

GAMIFICATION IN PRESENTATION AND TESTING AND THEIR EFFECT ON PERFORMANCE

Alan Franklin, Southern Illinois University, alanfranklin@siu.edu
Eric Jacobs, Southern Illinois University, eajacobs@siu.edu

ABSTRACT

The purpose of this study was to ascertain whether presentation and testing methods, conducted in traditional and gamified manners, would affect the performance on standard or gamified tests for college students. This study was conducted to examine if motivation and engagement derived from gamified conditions would lead to better performance on tests. Both gamified presentation and gamified testing was used to see if using similar testing methods would affect performance on the tests. This study was conducted using a participant pool of 100 students from a public Midwestern university, though 9 were dropped from analyses. Participants went through a traditional presentation or a game-based presentation followed by a traditional or game-based test to examine their performance. An MANCOVA was run on the data collected, and while the overall model was not significant, presentation and testing methods were found to affect participants first attempts on a test, as well as the covariate of Cognitive Engagement was found to have an effect on the highest score participants would earn in up to five attempts on a test. While participants did not significantly perform better in the gamified groups, they did express greater interest in the game-based presentations and tests, as well as in the use of gamification in classrooms.

Keywords: Gamification

INTRODUCTION

The use of gamification within classrooms is enjoyed by students and used to increase student engagement and motivation. Defined by Deterding et al. (2011), gamification is the practice of taking concepts and ideas of video games and other game-based formats and implementing them in non-gaming areas. Gamification has been applied in a variety of different courses to engage students with the material, improve their motivation to learn and to improve their understanding and retention of the course material.

Multiple researchers have examined classrooms in which gamification has been implemented. One such example was Leaning (2015), who had examined a class on media theory that had been modified to include games and other activities. The researcher compared students' scores in the course with their scores in an advanced theory class that was taught traditionally. Another comparison was made against the prior year of students' scores in both courses, both of which were also taught traditionally. The gamified class performed better compared to the three other classes, however, the means of these scores were found to be not significant by an analysis of variance. Students from the gamified class later participated in focus groups to discuss their experience with the course. The students reported feeling that they performed better and retained the information of the course better compared to their traditional courses. While the analysis did not reflect the students' sentiments, it is important to consider that the students reported their preference for the gamified course.

Research on gamification has been conducted, not only in college courses, but in elementary school classes as well. Kingsley & Grabner-Hagen (2015) observed elementary school classrooms and implemented gamifications techniques along with iPads provided to the students. The iPads included apps developed to teach a science course. Students from one teacher's classes were observed and surveyed about their experience with technology used within the class as well as gamification within the class. In the survey the students reported that they preferred the gamified class compared to traditional classes. The majority of the class had marked that they preferred classes that involved technology, as well as felt that the design of the course as well as the tasks made it easier to learn the subjects that were covered. This design that was implemented in this study is of interest as they included a variety of aspects of video game progression for the students to see how they were doing in the course. This included the use of 'Mini-

Bosses' as quizzes and 'Bosses' as tests, that the student was to defeat in order to progress. Framing coursework from a game-oriented perspective may be a method that can increase engagement and motivation within students.

A major aspect of learning that classrooms and research have examined and attempted to implement prior has been motivation, whether through external rewards, alternative teaching methods, or through gamification. Cicchino (2015) attempted to use game-based learning methods in aiding students in learning critical reasoning skills and content retention in 8th grade social studies classrooms. In comparing videotapes of the experimental classes with the control classes, it was rated that the game-based learning classes, which held higher motivation, showed higher levels of critical thinking and provided more opportunities for guided reflection for the students.

The effectiveness of gamification has been called into question and evaluated in more recent studies. Alshammari (2020) conducted an experiment with elementary students learning Arabic to evaluate the effectiveness of gamification on learning outcomes and student motivation. The students in the experimental condition were taught using a gamified e-learning system, while the control group was taught using traditional methods. The students in the experimental condition performed better on learning outcomes as well as more motivated compared to students in the control condition.

Gamification has been implemented in course assignments as well, as seen in the study conducted by Armier, Shepherd, and Skrabut (2016). The study did not apply the gamification to any part of the graded course work and instead examined the effects of voluntary gameplay's influence on participation and engagement. The participants were students from a technology integration course that taught preservice teachers. The assignment where gamification was implemented was a thirty-minute presentation on a topic from the course material. The gamified sections were provided the opportunity to complete additional activities either individually or with a group to earn points that could be redeemed at certain times for rewards. Some rewards were related to the project and allowed students to either increase or decrease their allotted presentation time, wear casual clothing, or reduce the minimum number of slides needed for the project. However, not all rewards were related to the project, such as being allowed to display unrelated objects such as soup cans or actions figures, and rewards were allowed to be traded between groups. The tasks that were completed as gamified objectives were additional suggested tasks that the instructors believed students had not been completing in the standard course. Through the surveys the study found that the gamified sections had groups spent more time on project and reported spending more effort on the project. The students in gamified groups also reported feeling better prepared for the project, completing more tasks and having spent more time with their groups. The student's preference for gamified versions of courses and assignments is apparent across prior research, and this preference can be applied to other class formats as well.

Recently there has been a push to create more courses that are taught exclusively online, and with the increase in use of gamification, these courses require guidelines of best practice for these online courses. In a study conducted by Chang and Wei (2015) a focus group was held with twenty-five people who frequently participate in massive open online courses (MOOCs) in order to identify forty gamification mechanics used frequently in MOOCs. The researchers also provided a survey to students who had taken MOOCs and identified ten of the items to account for more the 50% of the self-reported feeling of engagement. This push toward implementing gamification in courses to boost engagement and motivation is an effort to improve student's ability to learn material as well as complete courses. Danka (2020) discussed connectivist massive open online courses (cMOOCs) and noted that completion rates in these courses are low, despite massive initial enrollment. Danka argued that aspects of massively multiplayer online role-playing games (MMORPGs) can be incorporated into cMOOCs to boost motivation of students if formative assessment does not conflict with connectivist principles. MOOCs implementing gamification differs from implementation within a standard classroom, as the two course types have different services they provide to the students. What best serves a standard classroom will differ from what best serves MOOCs, and should be taken into consideration.

What components of a gamified classroom benefit students has been researched as well. Chapman and Rich (2018) surveyed students in a gamified course to evaluate components of a gamified course and their impact on the motivation of the students. Over two-thirds of their participants reported that they felt gamified courses were more motivating compared to traditional courses. The students rated all the components of a gamified course favorably, with most students indicating that each component improved their motivation or did not affect it. Similarly, Sailer and Homner (2019) conducted a meta-analysis of research findings to examine the effects of gamification on

cognitive, motivational and behavioral learning outcomes. In their analysis gamification was found to be moderated by game fiction and social interaction for behavioral outcomes. They also noted effects from combining competition with collaboration for motivational learning outcomes.

Gamification is used in the hopes that it will help improve retention of information as well as provide an alternative to standard teaching methods, it is important to understand how people learn. Studies have attempted to understand how people learn and what methods they use, by having the participants play games and see how they either learn the rules of the game presented to them or what methods they use to optimize their performance.

Research for Gamification has been typically done for courses where gamification was a supplemental and optional tool rather than the primary teaching method, meaning that the students who may perform well participated in the additional activities, while students that could have benefitted from these techniques may have opted out. It is useful to examine the effects of gamification as a primary teaching method for a topic compared to a standard method. This study aimed to address the following research questions:

1. Will the inclusion of gamification in the teaching process significantly improve performance on tests?
2. Will the use of gamification in the testing process significantly improve performance on tests?
3. Will performance on tests be better if both the teaching and testing process are conducted with similar methods?

The difference between these two teaching methods were examined in this study through a comparison of a game presenting material against presenting material through a standard PowerPoint presentation. In addition, the participants would take the posttest through either a standard format or through a gamified test format. The researchers hypothesized that participants who were presented information in a gamified method would be more engaged and perform better on either posttest compared to those presented the material in a standard method. In addition, the researchers hypothesize that participants that are tested on the material with a gamified method will be more engaged and will perform better as well compared to those tested in a standard method.

RESEARCH METHODOLOGY

The study's participants were 100 college students from a Midwestern university enrolled in an introduction to psychology course. These students were randomly assigned between four groups in a 2x2 experimental design—(1) Gamified Teaching and Gamified Testing, (2) Traditional Teaching and Traditional Testing, (3) Gamified Teaching and Traditional Testing, and (4) Traditional Teaching and Gamified Testing. All groups were expected to learn basic facts from the Psychology of Learning. Specifically, the material covered basic schedules of reinforcement and their effects on behavior. Schedules of reinforcement arrange for rewarding consequences to follow behavior when specific conditions have been met. Each schedule of reinforcement has specific effects on how responses are distributed through time. Students were expected to learn the defining features of each of four schedules and the typical response pattern generated by each schedule. Two groups were presented information through a PowerPoint presentation, while the other groups were presented the information through a video game. After reading through the presentation or playing the game, the groups were provided a test in either a traditional manner, through Qualtrics, or through a gamified version of the test. This population was chosen as it was an easily accessible population and, as students, the participants were able to provide feedback comparing the teaching methods used in the experiment with the methods they see used in a standard classroom. The participants chose to participate in the study as per the requirements of their introduction to psychology course. On average it took the participants approximately 75 minutes to complete their session.

The participants met in a small computer lab. As the participants were randomly assigned to four separate groups. The data were collected through Qualtrics, as well as save files from the game, and materials that were presented through multiple modes. A PowerPoint Presentation was prepared, covering the topic of schedules of reinforcement. Stimuli for the gamification group was presented using individual personal computers running software coded by a researcher in, a game creation program.

For Traditional Teaching conditions, a PowerPoint presentation was designed to present bullet points and definitions related to the topic of schedules of reinforcement. The material covered also went over material that was asked on

the final test of the study, with examples given of some of the points. The participants reviewed the slides at their own pace but were asked to be thorough in their reading. The presentation did not have a theme or audio to avoid making the presentation too complex or distracting from the information.

For Gamified Teaching Conditions, the material was presented in the context of a role-playing game coded in RPG Maker VX Ace. The game was designed to have the information related to the topic presented to the player through puzzles and conversations with various characters in a village. The starting location was designed as a town within the game where townspeople would discuss the material with the player and present various topics. The townspeople were grouped together by topic, along with the player starting at the far end of the town. The player was tasked with speaking to everyone before they could end the game and move onto the test.

During either the PowerPoint or the game the participant was encouraged to take notes to study from once they were finished with the material. Once the participant had read or played through the material, the participant was allowed to review their notes before attempting the test. After studying, the participant's notes were collected, and they were presented a short test with fifteen questions based on the topic that was presented or played through. The participants' responses were recorded and their score on the test would not affect their grade.

For Traditional Testing Conditions, learning the material was assessed through fifteen question multiple choice test that was developed in Qualtrics. To test for mastery of the material, the students were required to take the test multiple times until they achieved a perfect score on the test, or if they failed to answer all the questions correctly after five tries they were moved onto the final phase. All responses for the multiple attempts were recorded to see if the students were improving over time or if they scored similarly throughout the multiple attempts.

For the Gamified Testing conditions, participants had to complete a maze within the role-playing game where they had to answer questions on the material correctly to progress through the maze. The questions were the same as those on the traditional posttest. After each answer, the participant was informed if they were correct or incorrect, unlike the traditional posttest. Following correct answers, participants were told which direction to travel in the maze (i.e., north, south, east, or west). If the participant answered any of the questions incorrectly, at the end of the maze the participant was sent back to the beginning. If the participant was able to answer all questions correctly, they were sent to a final room and informed that they finished the test.

After the participants had finished one of the tests, they were then presented a questionnaire with items from the Student Engagement Instrument (SEI) to collect scores for cognitive engagement. The Student Engagement Inventory consists of 35 questions on a four-point scale. While the instrument has thirty-five questions only questions that measure the construct of Cognitive Engagement were used for this study. The SEI can be obtained for free from the University of Minnesota, Check and Connect Program. The SEI is normally used to measure student engagement levels in learning and focuses on engagement beyond the classroom as well.

The average amount of time that each participant needed for the experiment was approximately 75 minutes. Time was allotted to reading the consent form as well as signing onto a computer to have access to Qualtrics. The participants typically spent thirty minutes reading through the PowerPoint or playing the game, as well as taking notes, though they were able to move on to the next step once they had finished. The participants needed approximately a half hour to complete the multiple attempts on the tests. The participants were also asked to complete the SEI and a brief demographics section following the testing attempts. Once the participants had finished with the demographics, they were provided a debriefing form and informed of their completion of the study.

Notes were collected from the participants and analyzed by two blind reviewers who read through the notes to see if participants wrote unnecessary information, used arrows to guide themselves through the maze rather than answer the questions presented in the videogame test, wrote questions down as well as answers, wrote fake notes or admitted to intentionally disrupting the study, or left their notes blank. The reviewers compared their analyses of the notes and would discuss any difference in opinion, however no major differences between the reviewers were found.

Twenty-five participants were recruited for the Traditional Teaching and Traditional Testing Group; however, one was removed from analysis as the participant left early and the data set was incomplete. Participants in the Traditional Teaching and Gamified Testing Group had the material presented through PowerPoint and were tested

with questions presented through the videogame test. Again, 25 participants were initially assigned to this group, but five were excluded. Review of the notes of four of the excluded participants revealed they had solved the maze by making note of the correct direction to travel at each branch, rather than noting the facts they were required to learn. One participant was removed as he indicated a desire to undermine the research in his notes. Thus, data from 20

participants were included for the Traditional Teaching and Gamified Testing group. Participants in the Gamified Teaching and Standard Testing group had the material presented through the videogame and had been tested with multiple choice questions presented through Qualtrics. There were 25 participants initially within this group, but one was removed from analysis due to not completing all five attempts at the test or achieving mastery, for a total of 24. Participants in the Gamified Teaching and Gamified Testing group had the material presented through the videogame and were tested with the videogame test as well. This group initially had 25 participants, but two were removed from analysis due to using arrows to guide themselves through the maze rather than learning from the material.

RESULTS

To assess if presenting material in the context of a video game would increase engagement, which in turn should increase performance, analysis of the results focused on three primary dependent measures---engagement as measured by the SEI, performance during the first attempt on the posttest, and highest score achieved on the posttest. We initially intended to use the number of attempts at the posttest required to achieve mastery of the material (i.e., a perfect score) as a key dependent measure, but few participants achieved mastery by this criterion. Thus, the highest score earned on the posttest was used as a dependent measure instead of the number of attempts to mastery. Three ANOVAs, seen in Table 1, were conducted to assess if the different training and testing conditions affected these three key dependent measures. In addition, a MANCOVA, seen in Table 2 and Table 3, was conducted to-examine how testing method, presentation method, and the interaction between the two affected two dependent measures, while note taking, engagement and intrinsic motivation were held as covariates.

Skewness and kurtosis of the data was checked for the three ANOVAs through SPSS. The distribution of highest score had skewness and was leptokurtic in the Traditional Teaching and Traditional Testing group, but for the other three groups, as well as the overall analysis of the variable, skew and kurtosis were within acceptable ranges. To test for homogeneity of the variances, a Levene’s test was run on the ANOVAs. The Levene’s test was used to help identify outliers and was significant only for the Highest Score on the test, indicating that the variances between the groups were not equal for that variable. However, because the sample sizes were roughly equivalent, ranging from twenty to twenty-four, running the analyses was still possible (Hair et al., 2010, p. 365).

Table 1. ANOVAs of First Score, Highest Score, and Engagement by Group

		Sum of Squares	df	Mean Square	F	Sig
Cognitive Engagement	Between Groups	116.986	3	38.995	.988	.402
	Within Groups	3432.201	87	39.451		
	Total	3549.187	90			
First Score	Between Groups	63.797	3	21.266	2.916	.038
	Within Groups	678.265	93	7.293		
	Total	742.062	96			
Highest Score	Between Groups	19.800	3	6.600	1.189	.319
	Within Groups	483.101	87	5.553		
	Total	502.901	90			

Table 2. MANCOVA of Test Performance

	Effect	Value	F	Hypothesis df	Error df	Sig.
Intercept	Pillai's Trace	.625	43.293b	3	78	.000
	Wilks' Lambda	.375	43.293b	3	78	.000
	Hotelling's Trace	1.665	43.293b	3	78	.000
	Roy's Largest Root	1.665	43.293b	3	78	.000
IM	Pillai's Trace	.080	2.272 ^b	3	78	.087
	Wilks' Lambda	.920	2.272 ^b	3	78	.087
	Hotelling's Trace	.087	2.272 ^b	3	78	.087
	Roy's Largest Root	.087	2.272 ^b	3	78	.087
Word Count	Pillai's Trace	.016	.423b	3	78	.737
	Wilks' Lambda	.984	.423b	3	78	.737
	Hotelling's Trace	.016	.423b	3	78	.737
	Roy's Largest Root	.016	.423b	3	78	.737
Cognitive Engagement	Pillai's Trace	.037	1.010b	3	78	.393
	Wilks' Lambda	.963	1.010b	3	78	.393
	Hotelling's Trace	.039	1.010b	3	78	.393
	Roy's Largest Root	.039	1.010b	3	78	.393
Presentation	Pillai's Trace	.080	2.274 ^b	3	78	.087
	Wilks' Lambda	.920	2.274 ^b	3	78	.087
	Hotelling's Trace	.087	2.274 ^b	3	78	.087
	Roy's Largest Root	.087	2.274 ^b	3	78	.087
Testing	Pillai's Trace	.162	5.021b	3	78	.003
	Wilks' Lambda	.838	5.021b	3	78	.003
	Hotelling's Trace	.193	5.021b	3	78	.003
	Roy's Largest Root	.193	5.021b	3	78	.003
Presentation * Testing	Pillai's Trace	.033	.894b	3	78	.448
	Wilks' Lambda	.967	.894b	3	78	.448
	Hotelling's Trace	.034	.894b	3	78	.448
	Roy's Largest Root	.034	.894b	3	78	.448
a. Design: Intercept + CE Score + Word Count + Group						
b. Exact statistic						
c. The statistic is an upper bound on F that yields a lower bound on the significance level.						

Table 3. Between-Subjects Effects for MANCOVA of Test Performance

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Observed Power
Corrected Model	First Score	101.558a	6	16.926	2.375	.037	.785
	Highest Score	69.108b	6	11.518	2.241	.048	.756
Intercept	First Score	261.704	1	261.704	36.724	.000	1.000
	Highest Score	543.789	1	543.789	105.786	.000	1.000
Intrinsic Motivation	First Score	29.375	1	29.375	4.122	.046	.518
	Highest Score	26.051	1	26.051	5.068	.027	.604
Cognitive Engagement	First Score	4.239	1	4.239	.595	.443	.051
	Highest Score	11.886	1	11.886	2.312	.132	.151
Word Count	First Score	.044	1	.044	.006	.937	.119
	Highest Score	4.466	1	4.466	.869	.354	.324
Presentation	First Score	8.541	1	8.541	1.198	.277	.191
	Highest Score	6.966	1	6.966	1.355	.248	.210
Testing	First Score	39.355	1	39.355	5.523	.021	.641
	Highest Score	.104	1	.104	.020	.887	.052
Presentation * Testing	First Score	5.704E-6	1	5.704E-6	.000	.999	.050
	Highest Score	6.797	1	6.797	1.322	.254	.206
Error	First Score	570.098	80	7.126			
	Highest Score	411.237	80	5.140			
Total	First Score	5365.000	87				
	Highest Score	10745.000	87				
Corrected Total	First Score	671.655	86				
	Highest Score	480.345	86				
a. R Squared = .151 (Adjusted R Squared = .088)							
b. R Squared = .144 (Adjusted R Squared = .080)							
c. Computed using alpha = .05							

The first ANOVA was conducted on the engagement scores from the SEI for each of the groups to assess if the different training and testing groups had a significant effect on engagement. Nine participants were removed for this analysis, as they had either not completed the study or they “cheated” during the study by writing answers or similar topics in their notes during the testing phase. In other words, these latter participants failed to learn the material during the training phase but recorded the feedback of the gamified testing groups and used that information to better their performance on their latter attempts on the posttest. These nine participants will be referred to as the “Cheaters”. This left 91 participants across the four groups for analysis. Table 1 shows that engagement was not differentially affected by the training and testing groups. These findings demonstrate that the form of gamification used in this study did not produce significantly different levels of engagement than reviewing the PowerPoint presentation or the Qualtrics version of the test.

The second ANOVA was run on the first score earned on the test by each participant. The first score was examined to check if any of the four groups produced better performance on the first attempt on the test, as students would only have one opportunity in a traditional classroom. For this analysis, the Cheaters were included as they would not have been able to use their notes to cheat on the first attempt and would have been writing the notes in order to cheat on future attempts. Three participants were still excluded as one had written within his notes intent to undermine the study, while the other two had failed to complete the study. The participants in Gamified Teaching and Traditional Testing group performed significantly better than those in Traditional Teaching and Gamified Testing group. To check if these results held, a Bonferroni correction was conducted to reduce the chance of a Type I error, with the initial alpha level set at .05 and with 3 groups, resulting in the individual hypotheses’ alpha to be approximately .0167. After running the Bonferroni correction, the significance of the difference between these two groups held.

The third ANOVA was run on the highest score from each participant across all groups. For this analysis, the Cheaters were removed as they may have used their notes to artificially inflate their scores. There were no significant differences between the different testing or presentation groups on cognitive engagement. As engagement was not skewed or kurtotic, it can be reasoned that the study had failed to place engagement under control.

SUMMARY

The gamified groups were expected to report higher motivation and engagement compared to the PowerPoint groups. In turn, participants who reported higher motivation and engagement were expected to perform better on the posttest. Finally, maintaining consistency across the training and testing modalities was hypothesized to improve performance. That is, those participants who received the material through PowerPoint were expected to perform better on the traditional test and those who received the material through the game were expected to perform better on the game-based test.

In the present study, we attempted to assess if gamification would increase engagement in students learning basic facts about schedules of reinforcement and, in turn, if higher levels of engagement would improve retention of those facts. However, there was little evidence to suggest that motivation was affected by gamification and, in turn, we are unable to determine if engagement is related to performance on the tests. Although intrinsic motivation was included in the MANCOVA, the SEI only asks two questions related to the construct of intrinsic motivation in a very general way. Future research should use instruments that measure intrinsic motivation that is specific to the task at hand to assess the relationship between motivation to perform well on the task and actual performance on the task in relation to gamification.

One aspect of the study that should be addressed was that six of the participants within the groups that used gamified testing, either attempted to or did complete the game by using arrows or directions to guide them through the maze rather than remember the material and apply it. The main issue with participants “cheating” during the test is that the goal of the study was for the participants to learn and demonstrate their understanding of the material, therefore cheating on the tests remove the ability to see how well they understood the material. The fact that the cheaters were from the gamified testing groups is important as none of the participants in the standard testing groups attempted to cheat by writing questions or marking which answers they thought were correct. This difference may be due to the gamified groups of the test providing feedback to the participants informing them immediately if they had answered the question correctly or not, whereas the Qualtrics version of the test did not provide immediate feedback to the students, aside from if they scored perfectly or not. If this study were to be conducted again, feedback on the participants’ answers should be the same across all versions of the test.

Another shortcoming of the game used in the study was that some of the cheating participants would not read the questions and would guess which path lead to the correct answer. Making serial guesses until the correct answer was identified and noting the correct directions may have been viewed as easier or more time efficient than learning the material. The cheaters brute forcing their way through the test raises an important consideration for using gamified methods. Some people are likely to try to defeat the intended purpose of the game even for minor benefits, rather than use the system for its intended purpose. The only incentive the six participants who engaged in this behavior received for doing this was that they might be allowed to leave the study early if they obtained a perfect score. It should be noted, however, that only two of the participants who used this strategy achieved a perfect score and one of those two did so on the last possible attempt. Thus, the mere possibility of completing the task in a shorter time or with less effort was sufficient for these participants to employ this alternative strategy rather than simply try to learn the facts. A potential solution to this problem is to punish guessing by adding in a timeout following incorrect answers within the test. For example, transporting participants to a “holding cell” for a brief, but functionally aversive, time (e.g., 90 s) may encourage the participants to actually learn the material because doing so would be the most time efficient way to escape the task.

Using Qualtrics to present the standard testing group also presented some methodological problems. When examining how the participants responded during the study, it was difficult to examine how the participants responded in the Qualtrics version of the test. This was due to the participants in the Qualtrics version of the test being able to quickly choose random answers in the multiple-choice test, meaning that the researcher was unable to observe if they were reading the question or not. During the time of data collection Qualtrics did not support

measurement of reaction time. Thus, alternative research design packages may prove to be more versatile for future research. Comparatively, performances in the gamified test groups could be more easily observed by the researcher to assess if participants were running past the sign with the question and answers or were running in a systematic or random direction within the game. However, the participants run through the gamified test could not be observed afterward, but this could be rectified through screen recording software. Future researchers should set checks in the test, to ensure that the participants are actively reading the questions for each attempt on the test.

Whereas some participants attempted to cheat or just go through the motions until their participation in the experiment was deemed complete, other participants displayed various levels of observable motivation to do well on the test. Some of the participants displayed signs of anxiety or frustration when informed they did not achieve a perfect score on the test, while those that did achieve mastery showed pride that they were able to understand what they perceived as a complicated topic. Of interest is that though the participants had been informed that the study would not affect their grade in the course, a portion of the participants still wanted to do as well as they could on the test. The desire to perform well was also demonstrated through participants from both testing groups requesting to know how well they performed on the test, when they had failed to achieve a perfect score. Feedback on each question was available only for participants in the videogame test groups, whereas participants in the standard testing groups were only informed if they mastered the material or not upon completing each attempt. Comparatively other participants showed little interest in the study and would attempt to finish the tests quickly, either by choosing random answers in the Qualtrics version, or by running in random directions without reading the question in the videogame version. Although the participants were asked to do their best, their performance may be related to their own intrinsic motivation, meaning they were motivated by their own desire to perform well. Future research should consider comparing these types of students against each other or ask participants to provide feedback on what they had thought of the study and compare those that enjoyed the study against those that did not. Another potential solution to motivate participants would be to offer extra credit based on their performance on the post test, or implement the presentation and test into a course, with performance on the posttest added into the overall grade of the course. Adding a grade that affects the participants' course grades may motivate them to take the presentation and test more seriously.

Despite the many methodological difficulties in the present study, there were a couple of performance differences that warrant discussion. The most straight forward assessment of the effects of the different presentation methods is to compare performances on the first attempt of the test, as the first score is unaffected by the feedback inherent in the testing groups. Presenting the material in a videogame increased performance on the first attempt on the posttest when performance was assessed with multiple choice questions presented through Qualtrics relative to other training and testing groups. The average score for those in Gamified Teaching and Traditional Testing group on the first attempt was 8.25 out of a maximum of 15, compared to the Traditional Teaching and Traditional Testing group, which had an average score of 7.75. When examining those that were tested with the gamified test, the Gamified Teaching and Gamified Testing group had an average score of 6.96 on the first attempt, whereas the Traditional Teaching and Gamified Testing group's average score on the first attempt was 6.05. The increase in performance when the material is presented in the game context is similar to performance improvements reported by Tews, Jackson, Ramsay & Michel (2014). In their study, students had increased student engagement when they were presented the material through humor, storytelling or creative examples as compared to activities that used hands on exercises or promoted social involvement. The use of gamification may similarly increase engagement but do so in a manner that frees a teacher to attend to other aspects of instruction, such as providing one on one attention to struggling students. To examine how a teacher may affect the use of gamification it would be wise to research how students performed on a test when presented the material through different methods by the same teacher.

Turning attention to the highest score achieved across all posttest attempts, the standard presentation groups may have resulted in slightly better performances, at least descriptively. In comparison, the Traditional Teaching and Gamified Testing group performed best, with the average highest score of 11.65, whereas the Gamified Teaching and Gamified Testing group performed worst with an average highest score of 10.35. As for the Traditional Teaching and Traditional Testing group, the average highest score was 11.00, whereas the Gamified Teaching and Traditional Testing group had an average highest score of 10.67. Although there may be many reasons for the performance differences, one potential difference across the presentations that may have contributed to these differences was that the PowerPoint presentation presented the material in a systematic order, whereas the game did not. The gamified presentation used an "open world" approach where participants gained information as they

wandered through the village, with little direction. Performance in the gamified groups may have been better if the participants' movements were constrained such that they had to encounter the material in a fixed order.

The participants' notes suggest that this may be the case. In the slide presentation groups, the notes would typically be the information alone presented in the same order that was presented in the presentation. In comparison, the gamified presentation notes tended to include numbering or descriptions of the villagers within the game as participants would use the descriptions or a count of the villagers to determine if they had spoken with everyone. Another difference in the notes was that the gamified notes would have the information in the order that the participant saw the material as the participant could learn the information in an idiosyncratic order. Future research should provide more direction within the gamified presentation to allow for consistent presentation of the material. Providing a stimulus (i.e., a unique font or color) to indicate which information is noteworthy and may promote the students to record only relevant information. Doing so may decrease the distractions generated by the gamified presentation groups by directing attention to the course material instead of trivial features of the game.

Future research on the use of gamified teaching methods should consider conducting the research in a standard classroom with typical demand characteristics associated with that context. Although such research needs to be done cautiously to not harm the students' education, the research would have greater ecological validity. In the present research, the participants were students in an Introduction to Psychology class. Although the research participation was a course requirement, earning credit for participation was not tied to the performance on the posttest. If gamified education techniques were employed in the classroom, the students may be more motivated to learn the material because their grades and academic advancement depended upon it. Thus, the relaxed demand characteristics of the present study may have undermined the ability to cleanly assess the utility of gamified teaching methods. That said, several students asked questions about their performance in the present study, suggesting that at least some participants were motivated to learn the material.

Conducting research on gamified educational techniques in the classroom will also ensure that the material to be learned is appropriate for the students' educational and developmental levels. The material to be learned in the present study was arbitrarily selected from the curriculum of an upper division psychology course on learning. Although this material was selected because the participants were unlikely to have pre-experimental knowledge of the material, learning these facts may have been difficult outside of the context of the rest of the course curriculum and, perhaps, the level of difficulty exceeded what a typical freshman could be expected to learn. Conducting research in actual classrooms may avoid some of these difficulties.

Despite the limited conclusions that can be drawn from the present study, additional research on gamified teaching methods should be encouraged. Robinson & Kakela (2012) discussed how creating an educational environment that allowed for fun interaction and trust helped build engagement, deep learning, and meaning for the students. Examining how students learn and what ways to increase their motivation and drive to learn is important, not only to improve their performance, but to also improve how they understand and grasp the material that they are tasked to learn. Also, of importance was the participants' expressed interest in the use of gamification which matched with a major trend in the literature of the field.

CONCLUSIONS

Participants in the current study showed much interest in the research and during debriefing asked many questions about gamified instructional methods, suggesting that there is strong social acceptability of these approaches. In addition, there are many advantages to the use of computer-based instruction and gamified educational approaches. Using a computer-based curriculum could enable self-paced instruction where progress through the curriculum is contingent upon mastering prerequisite material. Using computer-based instruction could also lead to more efficient use of the teacher's time by allowing the teacher more one on one instruction with struggling students, while the rest of the class is engaged in computer-based instruction. Indeed, there is much potential upside to gamified computer-based instruction, but to achieve that potential will require much additional research. Although this study found the use of gamification unable to improve performance on a test, the two versions performed similarly on the test. As

participants also expressed interest in the gamified version of the presentation and test, gamification could still be implemented within classrooms. Students' enjoyment of these techniques could lead to other long-term benefits that this study was unable to assess. Though these techniques may not significantly improve test performance, the interest and preference that students show for these techniques are good starting points for how gamification could best be used.

REFERENCES

- Alshammari, M. T. (2020). Evaluation of Gamification in E-Learning Systems for Elementary School Students. *TEM Journal*, 9(2), 806–813. <https://doi-org.proxy.lib.siu.edu/10.18421/TEM92-51>
- Armier, D. D., Shepherd, C. E., & Skrabut, S. (2016). Using Game Elements to Increase Student Engagement in Course Assignments. *College Teaching*, 64(2), 64-72.
- Chang, J. & Wei, H. (2016). Exploring Engaging Gamification Mechanics in Massive Online Open Courses. *Journal Of Educational Technology & Society*, 19(2), 177-203.
- Chapman, J. R., & Rich, P. J. (2018). Does educational gamification improve students' motivation? If so, which game elements work best? *Journal of Education for Business*, 93(7), 315. <https://doi-org.proxy.lib.siu.edu/10.1080/08832323.2018.1490687>
- Cicchino, M. (2015). Using Game-Based Learning to Foster Critical Thinking in Student Discourse. *Interdisciplinary Journal of Problem-Based Learning*, 9(2), 57-74.
- Danka, I. (2020). Motivation by gamification: Adapting motivational tools of massively multiplayer online role-playing games (MMORPGs) for peer-to-peer assessment in connectivist massive open online courses (cMOOCs). *International Review of Education / Internationale Zeitschrift Für Erziehungswissenschaft*, 66(1), 75. <https://doi-org.proxy.lib.siu.edu/10.1007/s11159-020-09821-6>
- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: defining "gamification." *MindTrek*, 9–15. doi:10.1145/2181037.2181040
- Hair, Joseph F., Jr., William C. Black, Barry J. Babin, and Rolph E. Anderson (2010), *Multivariate Data Analysis*, 7th ed., New Jersey: Prentice Hall.
- Kingsley, T. L., & Grabner-Hagen, M. M. (2015). Gamification. *Journal of Adolescent & Adult Literacy*, 59(1), 51-61.
- Leaning, M. (2015). A study of the use of games and gamification to enhance student engagement, experience and achievement on a theory-based course of an undergraduate media degree. *Journal of Media Practice*, 16, 155-170. doi:10.1080/14682753.2015.1041807
- Robinson, C. & Kakela, P. (2012) Creating a Space to Learn: A Classroom of Fun, Interaction, and Trust, *College Teaching*, 54(1), 202-207, DOI: 10.3200/CTCH.54.1.202-207
- Sailer, M., & Homner, L. (2020). The Gamification of Learning: a Meta-analysis. *Educational Psychology Review*, 32(1), 77–112. <https://doi-org.proxy.lib.siu.edu/10.1007/s10648-019-09498-w>
- Tews, M. J., Jackson, K., Ramsay, C., & Michel, J. W. (2015). Fun in the College Classroom: Examining Its Nature and Relationship with Student Engagement. *College Teaching*, 63(1), 16-26.