Assessing the Efficacy of Incorporating Game Dynamics in a Learning Management System

Raymond D. Frost
Vic Matta
Erin MacIvor
Management Information Systems Department
Ohio University
Athens, OH 45701, USA
FrostR@ohio.edu

ABSTRACT

The goal of the study was to see if gamification of a Learning Management System (LMS) would increase a number of desirable outcomes: student interest, motivation, satisfaction, student learning and perception of pedagogical affect. These constructs were measured in a survey, except for learning, which was measured by grades. Gamification of the LMS included the addition of all of the following: (1) An illustrated hero's adventure storyline with monsters to overcome by completing quests (assignments and assessments), (2) Olympic colored badges to represent individual grades as well as overall progress, (3) Points earned on a game-like scale—e.g., 100,000 points for the course, (4) A leaderboard with anonymous names and avatars, (5) Lives which allowed students to turn in a fixed number of late assignments without penalty. While open-ended responses suggested that students appreciated some gamification aspects, the quantitative data suggested that gamification has virtually no effect on the constructs measured. Only relatedness (a sub-construct of motivation) and student interest were found to be significant, although with small effect sizes. This study contributes to existing literature by exploring the impact of gamification of an LMS for a required introductory course in information systems.

Keywords: Gamification, Self-determination theory, Motivation, Game dynamics

1. INTRODUCTION

The video game industry has created a user base quickly and grown exponentially. In 2012, computer and video games sales reached $13.3 billion (Ellingson, 2013) and 211.5 million Americans play video games (Boorstin, 2012). Games, based on the market demands and industry, have been created specifically to engage the millenials (Dickey, 2005). People become enmeshed with these games and spend countless hours poring over the storylines and content. The video game industry has established a strong connection with the youth market.

The millennial generation is the largest and most diverse group to ever attend college. They are characterized as achievers who rely on technology and an extensive support system (Strauss and Howe, 1991). What is clear is that video games engage the millennial students. By some estimates, millennial students spend 10,000 hours playing video games by the time they are twenty-one (McGonigal, 2011).

The question is whether it could be beneficial for educators to integrate video game dynamics into classroom instruction, an idea known as gamification. Gamification is not the same as a game. Playing a game is a **voluntary** attempt to overcome **unnecessary** obstacles (Suits and Hurka, 2005). Required coursework is not voluntary, and the obstacles (assignments, tests, attendance, etc.) are necessary. To gamify is to use game elements such as points, badges, and leaderboards in a non-game context. Adding game elements such as a Leaderboard, or a Storyline to Blackboard is an example of gamification. Such an experiment should look at whether gamification increases student engagement and learning outcomes. The next section discusses extant research in this area and shows how it falls short of measuring such outcomes.

At the college level do students benefit from dynamic and entertaining teaching strategies to achieve the same levels of engagement as that of prior generations? Could there be some value in introducing some video game elements into a course – specifically into the learning management system? More specifically, the question is whether there is any value in gamifying the content of a **required** course.

2. LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

Gamification is the use of game mechanics and thinking in non-game contexts such as education and was first coined in 2002 by Nick Pelling (Marczewski, 2012). One theoretical measure of engagement that fits well with gamification and

education is Self-Determination Theory (SDT). SDT is a cognitive autonomous theory that identifies deeper psychological needs that, when fulfilled, create engagement and produce self-regulated behavior (Ryan, Kuhl and Deci, 1997). SDT is based on three different needs: **relatedness**, **competence**, and **autonomy**. SDT claims that relatedness, competence, and autonomy are psychological needs that should be fulfilled as the foundation of someone's self-motivation. SDT has been used to study motivation in many settings including health care (Ryan et al., 2008; Williams et al., 2011), business (Vansteenkiste et al., 2004; Baard, Deci, & Ryan, 2004), and education (Ryan & Deci, 2002; Ryan & Brown, 2005).

However, most courses use extrinsic motivation almost exclusively in the form of points and grades--especially in required courses. When a student completes a task due to extrinsic motivation, he/she is driven by external environmental factors such as competition, physical rewards, and threats (Benabou and Tirole, 2003). Extrinsic motivation reduces autonomy and creates a controlled setting. Extrinsic motivation may produce positive results, but it is usually accompanied with adverse emotions such as pressure to perform or deliver (Pelletier et. al, 1995). Although this type of motivation has some potential, it does not create the ideal learning environment for students because of the negative impact it has on emotions.

By contrast, intrinsically motivated individuals complete tasks with a sense of choice and eagerness (Hagger, Chatzisarantis & Harris, 2006). Intrinsic motivation can be difficult to accomplish in education; however, it has been associated with lower dropout rates, better learning strategies, and more interest in school (Carlton and Winsler, 1998; Deci and Ryan, 1985; Kauffman & Husman, 2004; Moneta, 2004). Achieving balance between extrinsic and intrinsic motivation in the classroom is essential for educators to create a beneficial learning environment for students.

Prior research found positive outcomes from classroom gamification (Sheldon, 2011). However, on further examination, that research focused on elective courses, in some cases where the very subject of the courses was gamification. At times the classroom activities were also conducted as a game. Therefore the prior research benefited from intrinsic motivation. By contrast, this study looked at gamifying a Learning Management System (LMS) of a **required** introductory core course in information systems with a **standard** curriculum, and therefore attempted both extrinsic (points and grades) and intrinsic (gamification) motivation.

Rigby and Ryan (2011) established a connection between video games and SDT in *Glued to Games*. In the beginning of video game development, the focus was on competence and then it progressed to include autonomy and relatedness. The first home video game, Pong®, was an example of the focus on competence. Without much graphic design, choice, or connectivity, Pong® rewarded players for keeping a ball in the scope of the screen. Based solely on competence, the game became the most popular game in 1975 with sales reaching $40,000,000 (The Great Idea Finder, 2007). As the industry expanded and grew to include more complex games, Pong® was eclipsed by Lunar Lander and Hunt the Wampus. Besides competence, both of these games infused autonomy with their game play by removing strict rules and allowing the user to choose a path that affected the outcome of the game. In the next generation of video games, the development went one step further and incorporated all three dimensions. One of today's top multiplayer online role playing games, World of Warcraft, demonstrates the relatedness of people through online interactions and through characters, autonomy in the creation of one's character and actions, and competence in rewarding achievements in the game. The market now consistently develops games that leverage the motivations of SDT. The fact that the theory had been validated in both education and gaming made it particularly attractive for the current study which attempts to merge the two.

To achieve a balance between extrinsic and intrinsic factors, the educators' actions should be grounded in a theory based on both. SDT is a broad framework that defines three fundamental needs essential to human motivation and engagement. When the three needs (relatedness, competence, and autonomy) are fulfilled, students feel more engaged. Relatedness and autonomy increase the amount of intrinsic motivation (Weinstein, Przybyski, and Ryan, 2009). Competence has been linked to extrinsic motivation (Harter, 1981). Satisfying these needs results in a higher level of engagement for students, meaning that there is an improvement in the quality of learning, as well as better retention (Czubaj, 2004). These three needs, relatedness, competence, and autonomy, are vital for producing the most effective learning atmosphere, and can be said to impact motivation. SDT has been applied to the education environment to study motivation (Deci & Ryan, 2009) and self-regulation of learning (Ryan, Deci, & Williams, 2008). SDT is a good lens through which to look at education and video games, to combine the best practices of both areas and create a new teaching style. Relatedness, competence, and autonomy combined create self-determination theory. While this is a broad framework to define human motivation, it has been applied to describe a productive learning environment and to explain the success of the video game industry. We therefore hypothesize:

H1-motivation: The use of game dynamics will increase motivation as measured by the SDT

The first psychological need, relatedness, refers to the need to have meaningful connections between one's self and others. When people feel interconnected, they are more likely to feel motivated. People require quality relationships as a support system in their environment in order to act with intrinsic motivation (Markland et al., 2005). The relatedness dimension had often been deemed unimportant due to early research in SDT, which focused on intrinsic motivation. These studies demonstrated that people could be intrinsically motivated to do solitary activities (Koestner & Losier, 2002; Ryan and Deci, 2002). However, SDT still holds that relatedness is an essential piece and should be included in research. Individuals will not adopt structure from someone they do not think cares about them (Ryan and Deci, 2003). We therefore hypothesize:

H1a-relatedness: The use of game dynamics will increase relatedness in the course.

The second psychological need is competence. This refers to the human necessity of challenge and recognition of accomplishments (Rigby and Ryan, 2011). The tasks a person is assigned should be challenging, but possible. People need to know clear and compelling standards with affirmation of performance (Schlechty, 1997; Dickey 2005). In fulfilling the need for competence, it is important to be clear, consistent, and challenging in order to make people feel the task is worth their time. We therefore hypothesize:

H1b-competence: The use of game dynamics will increase competence in the course.

The third psychological need is autonomy. Autonomy reflects the inner need to take control of a situation and react with personal choice and without constraints or fear of consequences (Rigby and Ryan, 2011). The need for autonomy is fulfilled when a participant has the freedom to make meaningful choices and influence the outcome. Novelty, variety, and choice are identified as critical for encouraging autonomy (Schlechty, 1997; Dickey, 2005). In a video game, allowing a character to die and the player to try again is an example of an autonomous action. Characters are free to jump off a building, fight other characters, or explore a new cave with heavy breathing sounds exuding from it. All of these, in real life, could lead to death. In the game, players don't worry much about death—and in some cases enjoy a glorious demise. Incorporating autonomy allows a person to make choices and experience the results without experiencing serious ramifications for their actions. We therefore hypothesize:

H1c-autonomy: The use of game dynamics will increase autonomy in the course.

A gamer is interested in achieving a higher score against the odds presented by the game. The analog of this interest and drive in an academic environment is in being challenged and becoming more competent. In his research, Marks (2000) shows that interest is synonymous with a perception of learning. An individual interested in a game would perform better in the game than a disinterested student. In their research, Paswan and Young (2002) also show that student interest is central to multiple pedagogical outcomes, including performance (Abrantes, Seabra, and Lages, 2007). We test the theory that gamification would have a similar increase in interest, and therefore hypothesize:

H2-interest: The use of game dynamics will increase interest in the course.

This research also examines student satisfaction as an overall measure to indicate desirability of the gamified LMS. Ioannou and Artino (2009) examine satisfaction as an overall outcome of their pedagogical exercise on online collaborative learning. This research proposes that gamification of the LMS will have a similar increase in satisfaction, and purports to use the same instrument to test the following hypothesis:

H3-satisfaction: The use of game dynamics will increase satisfaction with the course.

A gamified interface is different from what students have seen, and suggests that the instructor is interested in improving the learning environment, has made an effort to provide a gamified interface, and is organized enough to deploy a gamified interface in addition to the academic preparation needed to teach the course. Abrantes, Seabra, and Lages (2007) have shown that students perceive the instructor's increased effort as an indicator of good pedagogy. Abrantes, et al. call this construct pedagogical affect.

H4-pedagogy: The use of game dynamics will increase the perception of pedagogical affect in the course.

In addition to all of the above, we were also interested in measuring, as a practical matter, whether gamification would lead to real gains in learning, as opposed to perceptions of learning. Therefore we hypothesize:

H5-learning: The use of game dynamics will increase learning as measured by test and assignment scores.

3. METHODLOGY

An experiment was conducted with two sections of a sophomore level business core course in information analysis and design at a large Midwestern university in fall 2012. All the registered students were in the College of Business. The course has both a lecture and lab component. This course introduced students to basic graphic design principles, allowed them to design an iPhone App, sell it in a simulated marketplace, conduct sales analysis using business intelligence tools, and conduct financial analyses on a fictional company marketing the app. Personal software tools including Microsoft PowerPoint, Microsoft Excel, Microsoft Access, Google Analytics, Google Sites, and Google Docs were all used to complete assignments. At the end of the semester, students were required to create a report aimed at a venture capitalist, combining all prior assignments.

To test the hypotheses, one section of the course interacted with a gamified version of the LMS, while the other section interacted with the unmodified LMS. A survey was administered to test for the differences in outcomes stated above (see Appendix). These are explained in more detail below.

3.1 Design of the LMS with Game Dynamics

The LMS used in this experiment is called Integrated Site Management System (ISMS). It is a non-commercial product designed in-house. It has been in use for over ten years in our college and is preferred over Blackboard due to its improved functionality, reliability, and stability as compared to Blackboard. ISMS has standard features similar to commercial products such as Blackboard. These include dropboxes, grade placeholders, wikis, online quizzes,

announcements, as so forth. ISMS was used, however, because it is more flexible and extensible. It was extended with deliberately selected game dynamics to increase student interest, good pedagogy, and satisfaction, along with the components of SDT (competence, autonomy and relatedness) that impact motivation. In order to ensure that the design would appeal to students, students in an upper level elective class designed the interface. They came up with competing designs, and then as a class picked the best one. The subject of the elective class was Gamification of Education, and students who enrolled tended to be gamers.

Since relatedness is the need to feel interconnected, it was introduced through two main elements: a storyline and a leaderboard. The storyline was intertwined with the material to provide a holistic and immersive experience for the students. Each week, the students were challenged by an assignment with the learning objective represented by a "monster". When the students completed a task, the monster was defeated. The second element, a leaderboard, is an anonymous listing of all the students and their grades. Each student has an avatar and "scholar tag" to preserve anonymity and comply with the U.S. Family Educational Rights and Privacy Act (FERPA) regulations. The leaderboard ranks each student's performance against the rest of the class from highest to lowest grade. It is available for all students to see.

The leaderboard also contributes to the competence construct as it demonstrates the overall skill of a student relative to others in the course.

Aside from the leaderboard, competence was built into the system by using Olympic colored medals and a positive growth attendance policy. The medals were a way to quickly show a student his or her level on any assignment. Every grade was designated a color: 93% and up was highlighted in gold, 83% - 92.9% was in silver, 73% - 82.9% was in bronze, 63% - 72.9% was in dark red, and 0% - 62.9% was in bright red. The colors were applied to individual assignments, the overall grade, and the leaderboard. Students were able to easily see their grade levels and get a sense of how they fared as compared to the class. If a student only had grades above a 93%, the screen would show all gold medals, reaffirming success and competence. A screenshot of this interface is included in the appendix.

In order to cover all learning types, the learning outcome component of this research was measured using three types of assessments: homework assignments (HW), lab assignments (Labs) that are done at school under supervision and the midterm exam (exam) comprised of multiple choice questions.

The fact that the LMS showed overall average turned out to be a problem. Most video games show only points and those points always increase over time. By contrast a student's average can go down over time. To combat this, the attendance grade was factored in such a way as to virtually guarantee that a student's overall average would increase as the semester progressed. We call this positive grade growth.

To incorporate autonomy, a system of using lives was developed. Each student began the semester with three lives. A life could be used to turn in an assignment up to 48 hours late with no consequences. In true game style, the lives were represented as three hearts at the top of the grade view screen for the student. If the student did not use all of his or her lives, the lives were redeemed at the end of the semester for extra credit. Students were given the freedom to choose when to use these lives. The control group also had the same benefit of turning assignments in late, though they were not called lives, nor did they have "the hearts" visual representation.

Autonomy was also worked into the leaderboard. As previously stated, the leaderboard ranks students against the rest of the class in an anonymous manner. To keep the students' identities hidden, they were asked to generate a creative name. In many environments, this is called a screen name. This research designates it as a scholar tag. As a companion to the scholar tag, students were asked to create an avatar. An avatar is a simple picture that is placed next to the student's scholar tag on the leaderboard. Students were given few guidelines to complete this task and encouraged to be creative.

In summary, the LMS was extended to incorporate game dynamics. These include a storyline, leaderboard, medals, lives, and positive grade growth.

3.2 Experimental Design

A convenience-based two-group sample was used for this experiment. There were 39 students in the control group and 41 students in experimental group. Each group used an LMS to turn in assignments, receive grades, and find class materials. Only the experimental group had the gamified elements of the LMS activated.

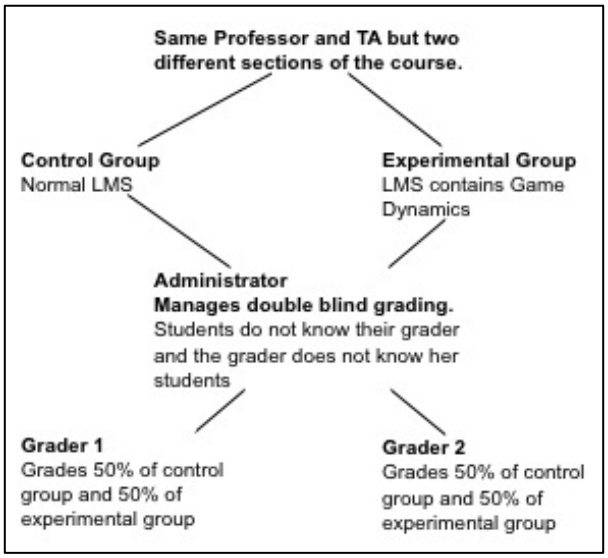


Figure 1: Grading and Teaching Setup

Factors not under investigation were held constant. For instance, each of the two sections was taught by the same professor in lecture and the same Teaching Assistant (TA) in lab. To avoid bias, the grading was done double blind as explained by Figure 1. Two teaching assistants, labeled the "graders", were randomly given 40 students to grade each week from both groups. After completing the grading, the graders submitted their grade book to the administrator, who posted the grades to the students. Any questions about

individual grades were sent to the administrator to avoid contaminating the graders' perception of a particular student. The administrator either forwarded the question on to the respective grader or answered it herself.

The Experimental Group contained all of the game dynamics. These were based on the Rigby and Ryan's *Glued to Games* (2011), see Table 1.

Construct	Experimental Group	Control Group
Competence	Leaderboard, Positive Grade Growth, Medals/Colors Game like 100,00 point scale	Medals/Colors (could not remove from system), Grades fluctuate up and down Regular 1,000 point scale
Autonomy	Leaderboard, Lives System, Avatars and Scholar Tags	Three late assignments accepted but not called "Lives"
Relatedness	Leaderboard, Storyline	n/a

Table 1: Game Dynamics in the LMS

Each of these game dynamics was specifically chosen to fulfill a need of self-determination theory and simulate a game-like environment. The experiment implemented gamification in the LMS while fulfilling all three needs of self-determination theory.

3.3 Survey Development

To understand the benefits of gamification in a LMS, a survey (see appendix) was administered to the two groups, to test student motivation (using the three dimensions of SDT), their perception of course quality (pedagogy), and their satisfaction and interest in the course.

To measure all the constructs of this study previously validated instruments were used. The SDT scale measured constructs for competence, autonomy and relatedness in the work environment (Baard, Deci & Ryan, 2004; Deci & Ryan, 2000). It was modified slightly to change the phrase "at work" to "in class". There were 21 statements, with random reverse coding, using a seven point Likert scale from "Not at all true" to "Very true". These dimensions nurture intrinsic motivation and lead to desired educational outcomes (Deci et al., 1991).

The other subscales that were measured included satisfaction with the learning management system, student interest, and perception of pedagogy affect. These subscales were included to expand on the current research. Constructs for student interest, and pedagogy were taken from the work of Abrantes, Seabra, and Lages (2007). The satisfaction subscale was based on a study by Ioannou and Artino Jr. (2009) in which they examined satisfaction in a collaborative learning environment. It was adapted to reference the LMS system using a seven point Likert scale from "Strongly disagree" to "Strongly agree". The stem of each statement began with "Overall, in this class the methods of instruction were", but used different anchors on the Likert scale (e.g. "Ineffective" to "Effective", "Useless" to "Useful", "Unsatisfactory" to "Satisfactory", and "Bad" to "Good"). Performance was measured using grades from three categories: homework assignments, applied lab assignments and an exam. Students in the experimental group were also asked seven open ended questions about the strong points and weak points of the game dynamics: the leaderboard, the lives, and the gaming language.

In order to ensure that the items performed well as a group, Cronbach Alpha was assessed. Two items were found to not work well, one in the Autonomy construct and one in the Competence construct. The Autonomy item was "When I am in this class, I have to do what I am told." This is a very strong statement in an American university classroom. Some work environments can be far more stringent, for example, a worker on an assembly line. The remaining Autonomy items were not as extreme in their wording, and their adaptation to the classroom environment was equivalent.

The Competence item that was flagged was "People in this class tell me I am good at what I do." The adaption from a work environment to an American university was not equivalent. Federal Education Rights and Privacy Act (FERPA) prevents any sharing of grades and therefore reduces the impact of social recognition implied in the statement. In addition, Americans often hide their talents, especially at the introductory level, continuing the high school tradition that it is not cool to be smart.

4. RESULTS AND DISCUSSION

4.1 Data Descriptives and Assumptions

In the Table 2, N refers to the group sizes. The group size varies by construct based on the number of usable responses for each. Responses that were incorrectly entered or missing were not counted. The extent of learning imparted from gamification is called Performance, and was assessed by running tests on homework, lab work and an exam.

Constructs	Group	N	Mean	Std. Dev.	Std. Err. Mean
Autonomy	Control	39	4.716	.6377	.1021
	Experimental	41	4.752	.6631	.1035
Relatedness	Control	39	4.300	.7867	.1259
	Experimental	41	4.624	.8981	.1402
Competence	Control	39	4.923	.7471	.1196
	Experimental	41	4.926	.8277	.1292
Motivation	Control	39	4.646	1.6868	.2701
	Experimental	41	4.767	1.8972	. 2963
Satisfaction	Control	39	5.247	1.0673	.1709
	Experimental	41	5.477	1.2711	.1985
Interest	Control	39	3.814	.7602	.1217
	Experimental	40	4.106	.6884	.1088
Pedagogy	Control	39	5.790	1.0558	.1690
	Experimental	41	5.872	1.1714	.1829
Learning - HW	Control	44	90.44	4.322	.6516
	Experimental	43	90.63	3.801	.5797
Learning - Labs	Control	44	93.93	4.383	.6608
	Experimental	44	94.31	3.838	.5787
Performance - Exam	Control	44	76.96	11.665	1.758
	Experimental	43	79.79	8.874	1.353

Table 2: Descriptive Statistics

4.2 Assumptions

Two pre-test assumptions were evaluated prior to the analysis: equal variances and normally distributed data. To detect the equality of variance, Levene's test was used for each construct. The null hypothesis of Levene's test states that the variances for the two groups are equal. Significant results ($p<.05$) lead to a rejection of equality of variance. Table 3 shows the results of Levene's test, $p>.05$ for all factors but one – Performance on Exams.

The assumption of normality was assessed using the Kolmogorov-Smirnov test. The null hypothesis assumes a normal distribution. Significant results $p<.05$ lead to a rejection of normality. Table 4 shows the results of the Kolmogorov-Smirnov test. The pedagogy and learning on exam constructs failed this test for both the control and the experimental groups. The interest, competence and learning on homework constructs failed this test for the experimental group while the performance on lab assignments construct failed for the control group. Mann-Whitney's test was therefore used to test differences.

Levene's Test for Equality of Variances	F	Sig.
Relatedness	.382	.538
Competence	.087	.769
Autonomy	.662	.418
Motivation	.299	.586
Satisfaction	.134	.715
Interest	.203	.653
Pedagogy	.000	.992
Learning – Homework	2.007	.160
Learning – Lab work	.084	.773
Learning – Exam	5.371	.023*

*p<.05, +This is a lower bound of the true significance.

Table 3: Levene's Test for Equality of Variance

Table 5 and 6 show the results of appropriate independent samples tests. On most dimensions, this analysis rejected the alternate hypothesis and supported the null: The use of game dynamics in the LMS did not affect student autonomy, competence, satisfaction, motivation, learning and perception of pedagogical affect.

Constructs	Group	Kolmogorov-Smirnov		
		Statistic	df	Sig.
Relatedness	Control	.124	39	.136
	Experimental	.109	40	.200+
Competence	Control	.104	39	.200+
	Experimental	.168	40	.006*
Autonomy	Control	.093	39	.200+
	Experimental	.075	40	.200+
Motivation	Control	.078	39	.200+
	Experimental	.094	40	.200+
Satisfaction	Control	.093	39	.200+
	Experimental	.121	40	.145
Interest	Control	.109	39	.200+
	Experimental	.141	40	.043*
Pedagogy	Control	.168	39	.007*
	Experimental	.170	40	.005*
Learning -HW	Control	.120	44	.114
	Experimental	.150	42	.018*
Learning -Lab	Control	.201	44	.000*
	Experimental	.127	42	.087
Learning -Exam	Control	.150	44	.014*
	Experimental	.150	42	.019*

*p<.05 (one-tailed)

Table 4: Kolmogorov-Smirnov Test of Normality

However, on the dimension of relatedness and interest, the analysis rejected the null and supports the alternative hypothesis: The use of game dynamics in the LMS increased in relatedness and interest ($p<0.05$: sig=0.046 and sig=0.043).

So in summary, only relatedness and interest were statistically significant. Next, we looked at the effect size for relatedness and interest. Effect size is a measure of strength for a phenomenon. According to Cohen (1988), it is calculated using the means and standard deviations of two groups, and ranges from .2 for small, to .5 for medium, to .8 for large effects. Though statistically significant, the effect sizes were tiny for both relatedness (.036) and interest (.023).

T-test for Equality of Means	t	df	Sig. (2-tailed)	Sig. (1-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower	Upper
Relatedness (H1a)	-1.71	78	.091	.046*	-.32387	.18916	-.70046	.05272
Autonomy (H1c)	-.050	78	.960	.48	-.00730	.14558	-.29712	.28253
Satisfaction (H3)	-.875	78	.384	.192	-.23023	.26310	-.75401	.29356

*p<.05 (one-tailed)

Table 5: Student's T-test for Equality of Means

Mann-Whitney U Test	Competence (H1b)	Motivation (H1)	Interest (H2)	Pedagogy (H4)	Learning (H5)		
					HW	Lab	Exam
Mann-Whitney U	771.500	709.50	606.00	752.00	916.50	921.50	858.00
Wilcoxon W	1551.500	1489.50	1386.00	1532.00	1906.50	1911.50	1848.00
Z	-.271	-.866	-1.717	-.463	-.251	-.388	-.750
Asymp. Sig. 2-tail	.787	.386	.086	.643	.802	.698	.453
1-tail	.394	.193	.043*	.321	.401	.349	.226

*p<.05 (one-tailed)

Table 6: Mann Whitney Test Statistics

4.3 Discussion

The results of the experiment were surprising, because the only elements that had significant impact were relatedness (H1a) and student interest (H2). But even though these were statistically significant, their effect sizes were tiny. In other words, they were only weakly confirmed. The remaining elements: competence (H1b), autonomy (H1c), motivation (H1), satisfaction (H3), pedagogy (H4) and learning (H5) constructs were not significant.

As noted earlier, relatedness is about feeling interconnected, and was operationalized through the leaderboard and the storyline. These elements appeared to fuel students' ability to associate with other individuals as well as the class as a whole. Students commented on the comparative and competitive influence of the leaderboard, and showed appreciation for its ability to display how the rest of the class performed, and as a result felt better connected with the class, no matter how they ranked on the leaderboard, as also noted by Banfield and Wilkerson (2014). Next, student interest was seen as significantly impacted by gamification of the LMS. It was indicated by increase in interest in the course material, attentiveness, intellectual challenge and competence. Language and the storyline appeared to contribute most to this increase of interest. Students believed that gamification made the course more exciting and fun.

In addition to the hypotheses, this study also explored grades on two types of assignments: those that are typical of most courses, and those that had fun elements (like designing an iPhone app) in them. For all assignments, both sections performed similarly. This suggests that there is no advantage in creating fun assignments that fit better with a game environment.

The experiment was carefully controlled as explained in the section on Experimental Setup. These included use of the same instructor, double blind grading, careful control for cross contamination of game concepts in the non-gaming class, as well as equal opportunities in both classes.

To help explain the results we also used qualitative data. Students were asked to name a positive and negative quality of each game dynamic (summarized in the Table 7). These comments give some insight into the increases in relatedness and interest. As is apparent in the table, each of the gamified elements (Leaderboard, Lives, Language and Medals) had net positive impact. The differences across all gamified elements was almost two to one positive.

The leaderboard was designed to fulfill the need for Relatedness, Competence and Competitiveness. However, students commented on the leaderboard's ability to display grades, but did not feel rewarded when they were ranked higher. Most students concurred that the leaderboard helped them understand how they performed in comparison to others. A few believed the leaderboard motivated them (three students), but more believed it was discouraging (nine students). In some cases, students felt negative pressure to perform. Not ranking in the top few was regarded as a negative experience, but the converse did not hold true -- it was not viewed as a positive experience to be highest ranked on the leaderboard. Other students stated their experience was "intimidating if you fell behind" or "you may feel discouraged if you are not doing so well in the class compared to others." This repurposing of the leaderboard from the fulfillment of competence to the fulfillment of relatedness may explain the increase in the relatedness construct without much impact on the competence construct.

The other game dynamics, the lives and the gaming language, experienced a repurposing in the eyes of the students as well. The lives represented the ability to turn in an assignment late. Each student was given three lives to use on whichever week they chose. If they were not used by the end of the semester, the lives were redeemed for points. While meant to fulfill the need for autonomy, many students saw lives as helpful in the event they needed a second

Positive Comments	**Leaderboard**	**Lives**	**Language**	**Medals**
Comparative / Competitive	25			1
Transparency	17	12		17
Motivating	3			2
Interesting / Innovative	2		11	8
Easy to use	2	2		
Helpful		17	3	
Extra freedom		2		
Rewarding		8		
Fun/Exciting		3	22	
Less Serious			8	
None (not in total)			5	3
Total Positive	49	44	44	28
Negative Comments	**Leaderboard**	**Lives**	**Language**	**Medals**
Comparative / Competitive	6			
Too transparent / visible	4			
Discouraging / Intimidating	9			
Not transparent / visible enough	3			1
Confusing / hard to use	6	7	25	4
Allows for laziness		10		
Felt pressured / invoked panic		3		1
Insufficient		2		
Devalues assessment			5	
Pointless / meaningless		2	4	3
None (not in total)	18	20	11	3
Other/Misc.			3	2
Total Negative	28	24	37	11

Table 7: Qualitative Comments

chance. The lives were not viewed as additional freedom, but as a safety net or an emergency resource. Often, students did not freely give up the lives, and elected to protect their lives instead of using them, even when using them would be in their best interest. The other main viewpoint from the students was that the lives allowed people to be lazy. Neither of these opinions supported the use of extra freedom (autonomy) in the course. Nonetheless, the lives serve as a convenient administrative tool to handle late assignments.

The gaming language also deviated from its original purpose. The language was incorporated to satisfy the need for relatedness, but was instead viewed as fun, exciting, interesting, and different by the students. None of the students mentioned a connection through a common experience caused by the language. Overwhelmingly, students stated it created a less serious atmosphere where "learning the material [was] a bit more enjoyable". The use of the gaming language lost its intended purpose of fulfilling the need for relatedness, but instead increased student interest. Further loss of interest could be explained by the fact that this experiment gamified just the LMS, not the entire course.

In fact, most of the game dynamics contributed to the increase in student interest. The leaderboard, the gaming language, and the medals all had "interesting" mentioned in the comments. Many students listed this as a benefit to the whole idea of game dynamics as well. While the game dynamics did not always achieve their intended goals, they increased perceived relatedness and interest.

Another possible explanation for the positive reception of the gamification is that the game elements were appreciated simply because they provide better feedback to the students and not because they produce engagement. The leaderboard gives students a clear indication of where they stand relative to their peers. The badges give students a quick visual of their performance on each assignment. The hearts give students a clear indication of where they stand relative to turning in late assignments. Indeed the authors have chosen to continue to employ these three game elements precisely because they are well received even if they do not increase engagement. The lives in particular, can be used to systemically allow for late assignments while monitoring and preventing abuse. We recommend clearly explaining all game elements at the beginning of the semester.

Finally, a truly disappointing outcome was that no gain in learning outcomes was found in homework assignments, labs or the final exam. So we are abandoning the expectation of learning outcome improvements through gamification, at least for now. Furthermore, a gamified course may create a false expectation that the assignments and exams should be fun and greater disappointment when they turn out to be work instead.

5. IMPLICATIONS

The researchers were surprised by the lack of statistical significance in the results. Even the constructs that were significant, relatedness and interest, had small effect sizes. The lack of results is even more surprising given that the prior pilot study and the open-ended responses from this study were overwhelmingly positive. We posit four explanations for this disparity.

The prior pilot study was performed on an upper level elective class whose very subject was gamification of education. Students entering the prior course were biased in favor of gamification and provided positive and enthusiastic feedback. By contrast, students in the current study were less mature (freshman and sophomores) and were in a required course. Secondly, Research suggests that new pedagogies must be sold to students (Smith, 2008). This would require explaining the reasons for gamifying the LMS. We made no attempt to sell the pedagogy in order to avoid biasing the study or introducing a novelty effect (Adair, 1984). Selling the LMS to students may make a good future study. Third, despite the fact that our system contained standard game elements (i.e. points, badges and a leaderboard) and had a story line to enhance the game environment, the design was constrained by the limitations of how far the retrofitted LMS could be extended. Perhaps an LMS designed with the sole objective of gamifying education would have yielded better results. Fourth, gamification could be extended into course content or assignment completion (e.g. using a virtual reality based system), and therefore increase its effect.

5.1 Limitations

Clearly more research needs to be done in this area. Generalizability of this research is limited by the fact that it was conducted at a Midwestern US university in a required undergraduate introductory Information Systems course (i.e. one university, one course). The experiment involved two sections of the course with 39 & 41 students. A larger sample size would enhance the results, though the effect size would not change. No known limitations were perceived in the experimental setup – as discussed above, it was followed carefully so as not to create confounding effects/biases. As noted earlier, a natively gamified LMS (versus a retrofitted one) may have a stronger effect. Lastly, it is possible that there is a better theory that would support gamification of an LMS. The literature pointed strongly in the direction of Self Determination Theory given its prior application in both education and game environments. However, that does not necessarily mean that the theory would provide the best explanation of gamification in education. Nor does it mean that an innovative teaching method will always improve learning.

5.2 Future Research

Nonetheless, the positive comments in the current study are some cause for hope for both teachers and researchers. Researchers may try measuring different motivation constructs. Maybe self-determination theory is not the answer. Here are some suggestions for further research: Work with upper level elective courses rather than required courses. Perhaps refining the gamification techniques with a sympathetic and more mature audience would yield better results. For teachers, we suggest selling the pedagogy on the first day of class. Students need to see how the pedagogy could be in their best interest. Given the competitive LMS market, maybe one of the players would distinguish itself by focusing on gamification (Villagrasa et al., 2014).

Fortunately, gamification has been a rapidly developing landscape. For example, Duolingo.com provides an excellent and engaging gamified platform for language acquisition. A

quick search reveals other new natively gamified LMSs under development. Additionally, mainstream LMSs are also developing increasingly immersive gamified environments. Blackboard and Moodle already have building blocks and plug-ins. Canvas has been providing grants for development in this area. It might also be interesting to shift the focus from gamification of the LMS, to gamification of individual assignments. Many teachers receive positive feedback when they play Jeopardy or some other game during one class session. That is very different than gamifying the whole class. Our assignments were standard across both the experimental and control sections. But no real game would remove the game element from the individual challenges.

6. CONCLUSIONS

Our research shows that gamification is not a panacea. It will not manufacture student engagement and could even have some negative effects. We believe that the key determinant of the value of gamification is voluntary motivation. In an involuntary setting, such as a required course, participants may appreciate the gamification elements but those elements may not significantly increase motivation. Our conclusion could be tested in future research. For example, there is now a gamified language learning website called Duolingo.com.

As mentioned previously, playing a game is a voluntary attempt to overcome unnecessary obstacles (Suits and Hurka, 2005). Required courses are not voluntary and the obstacles are necessary. Therefore turning a required course into a game problematically violates the very assumptions of a game. This may explain the weak results on the SDT constructs as well as the lack of improvement in grades.

7. REFERENCES

Abrantes, J., Seabra, C., & Lages, L. (2007). Pedagogical Affect, Student Interest, and Learning Performance. *Journal of Business Research*, 60(9), 960-964.

Adair, J. G. (1984). The Hawthorne Effect: A Reconsideration of the Methodological Artifact. *Journal of Applied Psychology*, 69(2), 334–345.

Baard, P. P., Deci, E. L., & Ryan, R. M. (2004). Intrinsic Need Satisfaction: A Motivational Basis of Performance and Well-Being in Two Work Settings. *Journal of Applied Social Psychology*, 34, 2045-2068.

Banfield, J. & Wilkerson, B. (2014). Increasing Student Intrinsic Motivation and Self-Efficacy through Gamification Pedagogy. *Contemporary Issues in Education Research (CIER)*, 7(4), 291-298.

Benabou, R. & Tirole, J. (2003). Intrinsic and Extrinsic Motivation. *Review of Economic Studies*, 70(3), 489-520.

Boorstin, J. (2012). Fewer People Now Playing Videogames. *CNBC.com*. Retrieved April 14, 2013, from http://www.cnbc.com/id/48917308.

Carlton, M. P. & Winsler, A. (1998). Fostering Intrinsic Motivation in Early Childhood Classrooms. *Early Childhood Education Journal*, 25(3), 159-166.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences (2nd ed.)*. Lawrence Erlbaum Associates.

Czubaj, C. A. (2004). Internet Review: Educator Concerns Regarding Cyberspace Curricula. *Education*, 125(1), 15-19.

Deci, E. L. & Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Plenum.

Deci, E. L. & Ryan, R. M. (2000). The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, 11, 227-268.

Deci, E. L. & Ryan, R. M. (2009). Promoting Self-Determined School Engagement: Motivation, Learning, and Well-Being. in Wentzel, K. R. & Wigfield, A. *Handbook on Motivation at School*, 171-196. New York: Routledge.

Deci, E. L., Vallerand, R. J., Pelletier, L. G. & Ryan, R. M. (1991). Motivation and Education: The Self-Determination Perspective. *Educational Psychologist*, 26(3-4), 325-346.

Dickey, M. D. (2005). Engaging by Design: How Engagement Strategies in Popular Computer and Video Games can Inform Instructional Design. *Education Technology Research and Development*, 53(2), 67-83.

Ellingson, A. (2013). Game Over? Video Game and Console Sales Take a Head Shot. *L.A. Biz*. Retrieved April 14, 2013, from http://www.bizjournals.com/losangeles/news/2013/01/15/2012-video-game-sales-dropped-22-percent.html.

Hagger, M. S., Chatzisarantis, N. D., & Harris, J. (2006). The Process by which Relative Autonomous Motivation Affects Intentional Behavior: Comparing Effects across Dieting and Exercise Behaviors. *Motivation and Emotion*, 30(4), 306-320.

Harter, S. (1981). A New Self-Report Scale of Intrinsic Versus Extrinsic Orientation in the Classroom: Motivational and Informational Components. *Developmental Psychology*, 17(3), 300-312.

Ioannou, A. & Artino, A. (2009). Wiki and Threaded Discussion for Online Collaborative Activities: Students' Perceptions and Use. *Journal of Emerging Technologies in Web Intelligence*, 1(1), 97-106.

Kauffman, D. F. & Husman, J. (2004). Effects of Time Perspective on Student Motivation: Introduction to a Special Issue. *Educational Psychology Review*, 16(1), 1-7.

Koestner, R. & Losier, G. F. (2002). Distinguishing Three Ways of Being Highly Motivated: A Closer Look at Introjection, Identification, and Intrinsic Motivation. in Deci, E. L. & Ryan, R. M. (eds.), *Handbook of Self–Determination Research*, 101-121. Rochester: University of Rochester Press.

Marczewski, A. (2012). *Gamification: A Simple Introduction & a Bit More (1st ed.)*.

Markland, D., Ryan, R. M., Tobin, V. J., & Rollnick, S. (2005). Motivational Interviewing and Self Determination Theory. *Journal of Social and Clinical Psychology*, 24(6), 811-831.

Marks, R. B. (2000). Determinants of Student Evaluations of Global Measures of Instructor and Course Value. *Journal of Marketing Education*, 22(2), 108–19.

McGonigal, J. (2011). *Reality is Broken: Why Games Make us Better and How they can Change the World*. New York, NY: Penguin Books.

Moneta, G. B. (2004). The Flow Model of Intrinsic Motivation in Chinese: Cultural and Personal Moderators. *Journal of Happiness Studies*, 5(2), 181-217.

Paswan, A. K. & Young, Y. A. (2002). Student Evaluation of Instructor: A Nomological Investigation using Structural Equation Modeling, *Journal of Marketing Education*, 23(3), 193-202.

Pelletier, L. G., Fortier, M. S., Vallerand, R. J., Tuson, K. M., Brière, N. M., & Blais, M. R. (1995). Measure of Intrinsic Motivation, Extrinsic Motivation, and Amotivation in Sports: The Sport Motivation Scale. *Journal of Sport & Exercise Psychology*, 17, 35-53.

Rigby, S. & Ryan, R. M. (2011). *Glued to Games: How Video Games Draw us in and Hold us Spellbound*. Santa Barbara, CA: ABC-CLIO.

Ryan, R. M., Patrick, H., Deci, E. L., & Williams, G. C. (2008). Facilitating Health Behaviour Change and its Maintenance: Interventions Based on Self-Determination Theory. *The European Health Psychologist*, Volume 10.

Ryan, R. M. & Brown, K. W. (2005). Legislating Competence: The Motivational Impact of High Stakes Testing as an Educational Reform. in Elliot, A. E. & Dweck, C., *Handbook of Competence*, 354-374. New York: Guilford Press.

Ryan, R. M. & Deci, E. L. (2002). An Overview of Self–Determination Theory: An Organismic-Dialectical Perspective. in Deci, E. L. & Ryan, R. M., *Handbook of Self-Determination Research*, 3-33. Rochester: University of Rochester Press.

Ryan, R. M. & Deci, E. L. (2003). On Assimilating Identities to the Self: A Self–Determination Theory Perspective on Internalization and Integrity Within Cultures. in Leary, M. R. & Tangney, J. P., *Handbook of Self & Identity*, 253-274. New York: Guilford Press.

Ryan, R. M., Deci, E. L., & Williams, G. C. (2008). Facilitating Health Behavior Change and its Maintenance: Interventions Based on Self-Determination Theory. *The European Health Psychologist*, 10, 2-5.

Ryan, R. M., Kuhl, J., & Deci, E. L. (1997). Nature and Autonomy: An Organizational View of Social and Neurobiological Aspects of Self-Regulation in Behavior and Development. *Development and Psychopathology*, 9(4), 701-728.

Schlechty, P. C. (1997). *Inventing Better Schools: An Action Plan for Educational Reform*. San Francisco: Jossey-Bass.

Sheldon, L. (2011). *The Multiplayer Classroom: Designing Coursework as a Game*. Cengage Learning PTR

Smith, G. A. (2008). First-Day Questions for the Learner-Centered Classroom. *National Teaching & Learning Forum*, 17(5), 1-4.

Suits, B. & Hurka, T. (2005). *The Grasshopper: Games, Life and Utopia*. Broadview Press.

Strauss, W. & Howe, N. (1991). *Generations: The History of America's Future, 1584 to 2069*. New York: Morrow.

The Great Idea Finder. (2007). Retrieved June 22, 2015, from http://www.ideafinder.com/history/inventions/pong.htm.

Vansteenkiste, M., Lens, W., DeWitte, S., DeWitte, H., & Deci, E. L. (2004). The "Why" and "Why Not" of Job Search Behavior: Their Relation to Searching, Unemployment Experience, and Well-Being. *European Journal of Social Psychology*, 24, 345-363.

Villagrasa, S., Fonseca, D., Redondo, E., & Duran, J. (2014). Teaching Case of Gamification and Visual Technologies for Education. *Journal of Cases on Information Technology (JCIT)*, 16(4), 38-57.

Weinstein, N., Przybylski, A. K., & Ryan, R. M. (2009). Can Nature Make us More Caring? Effects of Immersion in Nature on Intrinsic Aspirations and Generosity. Retrieved April 14, 2013, from http://psp.sagepub.com.

Williams, G. C., Patrick, H., Niemiec, C. P., Ryan, R. M., Deci, E. L., & Lavigne, H. M. (2011). The Smoker's Health Project: A Self-Determination Theory Intervention to Facilitate Maintenance of Tobacco Abstinence. *Contemporary Clinical Trials*, 32, 535-543.

AUTHOR BIOGRAPHIES

Raymond D. Frost is Professor of Management Information Systems in the College of Business at Ohio University. He chairs the Teaching Continuous Improvement Team for the College. Raymond also is a leader in the Bruning Teaching Academy for Ohio University. He is currently involved in promoting Team Based Learning (TBL) at the University.

Vic A. Matta is faculty member in the Management Information Systems Department at the College of Business at Ohio University. His interests lie in pedagogy and consumer behavior, including adoption and motivation. His research has been published in *Journal of Information Systems Education*, *Journal of Computer Information Systems* and *International Journal of Information Management.* He is a member of the Special Interest Group of eBusiness (SIGeBIZ) and the *Association of Information Systems.*

Erin MacIvor earned her Bachelors of Science in Management Information Systems and Finance from the Honors Tutorial College at Ohio University in 2013. She was a teaching assistant, resident assistant, and student ambassador. Her research interests include motivation, learning styles, and Gamification. She is currently working as a Business Consultant for Hyland Software, Inc. in Cleveland, Ohio.

APPENDICES

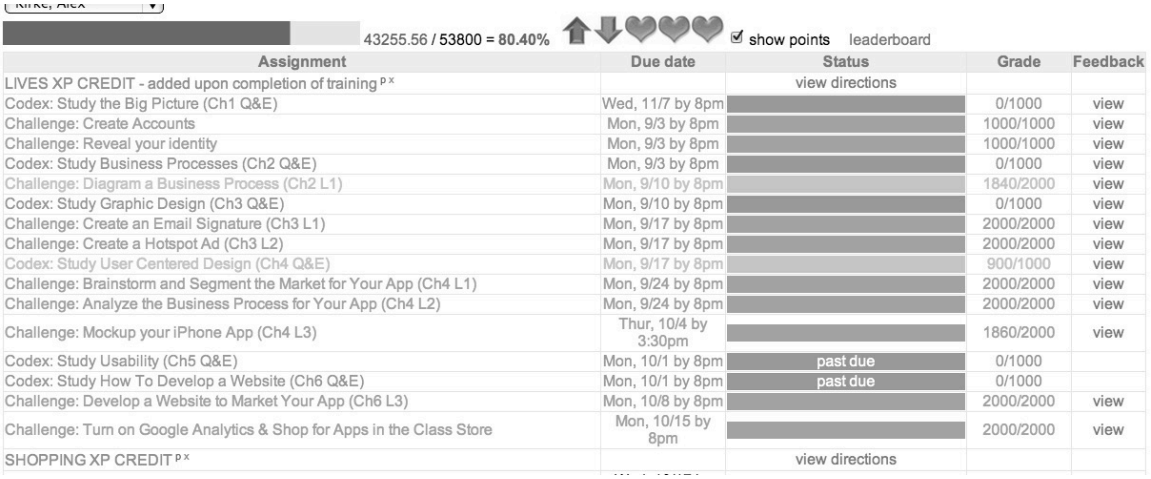
Survey

Respondents were asked to mark their opinions on a seven point scale ranging from "Not at all true" to "very true". Questions marked with (R) were reverse-coded.

Autonomy	Baard, Deci, and Ryan (2004)	• I feel like I have a lot of input in deciding how to complete my assignments. • I feel pressured in this class (R) • I am free to express my ideas and opinions in this class. • My feelings are taken into consideration in this class. • I feel like I can pretty much be myself in this class. • There is not much opportunity for me to decide for myself how to go about my work in this class. (R)
Competitive	Baard, Deci, and Ryan (2004)	• I do not feel very competent when I am in this class. (R) • I have been able to learn interesting new skills in this class. • Most days I feel a sense of accomplishment from this class. • In this class I do not get much of a chance to show how capable I am. (R) • When I am in this class I often do not feel very capable. (R)
Relatedness	Baard, Deci, and Ryan (2004)	• I really like the students I work with. • I get along with people in this class. • I pretty much keep to myself when I am in this class. (R) • I consider the people I work with in this class to be my friends. • People in this class care about me. • There are not many people in this class that I am close to. (R) • The people I work with in this class do not seem to like me much.(R) • People in this class are pretty friendly towards me.
Satisfaction	Ioannou and Artino (2009)	• Overall my learning experience using the ISMS system was positive. • I was satisfied with my learning experience using the ISMS system. • I would use this type of system if I were ever to teach a course. • I felt the ISMS system met my needs as a learner. • I would recommend this type of system if I ever had a friend who was teaching a course. • Overall, I enjoyed working with other students on the ISMS system.
Interest	Abrantes, Seabra, and Lages (2007)	• I was interested in learning the course material. • I was generally attentive in class. • I felt the course challenged me intellectually. • I have become more competent in this area of study.
Pedagogy	Abrantes, Seabra, and Lages (2007)	Overall, in this class the methods of instruction were: • Effective • Useful • Satisfactory • Good

Sample Screenshot of Gamified Interface

Gamified grade book of a student showing lives in the form of hearts, colored bars representing level of achievement (gray scale in this image), and a link to the leaderboard.



43255.56 / 53800 = 80.40% show points leaderboard

Assignment	Due date	Status	Grade	Feedback
LIVES XP CREDIT - added upon completion of training		view directions		
Codex: Study the Big Picture (Ch1 Q&E)	Wed, 11/7 by 8pm		0/1000	view
Challenge: Create Accounts	Mon, 9/3 by 8pm		1000/1000	view
Challenge: Reveal your identity	Mon, 9/3 by 8pm		1000/1000	view
Codex: Study Business Processes (Ch2 Q&E)	Mon, 9/3 by 8pm		0/1000	view
Challenge: Diagram a Business Process (Ch2 L1)	Mon, 9/10 by 8pm		1840/2000	view
Codex: Study Graphic Design (Ch3 Q&E)	Mon, 9/10 by 8pm		0/1000	view
Challenge: Create an Email Signature (Ch3 L1)	Mon, 9/17 by 8pm		2000/2000	view
Challenge: Create a Hotspot Ad (Ch3 L2)	Mon, 9/17 by 8pm		2000/2000	view
Codex: Study User Centered Design (Ch4 Q&E)	Mon, 9/17 by 8pm		900/1000	view
Challenge: Brainstorm and Segment the Market for Your App (Ch4 L1)	Mon, 9/24 by 8pm		2000/2000	view
Challenge: Analyze the Business Process for Your App (Ch4 L2)	Mon, 9/24 by 8pm		2000/2000	view
Challenge: Mockup your iPhone App (Ch4 L3)	Thur, 10/4 by 3:30pm		1860/2000	view
Codex: Study Usability (Ch5 Q&E)	Mon, 10/1 by 8pm	past due	0/1000	
Codex: Study How To Develop a Website (Ch6 Q&E)	Mon, 10/1 by 8pm	past due	0/1000	
Challenge: Develop a Website to Market Your App (Ch6 L3)	Mon, 10/8 by 8pm		2000/2000	view
Challenge: Turn on Google Analytics & Shop for Apps in the Class Store	Mon, 10/15 by 8pm		2000/2000	view
SHOPPING XP CREDIT		view directions		

STATEMENT OF PEER REVIEW INTEGRITY

All papers published in the Journal of Information Systems Education have undergone rigorous peer review. This includes an initial editor screening and double-blind refereeing by three or more expert referees.

ISSN 1055-3096