216–228. doi:10.1016/j.chb.2014.04.046

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.

Schiff, R., & Vakil, E. (2015). Age differences in cognitive skill learning, retention and transfer: The case of the Tower of Hanoi Puzzle. *Learning and Individual Differences*, *39*, 164–171. doi:10.1016/j.lindif.2015.03.010

Shute, V. J., Ke, F., Almond, R. G., Rahimi, S., Smith, G., & Lu, X. (2019). How to increase learning while not decreasing the fun in educational games. In R. Feldman (Ed.), *Learning Science: Theory, Research, and Practice* (pp. 327–357). New York, NY: McGraw Hill.

Shute, V. J., Ventura, M., & Ke, F. (2015). The power of play: The effects of Portal 2 and Lumosity on cognitive and noncognitive skills. *Computers & Education*, *80*, 58–67. doi:10.1016/j.compedu.2014.08.013

Shute, V. J., & Wang, L. (2015). Measuring problem solving skills in Portal 2. To appear. In P. Isaias, J. M. Spector, D. Ifenthaler, & D. G. Sampson (Eds.), *E-learning systems, environments and approaches: Theory and implementation*. New York, NY: Springer.

Shute, V. J., & Wang, L. (2016). Assessing and supporting hard-to-measure constructs. To appear A. Rupp, & J. Leighton (Eds.), Handbook of cognition and assessment. New York, NY: Springer.

Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z., & Stine-Morrow, E. A. (2016). Do "brain-training" programs work? *Psychological Science in the Public Interest*, *17*(3), 103–186. doi:10.1177/1529100616661983 PMID:27697851

Souders, D. J., Boot, W. R., Blocker, K., Vitale, T., Roque, N. A., & Charness, N. (2017). Evidence for Narrow Transfer after Short-Term Cognitive Training in Older Adults. *Frontiers in Aging Neuroscience*, *9*, 41. doi:10.3389/fnagi.2017.00041 PMID:28293188

Spires, H. A., Rowe, J. P., Mott, B. W., & Lester, J. C. (2011). Problem solving and game-based learning: Effects of middle grade students' hypothesis testing strategies on learning outcomes. *Journal of Educational Computing Research*, *44*(4), 453–472. doi:10.2190/EC.44.4.e

Steinkuehler, C., & Chmiel, M. (2006). Fostering scientific habits of mind in the context of online play. In S.A. Barab, K.E. Hay, N.B. Songer, & D.T. Hickey (Eds.), *Proceedings of the International Conference of the Learning Sciences* (pp 723-729). Mahwah NJ: Erlbaum.

Stiller, K. D., & Schworm, S. (2019). Game-Based Learning of the Structure and Functioning of Body Cells in a Foreign Language: Effects on Motivation, Cognitive Load, and Performance. *Frontiers in Education*, *4*, 18. doi:10.3389/feduc.2019.00018

Tower of Hanoi. (2019). Play Tower of Hanoi. The math is fun. Retrieved from https://www.mathsisfun.com/games/towerofhanoi.html

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

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics* (4th ed.). Academic Press; doi:10.1007/978-0-387-21706-2

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. doi:10.3389/fpsyg.2013.00852 PMID:24379790

Westera, W. (2019). Why and how serious games can become far more effective: Accommodating productive learning experiences, learner motivation and the monitoring of learning gains. *Education Technology & Society*, *22*(1), 59–69.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag.

Wickham, H., François, R., Henry, L., & Müller, K. (2019). dplyr: A Grammar of Data Manipulation. R package version 0.8.3. Retrieved from https://CRAN.R-project.org/package=dplyr

Benjamin Emihovich is an Assistant Professor of Educational Technology in the Education Department at the University of Michigan-Flint and is the program faculty coordinator for the online Educational Technology (M.A.) program. He currently teaches undergraduate and graduate students in the areas of Instructional Design and Technology as well as curriculum and instruction. His research area focuses on the following; game-based learning, assessments for learning in immersive environments, and emerging learning technologies.

Nelson A. Roque is a NIA T32 Postdoctoral Fellow, at Penn State's Center for Healthy Aging. Nelson earned his Ph.D. in Cognitive Psychology from Florida State University in 2018. Nelson has a strong background in visual attention, focusing on how to reliably measure it, how it relates to individual difference factors (e.g., age, sleep) and translating insights from theoretical work in visual attention to applied contexts (e.g. medication errors).

Justin Mason is a Postdoctoral Associate in Rehabilitation Science at the University of Florida. His research interests include interventions suitable for mitigating age-related cognitive and physical decline in older adults. Additionally, he's interested in factors that influence older adults' adoption and acceptance of emerging technologies.