

Gradequest – Evaluating the impact of using game design techniques in an undergraduate course

Bob De Schutter
Miami University
800 E High Street
45056 Oxford, OH
USA
b@bobdeschutter.be

Vero Vanden Abeele
KU Leuven, Campus Group T
A. Vesaliusstraat 12
3000 Leuven
Belgium
vero.vandenabeele@kuleuven.be

ABSTRACT

The use of game design techniques in a non-gaming context - or 'gamification' [6]- offers the promise to make education more motivating and more enjoyable to students. This study reports on both the design and evaluation of gameful class (N= 17) that incorporates a variety of game design techniques through an online application named 'Gradequest'. The course was evaluated using three methods. First, a quantitative survey was used to collect data to measure levels of intrinsic motivation and engagement for the course. Second, a focus group session was held to obtain qualitative feedback from the students. Third, a teaching log was recorded to capture the instructor's perspective. The project concludes that the gameful instruction did not necessary lead to higher levels of intrinsic motivation or engagement in comparison to traditional teaching methods, and that further improvements to the design and documentation of the course are necessary. Furthermore, the project found that mediating factors (e.g., the role of the teacher, the clarity of the teaching materials, etc.) outweigh the gameful format in facilitating intrinsic motivation. The findings of the study are used to formulate recommendations towards the design of gameful instruction

Categories and Subject Descriptors

I.6.8 [Types of Simulation]: Gaming - K.3.1 [Computer Uses in Education]: Computer-Managed Instruction (CMI)

General Terms

Measurement, Design, Experimentation, Human Factors, Verification.

Keywords

Gamification, game design, education, learning, intrinsic motivation, course design

1. INTRODUCTION

The idea that game design techniques can effectively be applied elsewhere as a means of increasing user motivation, engagement, enjoyment, etc., resonates with people from many different academic and professional disciplines. As a result,

'gamification' has been a hot topic in the last couple of years. The Gartner's 2013 Hype Cycle Special Report - a subjective overview of the relative maturity of an innovation - lists gamification at the very top of its 'peak of inflated expectations' [10]. Even though a playful approach to the design of pre-existing services or products is not a new idea [19, 20], gamification is a relatively young term, and its best uses are not fully defined yet.

While gamification is still being hyped by many, the concept has not been without its critics. In his much-cited blog post, game scholar and designer Ian Bogost discusses how gamification misses the point. By focusing on a number of easy to implement aspects of games (e.g. high scores, experience points, badges and achievements, etc.), it fails to capture what makes games engaging: *"It confuses the magical magnetism of games for simplistic compulsion meted out toward extrinsic incentives."*[3].

In reaction to the typical implementation of gamification that tends to focus on a limited set of interface design patterns and reward systems (such as the aforementioned), both designers [21] as well as game scholars [9] have distanced themselves from the term. They propose to shift the focus of gamification to game mechanics (e.g. turns, limited resources, time constraints, etc.), design principles (meaningful choices, clear goals, enduring play, etc.) and game models (MDA[15], CEGE[4], etc.). By successfully implementing these game mechanics, principles and models, 'gameful design' can arguably provide some of the 'magical magnetism' that Bogost felt to be lacking in gamification.

Aside from debating which elements of game design should or should not be applied to non-gaming contexts, academic literature on the topic has also been concerned with the extent to which gamification is effective in improving motivation, engagement, enjoyment, etc. In their literature review, Hamari, Koivisto & Sarsa [13] identify 24 studies on the implementation of gamification. Within this sample, a range of game design techniques have been studied, some of which are focused on classic gamification (i.e., points, leaderboards, achievements, rewards, progress, feedback, etc.), while others include the recommendations of gameful design (levels, story, clear goals, challenge, etc.). The authors conclude that the results are partially positive, but also point out that gamification often seems to depend on several other factors, such as the motivations of users or the nature of the gamified system. As a result, it has been challenging for research to make claims that move beyond descriptive findings. Furthermore, the research indicates that the effect of gamification is mainly short-term, and could be caused by the novelty effect.

When considering the bigger academic picture, these findings are not that surprising. At its core, the implementation of gamification – as criticized by Bogost – relies on extrinsic motivation, which is materialized through its leaderboards, achievements and point systems. Motivation research has reached a general consensus that extrinsic motivation can lead to weak but positive short-term effects, and potentially detrimental effects to the individual's desire to perform the activity in the long run [2, 5]. Furthermore, while gameful design attempts to add more emphasis on intrinsic motivation to the promise of gamification by adding *meaningful* gameplay, it is not easy to generate the same intrinsic motivation or long-term engagement that a well-designed game can generate, let alone to achieve this for any kind of non-gaming context.

2. GAMEFUL INSTRUCTION

As gamification and gameful design have been applied in a wide range of services across the world, education has not lagged behind in adopting its principles. Many educators have started to apply the concepts in their courses, often resulting in experimental research. Hamari et al.'s literature review [13] identified 9 studies that are using gamification for learning and education, which made it the most prevalent context of implementation in the literature review. While the findings within this subset of studies was found to be mostly positive with regards to motivation, enjoyment and engagement in the learning tasks, the field of education brought along additional difficulties in its implementation of gamification. In particular, the authors refer to the possible effect of increased competition in the class room [12], difficulties in evaluating a task [7] and increased work load in doing so [24], and design problems that are unique to very specific implementations [8].

Stott & Neustaedter's analysis [29] extends the findings of Hamari et al. and presents 4 underlying dynamics and concepts that *"are shown to be more consistently successful than others when applied to learning environments"*: 1) freedom to fail, 2) rapid feedback, 3) progression, and 4) storytelling.

While providing an ongoing scoring system and a story – as suggested by Scott & Neustaedter – could be seen as a step in the direction of successful gamified course design, the approach should be embedded in a player-centered process in order to avoid the pitfalls of a waterfall design process [23]. Nicholson's [22] work provides such a user-centered theoretical framework, while also focusing on a variety of theories and concepts that emphasize the importance of freedom of choice and meaningfulness.

Finally, Kim & Lee's Dynamic Model for Gamification of Learning (DMGL) [18] provides a design model that is closer to acclaimed game design models such as the MDA framework [15]. Basing itself on both game design theory, instructional design and the influential work of Thomas Malone [19, 20], DMGL aims to maximize educational effectiveness through four primary aesthetics: challenge (e.g., clear fixed goals, uncertain outcomes, appropriate difficulty levels, etc.), curiosity (e.g., progressive unlocking of new content, time-based patterns, thrills, comedy, etc.), fantasy (storytelling, audio, visuals, etc.), and control (i.e., offering the player control over the 'game').

In summary, the literature currently seems to indicate that there is potential value in adding game design elements to educational courses, while at the same time emphasizing the many issues and complexities that need to be considered in order to design a course using game design techniques. This article describes a design research project that attempts to facilitate engagement

and intrinsic motivation among undergraduate students through the use of gameful instruction (i.e., instruction that adopts the principles of gameful design).

3. THE GRADEQUEST PROJECT

During the fall 2013 semester, two 3 credit hour undergraduate courses in a liberal education program were to be redesigned. The first course was a course on the principles of game design (N = 23; 4 female students, 19 male students), from now on referred to as the 'non-gameful course'. The second was a course on game design for educational purposes (N = 17; 7 female students, 10 male students), from now on referred to as the 'gameful course'. Across both courses, the students majored in a wide variety of disciplines, with the most prevalent ones being interactive media studies (12), computer science (7), and communication (4). Surprisingly, only one student majored in education.

Both courses shared a similar structure and had the same kind of assignments: they required the students to write reflective blog posts, participate actively in class, and turn analytical or game design related assignments in at similar points of the semester. There was also overlap in the course materials, as both courses discussed game design theories and methods, with one course diverging towards learning theory, while the other emphasized entertainment theory and game studies. Considering the similarities between both courses, the decision was made to apply gameful instruction to the educational game design course, while teaching the general game design course using a more traditional didactic approach.

It should be noted that – due to some practical inconveniences – the students did not have access to the actual syllabus of the educational game course while signing up, and did therefore not know about the gameful approach.

3.1 Course design

Aside from the design recommendations that were mentioned earlier in this paper, other literature on player motivation [e.g., 17, 28] media enjoyment [e.g., 30] and game design [e.g., 1, 25, 26] was used to drive the design of the course as well. The course design strived for as much 'gamefulness' as possible, i.e. striving for activities that are fun in their own right, without having to rely on external reward systems to motivate students. A particular inspiration here were the elaborate backstories or in-class racing games as described in Lee Sheldon's book on his 'multiplayer class rooms' [27]. Influenced by the pre-existing literature, the course was designed using

- heroes (fantasy alter ego's for the students),
- guilds (a different term for a group of students),
- quests (a different term for the course assignments),
- a backstory (occasionally told by instructor during class),
- experience points (XP; gained by successfully completing quests and transferred to a grade at the end of the semester),
- character levels (based on the amount of XP a student gained),
- character skills (in-class super-powers chosen when reaching a certain level), and
- leaderboards (high-score tables).

The course offered 5 types of quests. *Main quests* were quests that students were required to take, such as 'a hero is born' (i.e., designing an avatar), 'a champion rises' (i.e., playing a custom

made version of LocalNo12's metagame¹), 'the gauntlet' (i.e., a playful midterm reminiscent of Hasbro's Taboo² game), 'Challenge the Fallen Ones' (i.e., present your project in front of an expert jury), etc. In addition; the students were also required to do two *side quests*, for which they were asked to choose 2 of 6 possible side quests (ranging from writing academic papers or creating an instructional YouTube video, to developing an educational game design).

Aside from the mandatory main and side quests, the course offered a number of *optional quests*. The students were encouraged to pursue repeatable 'grind' quests – such as 'on guard' (i.e., attendance and participation) and 'battle cry' (i.e. reflective blog posts) – and optional bonus quests. The bonus quests were very similar to achievements, as they were awarded to students who put in some extra effort. For example, 'uncovering the code' awarded students a small amount of extra XP if they wrote their blog post in the writing style of the week (ranging from 'challenging palindromes' to 'battle raps' or 'valley girl style'). Some of these bonus quests were also specifically designed to make the guilds more meaningful. For example, the 'First to Five' quest awarded extra XP to the first guild whose members all accumulated enough experience points to reach 'level 5' (see below). As a result, the quest was designed to encourage students to help each other with the course materials.

Finally, the course provided 4 *random quests*, which were quests that randomly took place at the beginning or at the end of a session. An example of a random quest was 'Goblins, I Tells Ya', a quiz on the materials that had to be read prior to class. Another example was 'Challenge the Sword Master', a quiz at the end of class that pitted the guilds against each other in a battle of knowledge about the materials that were presented in class. For these quests students had to roll dice, the outcome of which would change the quest's story and the difficulty level of the quest.

While the quests were designed as actual games as much as possible, an extra layer of gameful design was added through the character skills. When players leveled up they were allowed to pick skills that had consequences in class. For example, a 'bodyguard' hero could jump in for another guild member if he or she did not know the answer to a question during the goblins random quest. A 'dodger' hero could pass on a single question during a random quest and request a different question instead. As a result, the skills served two purposes: 1) flesh out the hero character for each student (to provide a meaningful fantasy), and 2) add an element of (controlled) chance to the quests to facilitate curiosity.

The evaluation of student work determined how much XP a student would get for a quest. The course offered 5 possible results:

- "Wipe" – The student did not do the assignment, or completely failed to meet the rubric standard.

- "Barely acceptable" – The student did the assignment, but the rubric standard is not fully met.
- "Acceptable" – The student met the rubric standard.
- "Good" – The student exceeded the expectations.
- "Exceptional" – The student exceeded the expectations to the extent that the work sets a new standard.

Depending on its relative importance, each quest was assigned a certain amount of XP for each of the 5 possible outcomes, with a 'wipe' often being 0 XP. Prior to the course, several simulations were run to determine which amount of XP would be appropriate for which final letter grade. A student who would receive "Good" evaluations on the most important quests would end in the B+ to A range.

3.2 Software Development

While the course design implemented many of the design recommendations in the literature (e.g., freedom of choice, fixed goals, storytelling, playful challenges, etc.), extra steps still needed to be taken to provide direct feedback as well as a sense of progression for the students.

Aside from GradeCraft – a learning management system by Holman, Aguilar and Fishman [14] that was unfortunately not operational at the time – no software applications were identified that would allow to achieve both design goals. Therefore, a custom built system was built: Gradequest is a PHP-based jQuery Mobile application that offers a back-end (allowing to grade the students and view their grades and skills) and a front-end that allowed the students to access the following main features:

- "My Hero" – A personal profile page where students could name their avatar, provide the avatar with a biography, buy skills using their experience points, see the amount of dice they are allowed to roll on quests, and see their current level and the amount of XP needed for the next level.
- "My Quests" – An overview of every quest the student did, along with its evaluation.
- "My Guild" – A group-based high score that showed which guild relatively accumulated the most XP.
- "Hall of Honor" – A leaderboard that showed the current ranking and the level of each avatar.

Some specific considerations had to be factored into the design of Gradequest, even before the semester started. For legal reasons, Gradequest could only be used as long as students would not be able to figure out each other's grades. Therefore, the Hall of Honor was capped at level 10 (even though students would accumulate experience points beyond level 10), and the overview provided by "My Guild" only showed the percentage of XP each guild had accumulated in relation to the total amount of XP as gained by the entire class.

While Gradequest's visual style was minimalistic, the application was fully usable on a smartphone or a tablet. This made handing out XP easier, which proved to be important considering the significantly higher work load that came with the gameful design of the course.

¹ The metagame is a card game in which players have to argue why the game on their card (e.g. Tetris) is better at a random comparison (e.g. which has the bigger affinity group?) than the game on their opponent's card (e.g. World of Warcraft).

² Taboo is a game in which one player explains a term without actually naming it. The other players attempt to guess the word.

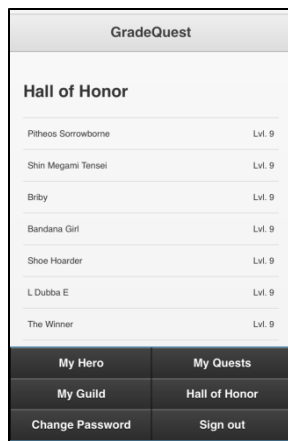


Fig. 1 - Screenshot of the 'Hall of Honor' on a smartphone

3.3 Evaluation

As Gradequest went through its first iteration during the semester, an evaluative study was in order. The research questions for the study were:

1. How does self-reported intrinsic motivation and engagement of students differ from the non-gameful course?
2. Which game design elements improve/worsen students' self-reported engagement, enjoyment and motivation?
3. How can the course design (as well as Gradequest) be improved?

Mixed methods were used in order to answer these questions. First, the students were asked to provide informal feedback whenever they saw fit. They could do this by talking directly to the instructor, but in addition, an online feedback form (using freesuggestionbox.com) was provided so they could provide feedback anonymously without having to fear any repercussions.

At the end of the semester, two sessions of the educational game design course were devoted to evaluating the gameful design of the course. The evaluation was done using both quantitative and qualitative methods. The quantitative part was done using a Qualtrics survey. The questionnaire consisted out the Situational Motivation Scale (SiMS) [11] and the core module of the Game Experience Questionnaire (GEQ)[16].The SiMS measures the following motivational concepts:

- intrinsic motivation (Cronbach's $\alpha=.870$; i.e., performing an activity for itself),
- identified regulation ($\alpha=.852$; i.e., performing a valued activity as a means to an end),
- external regulation ($\alpha=.831$; i.e., performing an activity for external rewards), and
- amotivation ($\alpha=.884$; i.e., an activity that is neither intrinsically nor extrinsically motivated).

The GEQ measures the following concepts:

- imaginative and sensory immersion ($\alpha=.830$; e.g. "I felt like a rich experience"),
- annoyance/tension ($\alpha=.956$, e.g. "I felt irritable"),
- flow ($\alpha=.776$; e.g. "I was fully occupied"),
- competence ($\alpha=.892$; e.g., "I was good at it"),
- positive affect ($\alpha=.854$;e.g., "I enjoyed it"),
- negative affect ($\alpha=.764$; e.g., "I was bored"), and

- challenge ($\alpha=.770$; e.g., "I had to put a lot of effort into it.").

Aside from these instruments, the questionnaire used 7-item Likert scales to evaluate to which extent the various design elements of the course led to enjoyment (e.g., "I enjoyed the XP-based grading system"), engagement (e.g., "The XP-based grading system was engaging."), motivation (e.g., "The XP-based grading system motivated me.").The questionnaire also asked if the students would prefer to have taken the course without the gameful elements (e.g. "I would have preferred to take the course without the XP-based grading system.").

The students were also asked about their playing behavior, identity as gamers, prior interest in the topic of the course, and how they would evaluate the course using common course evaluation questions (e.g. the instructor is an excellent teacher, the course materials were clear to understand, the course helped me develop competence, etc.). Finally, the majority of students entered their student IDs, so that their grades could be attached to their answers.

After the students filled out the questionnaires, a focus group session was held. This session was informed using the course materials concerning gamification. When students made subjective remarks about the course, a 'clicker' survey was launched (using the script that is provided at clicker.math.ntnu.no), which allowed to quickly survey the extent to which the entire classroom agreed with the statement.

Aside from gathering quantitative and qualitative data, the instructor kept track of his own perceptions of the course, the students' behavior in class, and any changes made to Gradequest, in a teaching log. This was done to provide the teacher perspective to the findings of the study.

4. FINDINGS

4.1 Summary of the Teaching Log

The gameful course started out with a mixed response by the students. The approach seemed a bit geeky for a subgroup of the students (e.g., they had to create hero avatars and sessions started with a back story about how the heroes traversed a fictional world) Nonetheless, after the first class seven students came up to voice how "awesome" they felt that this approach was. The first impression of how well gameful design would work well for this class was therefore one of mixed feelings.

One student also talked about his learning disability in (social) stress situations and verified whether this gameful approach would be able to accommodate his disability.. He expressed his fears, considering the social aspects of the gameful design of the course. However, the student also expressed a desire to stay in the course, in particular because of the experimental design. By the end of the course, it turned out that there was no need for accommodations. The student even participated in some stressful social evaluation moments in front of the class on road to a good grade.

A problem arose in the first couple of weeks, as it became evident that only a small group of students had read the syllabus and rule book as was specifically requested in the first sessions. As a result, they had absolutely no idea about which quests awaited them, what was expected of them, or how they would be evaluated. An extra session was organized to take the students through the texts paragraph by paragraph, as this seemed to be the only way to make sure that the students would get accustomed with the material at this point in the semester.

Two students dropped from the gameful class (bring the total from 19 to 17). The first student said that he did not know how big the workload would be prior to registering for the class, while the second student dropped from the university due to personal reasons.

During the first half of the semester, there were no major problems with the course design, although one minor issue occurred. The semi-random group assignments (i.e., students were asked about what skills they would bring to a game design team and were assigned to guilds according those to make sure that every guild had all the skills they needed to succeed) led to internal problems within some of the groups. It was agreed with these groups to resolve after the midterm.

By the middle of the semester changes needed to be made to the 'battle cry' quest (i.e. blog posts). At the beginning of the year it was designed as a competition in which each student was required to read the other students' blog posts and vote (online) for the ones they felt to be the best ones in correspondence to a rubric. This seemed to work well in the beginning of the semester, but at some point the voting responses decreased and it became apparent that some students were gaming the system by voting on each other without even checking if everyone actually wrote a blog post. The voting was therefore removed.

The optional writing styles that were intended to make the blog post assignments more fun were not very successful. While some students did them on and off, only 2 students did them each week. Furthermore, they seemed to genuinely have fun doing them. It is very inspirational for an instructor to see students write informed battle raps about their course work.

The other quests seemed to work as intended. In particular, the random quests seemed to engage the students, especially when they were given meaningful choices that are embedded in a backstory. The 'Goblins' quests is an example of this. In the quest students can earn an amount of XP - that is too small to be relevant towards the final grade, but big enough to be motivating - by answering three pop quiz questions (i.e., three goblins) correctly. During the quest, students could try to sneak past the goblins for little but almost guaranteed XP. If they decided to do so, then - with a lucky die roll - they would not be subjected to any questions and still receive a small amount of XP. However, if the die roll was not successful, then the students would still have to answer the quiz questions for even less XP. Thankfully, some students often came prepared and would simply decide to charge the goblins head first, which gave them a short at the maximum amount of XP that was allocated to the quest. While the story elements and the ability to make meaningful choices worked well, the special powers seemed to work even better. Students would activate them to help their guild members as well as themselves (e.g., by getting a +1 on their dice rolls or to get rid of a particularly tough question).

It took a while for some of the students to realize that the course required them to come to class prepared (by studying the course materials), and as a result, they would get 'wiped' (fail) for some of the random quests. Even though the course was designed to allow students to fail some of the random quests (and make up for it through optional quests), some students were upset with their failing in the random quests. Nevertheless, the results of the random quests steadily improved throughout the semester, as students learned that they needed to be prepared for the challenges that were presented to them in class.

The midterm was an important part of the course. At this point, students needed to be able to understand the core course

materials well, in order for them to be able to apply them in the second part of the semester. The midterm was designed as a game, similar to Taboo. Each student would go in front of class and explain a term from the course materials, while every other student needed to identify the term. In the end, a total score was calculated based on how many terms each student got right, how many students got the term right that the student explained himself, and how well the student did on a short non-gameful writing assignment (which was included as a backup plan in the event that the midterm game would fail). The results of the midterm varied widely. Four students failed, while 5 students received a "Good" or "Exceptional", with one student almost obtaining a perfect score. None of the students who failed the midterm dropped the course, and indicated that they would improve their efforts for this course. All of them argued that they "needed to pass to course in order to graduate on time".

After the midterm, the issues with some of the guilds were resolved by changing the rules and rubrics slightly. From this point forward, the students were allowed to do the main and side quests by themselves if they wanted to. In order to make sure that these students would still be able to participate in quests that were designed for guilds, they were assigned to a special 'solo-only' guild. Story-wise, the students' heroes became mercenaries for hire who sometimes fought together in the so-called 'guild of calamitous intent'. This change was accompanied with the option for the students who still wanted to do the quests as part of their own guild, to leave their old guild and reform new ones. This was met with positive responses by a large majority of the students.

At this point, the students had requested and were offered additional structure so they could gauge their progress better. While the students had access to all the information needed to know where they stood in terms of obtaining a certain grade at the end of the year, a majority of students did not do the necessary calculations to actually figure this out. Gradequest therefore received a 'projected XP' feature, that showed the students how much XP there was still left to earn, and how much of that XP they would earn if their evaluations for each type of quest would remain the same. Aside from adding a predictive algorithm to Gradequest, the students received a 'road map to a B' document that showed exactly what was expected to get a 'B' letter grade at the end of the year and what options there were if they would fail certain quests.

On route to the end of the semester, two more significant events occurred. First, the story at the beginning of the courses was dropped, as the students seemed to be less and less interested in it. In the beginning, the story elements were met with enthusiasm (especially when students were allowed to interject random location names in the story), but after the midterm students seemed to have lost interest in it, barely paying attention to it.

Second, during one of the last weeks, the class played a custom version of the metagame (see 3.1 Course Design) which was highly successful. Students were given five cards with games on them and they were told to make sure that they knew these games inside out. In groups of three, students debated why their games were best at implementing a randomly drawn learning theory. One student refereed the two other students and passed on the result to the instructor, before moving on to the next group. The event ended with a final battle between the two students with the best winning record in front of class.

After wrapping up the semester, it seemed that the course was successful in meeting its goals. In particular, the quality of the metagame discussions as well as some of the final projects was impressive. In comparison to the non-gameful course, there were little differences in the outcome - recall that the courses were quite similar in their learning content and objectives. In particular, the extent to which the students mastered the materials seemed similar. Even though the gameful course had more intended playfulness to it by doing playful activities with the students, the non-gameful course had plenty of impromptu playfulness to it as well, due to the very nature of any course that is about digital games. As a result, students of the gameful educational game design course did not seem more motivated, engaged or entertained than the students that followed the course on principles of game design.

4.2 Suggestion Box

The anonymous online suggestion box collected five anonymous responses by the middle of the semester. One student complained about how her guild let her down which resulted in her gaining less XP than she could have. Another student felt the wide variety in options were overwhelming and that the course could use more structure. Both issues were addressed in the middle of the semester, by allowing students to 'solo' the side quests (as already discussed in the "teaching log" section of this paper), and by providing extra documentation on the course schedule and specific quests.

Aside from these two issues, the comments did not address the gameful design of the course. Instead, students used the anonymous feedback system to address issues they had with other students in the class, to discuss their perceived difficulty level of the course (too hard for some/too easy for others), and to comment on my teaching style. However, 3 of the 5 comments specifically mentioned that they enjoyed the class as well as its gameful design.

4.3 Focus Group

The focus group followed a similar pattern as the suggestion box, in the sense that a good deal of comments were not about the gameful design of the course. The students mainly used the session to suggest changes to the schedule, to how the course information is presented, and to the rubrics. Nevertheless, from these comments, there were some valuable recommendations for the gameful design of the course. First, the students suggested to 'chain' the side quests and the final quest to each other, so that work of the side quests furthers the final quest. While there was nothing preventing from them doing this already, it would be better to have this pointed out more. Second, some students wanted the avatars and story to have a bigger role in the quests and in the course in general. Finally, some students asked for more quests that awards XP for collaboration.

4.4 Survey

4.4.1 Evaluating the Gameful Design Elements

The students reported to check Gradequest about 2 times per week ($M = 2.3$, $SD = 1.4$), which seems to coincide with the 2 sessions of class per week.

Table 1. Evaluations of the Gameful Design Elements

	Motivate	Enjoy	Engage	Without	Overall
Choice of Side Quests	5.3 (1.3)	5.3 (1.2)	5.9 (0.9)	2.3 (1.1)	5.55

Metagame	4.7 (2.1)	5.2 (1.6)	5.0 (2.1)	n/a ³	4.97
Special Skills	4.6 (2.3)	4.8 (2.2)	5.1 (1.5)	2.8 (1.7)	4.93
Achievements	4.8 (2.1)	4.5 (2.1)	5.2 (1.6)	3.2 (1.5)	4.82
Leaderboards	4.7 (2.1)	4.5 (2.0)	4.9 (1.9)	3.0 (1.5)	4.78
Story Elements	4.0 (2.0)	4.6 (1.9)	4.6 (1.8)	3.1 (1.6)	4.52
Blog - Writing Styles	4.1 (2.2)	4.2 (2.4)	5.2 (1.5)	3.5 (1.8)	4.50
XP grading	4.5 (1.8)	4.1 (1.8)	4.6 (1.8)	3.5 (1.8)	4.42
Random Quests	3.8 (1.9)	3.3 (2.2)	4.2 (2.2)	3.3 (1.7)	4.00
Guilds	4.2 (2.0)	4.3 (2.1)	4.1 (2.4)	3.6 (2.1)	4.00
Midterm (Taboo)	3.7 (2.3)	3.5 (2.2)	4.3 (2.4)	3.6 (2.4)	3.98
Blog - Polls	3.2 (2.2)	3.1 (2.1)	4.8 (1.8)	4.4 (1.8)	3.68

Table 1 provides the means and standard deviations (the latter between brackets) for the extent to which students reported to enjoy specific design elements, or the extent to which they were motivated and engaged by them. The 'without' column indicates whether or not the students would have preferred to take the course without the specific gameful element to it. As the survey used 7-item Likert scales, values below 4 are in disagreement, while values above 4 are in agreement with the statements.

In general, the results suggest that students appreciated the various gameful features (or games) that were part of the course. When calculating aggregate scores, being able to choose your side quest, the metagame, and the special skills scored the highest. The polls (i.e., voting on who had the best blog post) and the midterm were the only feature with a mean value below the neutral value of 4. The blog polls were also the only feature that the class preferred to have done without.

Aside from the blog polls, the midterm and the random quests received negative scores with regards to motivation and enjoyment. This could be due by the fact that both are closely related had a strong tie to the student's grades. It is plausible that enjoyment decreases when turning a quest into an evaluation, and failing to perform well at a (somewhat important) quest could be demotivating. Both elements also required the students to demonstrate their knowledge on the course materials in front of the class, which could be another reason for the lower means.

4.4.2 Comparing Both Courses

The survey also compared the gameful course with the non-gameful course, on measures of motivation (i.e., SiMS) and engagement (i.e., GEQ). Kolmogorov-Smirnov tests indicated that the distributions for all variables did not deviate from a normal distribution, and *t*-tests revealed two significant differences between the gameful and the non-gameful course. The non-gameful course scored significantly higher on intrinsic motivation ($t(16.163) = 2.802$, $p < .05$) and significantly lower on tension/annoyance ($t(22) = 2.942$, $p < .05$).

Table 2. Analysis of the SiMS and GEQ measures (t-test)

	Gameful	Non-Gameful	<i>t</i>
SiMS - Intrinsic Motivation	4.4 (1.5)	5.7 (0.7)	2.802
SiMS - Identified Regulation	4.6 (1.7)	5.5 (1.0)	1.732

³ The extent to which students would prefer not to have done the midterm was not asked in the questionnaire.

SiMS - External Regulation	3.3 (1.3)	3.6 (4.6)	.569
SiMS - Amotivation	3.0 (1.6)	2.5 (1.1)	1.151
GEQ - Tension/Annoyance	5.0 (1.3)	3.3 (1.3)	2.942
GEQ - Negative Affect	4.3 (1.2)	3.3 (1.2)	1.800
GEQ - Flow	3.5 (1.1)	4.1 (1.0)	1.376
GEQ - Immersion	4.7 (0.7)	5.1 (1.0)	.907
GEQ - Positive Affect	4.7 (0.9)	5.1 (0.9)	.886
GEQ - Challenge	5.0 (1.5)	5.1 (1.1)	.119

An analysis of covariance (ANCOVA) was performed to further explain the significant differences. The following covariates⁴ were entered in the analysis:

- *teacher effectiveness* (i.e., “The instructor is an effective teacher”; this covariate is included as the instructor had more experience teaching one of the two courses),
- *classroom atmosphere* (i.e., “I enjoy to be in the same class as the other students.”; this covariate is included as there was tension between some of the students in the gameful course),
- *clarity of the course* (i.e., “The objectives, expectations, requirements and content of the course were clear to me.”; this covariate is included as the students specifically asked for more structure during the semester),
- *competence development* (i.e., “The course helped me to develop competence.”; this covariate is included as to measure if students felt that the course was effective),
- *prior interest* (i.e., “I registered for the course because it matches my interests”; this covariate is included to control for people with an affinity towards the topic), and
- *playing time* (i.e. “the self-reported amount of hours the students play games per week”; this covariate was included to control for gamers).

The resulting model explained 83.6% of the total variance and the difference in intrinsic motivation between both courses disappeared ($F(0.335,1) = 4.688, p = n.s.$). The significant factors in the model were teacher effectiveness ($F(1.753, 1) = 8.838, p < .01$) and the clarity of the course ($F(1.711, 1) = 8.626, p < .05$). This finding can be explained considering the instructor had been teaching the non-gameful course for multiple years before, while it was his first time teaching the gameful course. (While the ratings for teacher effectiveness were still positive, the difference was significant.) The difficulties that some students had with the clarity and structure of the gameful course’s format were also documented in the qualitative feedback.

The annoyance/tension difference was no longer significant either. The analysis indicated that the difference could have been caused by the class atmosphere ($F(5.328,1) = 4.481, p < .05$), which is in line with some of the frustrations that were voiced

concerning the group work in the gameful course (see section 4.1).

4.5 Other Measures

Finally, the study compared the extent to which the students did their blog posts assignments in order to see if the gameful course managed to get better responses. However, significant difference between the amount of blog posts the students did in the two classes could be found ($t(25) = .216, p = n.s.$).

5. DISCUSSION

The first iteration (or pilot) of the gameful course was successful in revealing interesting dynamics. At first sight, it seemed that the gameful course relied too heavily on extrinsic motivation (e.g., leaderboards, experience points, etc.) in comparison to the non-gameful course. However, since the non-gameful course also used letter grades throughout the year, the difference in how students are rewarded (intrinsically versus extrinsically) should be minimal. If anything, the students in the gameful course received much more timely feedback.

Further analysis revealed that the lower reported intrinsic motivation could be due to a lower opinion of teacher effectiveness and less overall clarity concerning the expectations. The lower rating for teaching effectiveness could be related to disappointing grades during the midterm (as 4 students were ‘wiped’ in the gameful course as opposed to only one ‘F’ in the non-gameful course). The lower rating for clarity can be understood when comparing both course syllabi. The gameful course’s syllabus is substantially longer due to it having to explain the gameful elements of every quest similar to the rulebook of an actual (board) game. The clarity issues were also voiced across the various feedback channels.

Even though the motivational factors were not significant after controlling for the role of the teacher and other factors, the study did show how certain design elements could be related to higher engagement, motivation or enjoyment than others. In particular, based on these findings and the way the students responded to the course in class, the following recommendations should be made:

- *Know who you are teaching.* The extent to which gameful activities work depends for a large part to how the students are willing to sign up for them.
- *Provide students with freedom of choice in how they want to show their mastery of the materials.* Some of the most impressive and insightful side quests for this course ended up being YouTube videos.
- *Special skills and achievements go a long way.* The classes were brought to life when students could use their superpowers to influence the outcome of a class.
- *Provide extra structure.* The gameful course had the same amount documentation as the non-gameful course, but as the gameful course is an experimental format, students need extra explanation and structure.
- *Evaluate their knowledge of the game.* Do a quest that forces students to study the rulebook carefully.
- *Real-world quest names.* While it is fun to have fantasy names for quest, the students asked to add non-fantasy elements to the names. For example, a “Battle Cry Blog” quest is much clearer than just “Battle Cry”.
- *Communicate that the course will be challenging.* Some students were expecting an easy course since it

⁴ As the students’ grades were not recorded at the time of writing, we could not include them in the analysis.

was “a course about games that was designed as a game”.

- *Location matters.* To engage the students, it is best to have a classroom that accommodates this. (Also see [27].) As the classroom were the courses were taught in was a relatively deep and small, it made it more difficult to have students come up to the front of the class. Furthermore, the tables were positioned against the walls with computers on them, which not optimal for collaborations.
- *Prepare for a lot of work.* Turning every aspect of a course into a fun game and providing constant feedback while doing so, is very labor intensive.

While the evaluations led to interesting results, it should be noted that this was the first iteration of the courses, and the sample sizes were small (i.e., 39 students across both courses). In the end, the goal of this study is not to make generalizable claims, but to explore differences the specific design of this course. Furthermore, it is not too surprising that the study highlights the importance of the teacher and clarity of the course materials as being more important than the teaching format itself.

At the end of this first iteration, a lot of questions still remain. How do student personalities influence the student experiences? How would it be possible to accommodate a student with social and stress-related learning abilities for a gameful course? What happens if the experience points are not tied to a specific grade at the end of the year? Hakulinen et al. [12] found that behavior can be affected even when the badges have no impact on grading, so is it even necessary to have experience points? We are hoping to address some of these questions in the next iterations of these courses, and to contribute to a body of work on this topic that might lead to a manual for successful gameful instruction.

6. ACKNOWLEDGEMENTS

The researchers would like to thank the Armstrong Institute for Interactive Media Studies in their support to produce and publish this research. The study met the exemption criteria of Miami University’s Institutional Review Board.

We would also like to thank the students. Without their willingness to be guinea pigs and their support when things did not go as planned, this project would not have been possible. A ‘thanks’ also goes to Per Kristian Hove for providing me with his ‘clicker’ script.

7. REFERENCES

- [1] Adams, E. and Dormans, J. 2012. *Game mechanics: advanced game design*. New Riders.
- [2] Bénabou, R. and Tirole, J. 2003. Intrinsic and Extrinsic Motivation. *The Review of Economic Studies*. 70, 3 (Jul. 2003), 489–520.
- [3] Bogost, I. 2011. *Persuasive games: exploitationware. Gamasutra: The art & business of making games*. (2011).
- [4] Calvillo-Gámez, E.H., Cairns, P. and Cox, A.L. 2010. Assessing the Core Elements of the Gaming Experience. *Evaluating User Experience in Games*. R. Bernhaupt, ed. Springer London. 47–71.
- [5] Deci, E.L. 1975. *Intrinsic Motivation*. Plenum Press.
- [6] Deterding, S., Dixon, D., Khaled, R. and Nacke, L. 2011. From Game Design Elements to Gamefulness: Defining

“Gamification.” *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (New York, NY, USA, 2011), 9–15.

- [7] Domínguez, A., Saenz-de-Navarrete, J., de-Marcos, L., Fernández-Sanz, L., Pagés, C. and Martínez-Herráiz, J.-J. 2013. Gamifying learning experiences: Practical implications and outcomes. *Computers & Education*. 63, (Apr. 2013), 380–392.
- [8] Dong, T., Dontcheva, M., Joseph, D., Karahalios, K., Newman, M. and Ackerman, M. 2012. Discovery-based games for learning software. *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems* (2012), 2083–2086.
- [9] Fishman, B.J. and Deterding, S. 2013. Beyond Badges & Points: Gameful Assessment Systems for Engagement in Formal Education. (University of Wisconsin, Wisconsin, Madison, 2013).
- [10] Gartner’s 2013 Hype Cycle for Emerging Technologies Maps Out Evolving Relationship Between Humans and Machines: 2013. <http://www.gartner.com/newsroom/id/2575515>. Accessed: 2013-12-04.
- [11] Guay, F., Vallerand, R.J. and Blanchard, C. 2000. On the Assessment of Situational Intrinsic and Extrinsic Motivation: The Situational Motivation Scale (SIMS). *Motivation and Emotion*. 24, 3 (Sep. 2000), 175–213.
- [12] Hakulinen, L., Auvinen, T. and Korhonen, A. 2013. Empirical Study on the Effect of Achievement Badges in TRAKLA2 Online Learning Environment. *Learning and Teaching in Computing and Engineering (LaTiCE)*, 2013 (2013), 47–54.
- [13] Hamari, J., Koivisto, J. and Sarsa, H. 2014. Does Gamification Work?—A Literature Review of Empirical Studies on Gamification. *Proceedings of the 47th Hawaii International Conference on System Sciences* (Hawaii, USA, 2014).
- [14] Holman, C., Aguilar, S. and Fishman, B.J. 2013. GradeCraft: what can we learn from a game-inspired learning management system? *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (2013), 260–264.
- [15] Hunnicke, R., LeBlanc, M. and Zubek, R. 2001. MDA: A Formal Approach to Game Design and Game Research. (2008 2001).
- [16] IJsselstein, W., van den Hoogen, W., Klimmt, C., de Kort, Y., Lindley, C., Mathiak, K., Poels, K., Ravaja, N., Turpeinen, M. and Vorderer, P. 2008. Measuring the experience of digital game enjoyment. *Proceedings of the 2008 Measuring Behavior Conference* (2008).
- [17] Jansz, J. and Tanis, M. 2007. Appeal of playing online First Person Shooter Games. *Cyberpsychology & behavior*. 10, 1 (2007), 133–6.
- [18] Kim, J.T. and Lee, W.-H. Dynamical model for gamification of learning (DMGL). *Multimedia Tools and Applications*. 1–11.
- [19] Malone, T.W. 1980. What makes things fun to learn? heuristics for designing instructional computer games. (Palo Alto, California, United States, 1980), 162–169.

- [20] Malone, T.W. and Lepper, M.R. 1987. Making learning fun: A taxonomy of intrinsic motivations for learning. *Aptitude, learning, and instruction*. 3, (1987), 223–253.
- [21] McGonigal, J. 2011. We Don't Need no Stinking Badges.
- [22] Nicholson, S. 2012. A user-centered theoretical framework for meaningful gamification. *Proceedings GLS*. 8, (2012).
- [23] Royce, W.W. 1970. Managing the development of large software systems. *proceedings of IEEE WESCON* (1970).
- [24] Rozeboom, R. 2012. *Professor Clifford Lampe Talks about Gamification in Academia*.
- [25] Salen, K. and Zimmerman, E. 2003. *Rules of Play: Game Design Fundamentals*. MIT Press.
- [26] Schell, J. 2008. *The Art of Game Design: A book of lenses*. Morgan Kaufmann.
- [27] Sheldon, L. 2011. *The Multiplayer Classroom: Designing Coursework As a Game*. Course Technology/Cengage Learning.
- [28] Sherry, J.L., Lucas, K., Greenberg, B.S. and Lachlan, K. 2006. Video game uses and gratifications as predictors of use and game preference. *Playing video games: Motives, responses, and consequences*. (2006), 213–224.
- [29] Stott, A. and Neustaedter, C. Analysis of Gamification in Education.
- [30] Vorderer, P. and Hartmann, T. 2009. Entertainment and Enjoyment as Media Effect. *Media effects: advances in theory and research*. J. Bryant and M.B. Oliver, eds. Routledge. 532 – 550.