Improving Participation and Learning with Gamification

Conference Paper · October 2013

DOI: 10.1145/2583008.2583010

4 authors:

Gabriel Barata
Inesc-ID

Sandra Gama
Inesc-ID

Joaquim Armando Pires Jorge
Inesc-ID

Daniel Gonçalves
Inesc-ID

Improving Participation and Learning with Gamification

Gabriel Barata, Sandra Gama, Joaquim Jorge, Daniel Gonçalves
Dept. of Computer Science and Engineering
INESC-ID/IST/Technical University of Lisbon
Lisbon, Portugal
{gabriel.barata, sandra.gama}@ist.utl.pt, {joaquim.jorge, daniel.goncalves}@inesc-id.pt

ABSTRACT

In this paper we explore how gamification can be applied to education in order to improve student engagement. We present a study in which a college course was gamified, by including experience points, levels, badges, challenges and leaderboards. The study was five years long, where the first three were non-gamified years, and the last two regarded two successive experiments of our gamified approach. To assess how gamification impacted the learning experience, we compared data from both gamified and non-gamified years, using different performance measures. Results show significant improvements in terms of attention to reference materials, online participation and proactivity. They also suggest that our approach can reduce grade discrepancies among students and help them score better. Modeling course activities with game challenges and properly distributing those over the term seem to enhance this effect.

Author Keywords

Gamification; Education; Evaluation; Student Participation; Virtual Learning Environments

ACM Classification Keywords

H.5.2. User Interfaces: Evaluation/methodology

General Terms

Measurement; Experimentation; Human Factors

INTRODUCTION

In the last decade, the use of technology to improve learning and education has been widely explored, as a means to improve instruction delivery. It potentiated the emergence of learning experiences, like blended learning [11] and distance education [2]. Prominent examples of these are flipped classrooms [24], where content is delivered online and students work in class to solve problems, supervised by an instructor, and Massive Open Online Courses [17], which deliver class materials and lectures as online resources, and students may discuss and get help in interactive forums. Most of these learning advances rely on distributed systems to make resources available remotely, without any further efforts to make the experience more engaging and rewarding for students.

Gamification'13, October 2 – 4, 2013, Stratford, ON, Canada.

Education is also being shaped by the use of other techniques like gamification: adding game elements to non-game contexts [9, 10]. It is used to engage users to embrace specific behaviors, like working out or saving money. Gamification relies on the motivational power characteristic of good games which, unlike traditional learning materials, can deliver information on demand and within context [12], balancing challenge difficulties according to one's abilities. This prevents players from becoming bored or frustrated, and allows them to experience flow [3, 6]. Gamification has been used in many domains, such as healthcare, productivity and ecology, but its benefits to the learning experience when applied to education remain unclear.

Previously [1], we presented an experiment in which an MSc course, Multimedia Content Production (MCP), was gamified, and where several metrics of student behavior were compared between a gamified and the previous non-gamified year. Students perceived the course as being more motivating and interesting than other non-gamified courses, and we found evidence that lecture attendance, attention to reference material and both participation and proactivity on the course's online forums greatly improved, although final grades were unaffected. Here we present a new study where a new gamified trial of the course was deployed to gather additional student data. We compare data from both gamified experiments with that from three previous regular years, to confirm our findings about student engagement with more significance, and also try to assess the impact of gamification over student grades. We fine-tuned the course experience based on student feedback, mostly related with the insufficient rewards for the game component, lack of student cooperation and lack of rewards for oral participation. We aimed to address the following questions:

Does more complete data support our previous findings? Do students perceive the gamified editions as more motivating and interesting? Are attendance, downloads and post numbers better than those of non-gamified ones?

How did the gamified experiment affect the grades? Are there any significant correlations between engagement and the students' final grade?

How was student engagement affected by the second gamified edition of the course?

Answering these questions provided deeper insights into how gamification can be applied to education, improving student engagement and learning outcomes.

RELATED WORK

Previous experiments show that games can improve one's learning outcomes, motivation and diligence. McClean et al. [18] show that students learning college level cell biology with a video game performed 30% better in their assessments than those learning without. Moreno [19] had similar findings in his experiment, where students used a video game to improve their programming skills and performed 12% better in their final exam. In another study where videogames were used for homework as part of a numerical methods course, results suggest more "intellectual intensity, intrinsic motivation, positive effect and overall student engagement" [4].

Gamification tries to take advantage of these beneficial effects to motivate users into embracing certain behaviors. For instance, Fitocracy[1] uses game elements like points, levels, badges and challenges, as well as a strong social component, to encourage users to work out and stay in shape. Other interesting examples include improving one's productivity [23] or encouraging eco-friendly driving [13]. Gamification has also been used to motivate people to learn and train new skills. Adobe Level Up[2] for instance, engages users into exploring features from Adobe Photoshop. Research works on this subject show promising results. For example, GamiCAD [15] is a gamified tutorial for AutoCAD to teach how to perform line and trimming operations. As users complete tasks, they help NASA build a spacecraft to participate in a mission. Results show that users completed tasks faster and found the experience to be more engaging and pleasant than the non-gamified version.

Gamification of education is still recent, but there are already a few well-known services that use game elements to teach. Khan Academy[3] is a free service that allows students to learn about several topics online, by watching videos and performing exercises. Progress is rewarded with energy points and badges. Codeacademy[4] teaches online students to code in several programming languages. It also uses points and badges to track progress. While systems like these have been used in flipped classrooms [24], empirical data to vouch for this kind of method are lacking.

In his book, Lee Sheldon [22] describes how a conventional learning experience can be designed as a game without resorting to technology, to engage students and make classes more fun and interesting. Students start with an F and go all the way up to an A+, by completing quests and challenges, and gaining experience points. However, little statistical evidence is provided to support any potential benefits of this kind of approach. Due to this lack of data, we made a long-term study assessing the potential benefits of gamification in college-level education.

[1] https://www.fitocracy.com/

[2] http://success.adobe.com/microsites/levelup/index.html

[3] http://www.khanacademy.org

[4] http://www.codecademy.com/

THE MCP EXPERIMENT

MCP is an annual semester-long MSc course in Information Systems and Computer Engineering at Instituto Superior Técnico. The course runs synchronized across two college campuses, Alameda and TagusPark, and follows a blended learning program, in which students attend live theoretical lectures (3 per week) and work on projects in laboratory classes (1 per week), but they also engage in discussions and complete assignments in the course's virtual learning environment (Moodle[5]). The theoretical lectures cover multimedia concepts like capture, editing and production techniques, multimedia standards, Copyright and Digital Rights Management. In lab classes, diverse concepts and tools are taught on image, audio and video manipulation, and there are regular assignments. We have five years' worth of student data (from 2007/2008 to 2011/2012), including attendance to lectures, posts and downloads on Moodle, and grades. Measures like these have been used before to informally assess student engagement [16].

Grading

Course evaluation varied slightly over the years, but the evaluation components were consistent. In the three non-gamified years, typical course evaluation consisted of regular quizzes (25% of total grade), lab evaluations (20%, 20% and 15%), online participation in the course's forums (10%, 10% and 5%), a thematic multimedia presentation (20%), a final exam (25%, 25% and 35%), and 5% extra for lecture attendance. The final grade ranged from 0 to 20, with a passing grade of 10. In the two gamified years, instead of grade points, students were awarded experience points (XP), by meeting traditional evaluation criteria. It included quizzes (20% and 10% respectively), a multimedia presentation (20%), lab classes (15%), a final exam (35%) and a set of collectible achievements (10% and 20%, plus a 5% extra). Overall, the evaluation methods were similar, with achievements replacing the online participation and attendance bonuses in gamified editions.

Gamified MCP 1.0

The main motivation to gamify the course was to improve student engagement and make it more interesting. We added a few game elements, such as XP, progress levels, a leaderboard, challenges, and badges, which seem to be consensually used in gamification [14, 5, 26, 25]. As students perform course activities, they are awarded with XP, which provides direct feedback on how they are doing and motivates them by instant gratification [20]. In the first gamified year (2010-2011), every 900 XP corresponded to a progress level, which corresponded to a grade on the 20 points' scale. For example, a student with 1800 XP would be at level 2, which means her grade was 2 so far. To prevent rounding problems, students were given a head start bonus of half a level (450 XP) for enrolling in the course.

[5] http://www.moodle.org

The leaderboard webpage (see Figure 1) provided an entry point to the gamified experience, and it was publically available from the forums. Here, enrolled students were displayed, sorted by descending order of level and XP. Each row portrays the player's rank, photo and name, campus, XP, level and achievements, awarded as collectible badges for completing course activities, like attending lectures, finding resources related to class subjects, finding bugs in class materials or completing challenges. By clicking a row, the achievement history for that player was shown, which made progression transparent and allowed students to learn by watching others.

Challenges were tasks students had to complete to be granted XP and achievements. There were two main Challenge types. Theoretical Challenges were activities presented to students over the semester, at the end of some lectures. These consisted of small creative and time-limited tasks, designed to explore multimedia types and materials taught in those lectures. A special case of Theoretical Challenges were the Online Quests, which were more demanding, not particularly bound to any lecture, and with a longer deadline. Online Quests were the only challenges that already existed in the previous non-gamified course versions. The second type was the Lab Challenges, which were assigned during the first month of classes, before other lab evaluations started. These were meant to be fun and expressive, by allowing students to make creative content using multimedia tools introduced in lab classes.

Challenges were formally issued via posts to course forums by faculty. Achievements and graded activities were logged based on students' posts, except for attendance, lab grading, quizzes and exams, which were manually recorded. The main purpose of challenges and achievements was to model course activities into meaningful deeds while providing the autonomy to choose what tasks and achievements to pursue. Achievements also provided feedback on how proficient students became on specific subjects. Achievements could be single-level, (a task that should be performed once) or multi-level (several times with increasing difficulty). The whole scoring process was done manually. Data from lectures and lab classes were logged by faculty, and activity logs from Moodle were also daily downloaded. Two to three times a day, a script was manually run to process all log files and generate the updated leaderboard webpage.

There seems to be a strong bond between how intrinsically motivated students are and how well they perform in learning [21]. Self-Determination Theory (SDT) identifies three innate needs of intrinsic motivation [7]: Autonomy, a sense of volition or willingness when doing a task; Competence, referring to a need for challenge and feelings of effectance; and Relatedness, experienced when a person feels connected with others. We tried to align the gamified course's goals with those of students, in order to improve the experience's intrinsic value [8]. The rationale behind the selection and integration of game elements with the

Figure 1. The MCP course leaderboard.

course was thus based on SDT: Autonomy was promoted by providing options on what challenges to pursue and which achievements to level up. We tried to boost competence by displaying positive feedback and progress via points, levels and badges. We aimed to improve relatedness by providing means of competition, cooperation and online interaction among players.

Gamified MCP 2.0

In the academic year of 2011-2012 we performed another experiment. Our gamified course was improved based on student feedback from the previous year, with two new achievements to reward cooperation in the labs and one to reward oral participation. Additionally, a new achievement rewards students for timely responses to challenges and another rewards students for compiling challenge results. We also had critiques about the achievements being too much trouble for only 10% of total grade, which made us re-grade the course so that quizzes would be worth 10% less and achievements 10% more. Since achievements were now worth more, we created six new challenges, four theoretical and two lab challenges, with the intent of making the work load more even over the semester, which was also criticized before. For cosmetic purposes, we changed the amount of XP per level from 900 to 1200 XP.

THE SECOND TRIAL

In the first gamified experiment, we saw significant improvements in lecture attendance and in the number of downloaded lecture slides, with the most gains in both initiated threads and reply posts on the forums. This suggested that students were more participative and proactive, enjoyed attending classes more and paid more attention to support materials, which seems to reflect a deeper engagement with the course. However, no significant gains were seen in final grades, which questions if this approach affects learning outcomes or not. With our second experiment we tried to verify our previous findings and assess the impact of our approach over student grades.

We analyzed data from five academic years, beginning in 2007-2008 and ending with 2011-2012, with the last two being gamified years. We collected data regarding student lecture attendance, number of downloads of lecture slides, number of posts, as well as grades from quizzes, lab evaluations, the multimedia presentation, the exam, and the final grade. Students had similar backgrounds in all years, with the majority of them having finished their computer science undergraduate degree on the previous year, and a minority composed by one to three exchange students. For the five years, we had 52, 62, 41, 35 and 52 students respectively, excluding those that dropped out or only enrolled mid-semester and could not complete the course. The faculty staff for the theoretical lectures was composed of two professors and it remained the same across all years. As for lab classes, it varied between one and two instructors that changed from one year to the next. We had 18 lectures in every year except for 2009-2010, in which we had 19.

Student data in both experiments was collected in an uncontrolled environment. Variables like the composition of taught subjects, support materials and faculty staff could not be manipulated, as some of these changed and evolved on a five-year time span. Since our data did not appear to follow a normal distribution, all statistical differences between groups were checked using the non-parametric Kruskal-Wallis and post-hoc Mann-Whitney's U tests with Bonferroni correction. Correlations between variables were calculated using the Spearman's rank correlation coefficient. We have also collected qualitative feedback from students with a questionnaire by the end of each gamified experiment, which we will also present.

Lecture Attendance

We saw ***significant differences in mean attendance by student*** over the years (Kruskal-Wallis test, $H(2) = 20.14$, $p < 0.001$), with these being between 2007-2008 and 2010-2011 (Mann-Whitney's U, $p < 0.05$), respecting to an increase by 12% (see Figure 2), and between 2010-2011 and 2011-2012 ($p < 0.05$), reflecting a decrease of 11%. The lack of significance across years, and mainly among the gamified ones, seems to contradict our previous finding that student attendance would increase significantly with our approach. We saw a ***moderate correlation between lecture attendance and the amount of student posts*** (Spearman's coefficient, $\rho = 0.48$, $p < 0.001$) in 2011-2012. In 2010-2011 this correlation was weak and less significant ($\rho = 0.29$, $p < 0.1$) and in the non-gamified years there was no correlation ($\rho <= 0.1$, $p > 0.1$). We saw a similar pattern comparing lecture attendance with student posts on challenge threads, with a moderate correlation ($\rho = 0.47$, $p < 0.001$) not previously seen. This suggests that a positive correlation between lecture attendance and posts might be an emergent trend, but this matter requires further study.

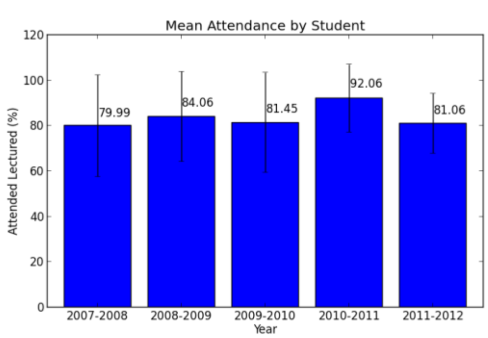

Figure 2. Mean attendance by student.

Support Material Downloads

The number of downloads of lecture slides, normalized to the number of students, presented significant differences (Kruskal-Wallis, $H(2) = 20.14$, $p < 0.001$) between every gamified and non-gamified year (Mann-Whitney's U, $p < 0.05$). Both gamified years presented an average of 4.17 normalized downloads per lecture, which is almost twice as much as the 2.32 seen in 2009-2010, and tree times more than the 1.2 and 1.41 seen in 2008-2009 and 2007-2008 (see Figure 3). Considering all support materials in the course, the numbers were less enlightening (see Figure 4). We found ***significant differences in mean downloads per student*** across years (Kruskal-Wallis, $H(2) = 58.20$, $p < 0.0001$), being these between 2010-2011 and every other year, and between 2007-2008 and all other years, except for 2010-2011 (Mann-Whitney's U, $p < 0.05$). We clearly have two cases where the number of mean downloads per student were above normal, but since these correspond to a non-gamified and a gamified edition, with three years apart, it is hard to evaluate the impact in this measure. Differences in course materials over the years might be behind this effect.

Posts on Forums

We found ***significant differences in mean posts by student*** (Kruskal-Wallis, $H(2) = 141.20$, $p < 0.001$), between every non-gamified and every gamified year, and between

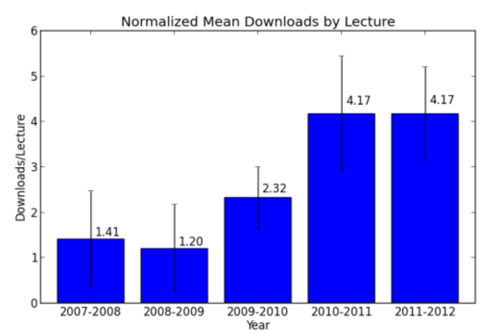

Figure 3. Normalized mean downloads by lecture.

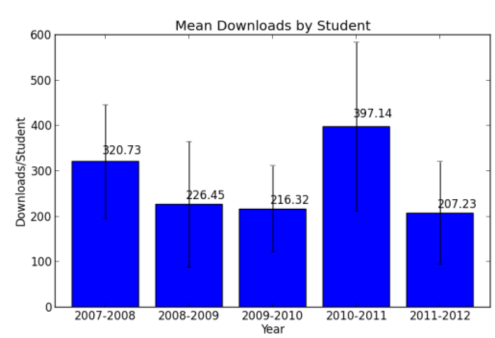

Figure 4. Mean downloads by student

gamified years (Mann-Whitney's U, $p < 0.05$). The number of posts per student increased significantly 4 to 6 times in the first gamified experiment, but in the second gamified year, the number increased again by 66% when compared to the previous one (see Figure 5), which suggests a significant growth in student participation. This effect can be explained by the ***significant increase in the number of reply posts by student*** (Kruskal-Wallis, $H(2) = 53.18$, $p < 0.001$) between the same years (Mann-Whitney's U, $p < 0.05$), and indicates that most of student participation consists of replies to other posts. This was expected as students have more opportunities to reply to others (e.g. challenges) than to create new threads.

There were also ***differences in terms of initiated threads by students*** (Kruskal-Wallis, $H(2) = 53.18$, $p < 0.001$) between every non-gamified and gamified year (Mann-Whitney's U, $p < 0.05$). These differences denote a significant increase of student proactivity in the gamified years, which were 3 to 4 times bigger when compared to the non-gamified year with the most first posts. Even though there was a small drop from the first to the second gamified year, this effect was not statistically significant but is likely to be related to the fact that here were more threads to reply to.

Challenges responded to 52% of student posts in the first gamified year and 73% in the second, which suggests that challenges are playing and increasingly strong role as participation engagers. In fact, there was a ***significant increase of 133% (13.0 vs. 30.3) in challenge posts by student*** (Mann-Whitney's U, $p < 0.001$), which further supports this finding. This effect can be in part explained by challenges being more attractive in the second year, as there were more of them and they were worth more. Hence, it was no surprise to find a ***very strong correlation between the number of student posts and student posts on challenge threads*** in both gamified years ($\rho > 0.85$, $p < 0.001$). We saw a ***significant increment of 55% (1.08 vs. 1.68) in the amount of posts per challenge thread normalized to the number of students***, between the two gamified years, although this result has limited significance (Mann-Whitney's U, $p < 0.1$). This indicates that student posts increased, not only because there were more challenges, but because they posted more per challenge.

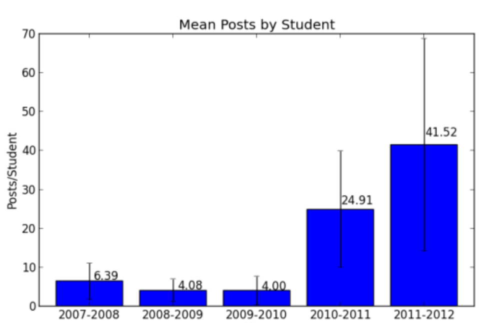

Figure 5. Mean posts by student.

We found an interesting ***moderate correlation between the amount of challenge posts and the amount of posts on forums that would not reward students with XP***, both in the first (Spearman's coefficient, $\rho = 0.40$, $p < 0.05$) and the second gamified years ($\rho = 0.53$, $p < 0.001$). These results imply that students that participate more on challenges might as well participate more on non-rewarded activities, which points to increased engagement. Furthermore, the mean number of non-rewarded posts by student over the years is 3.64, 1.67, 1.85, 5.31 and 4.34, which suggests that ***non-rewarded participation has also been increasing***. However, these differences were only significant between 2008-2009 and 2011-2012 (Mann-Whitney' U, $p < 0.05$).

Grades

Final grades were significantly different (Kruskal-Wallis, $H(2) = 24.55$, $p < 0.001$) between 2007-2008 and 2011-2012, 2008-2009 and 2010-2011, and 2010-2011 and 2011-2012 (Mann-Whitney's U tests, $p < 0.05$). While the first gamified year presented a grade drop, only significant relative to one non-gamified year, the second gamified year presented a significant improvement relative to both the previous year and one non-gamified year (see Figure 6). While it is not consistently significant, the second gamified year presents the highest mean final grade by student ever seen. The minimum grades per year, which were 12, 12, 13, 12 and 14, also increased in the second year, suggesting that students that would typically perform worst could now perform better. Moreover, we had the maximum relative

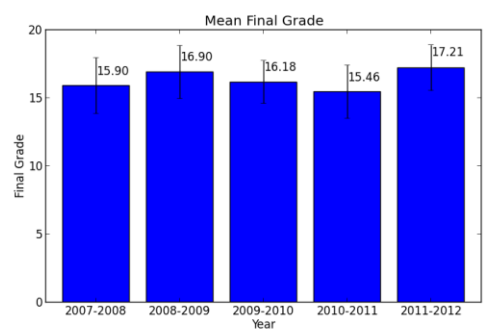

Figure 6. Mean student final grade.

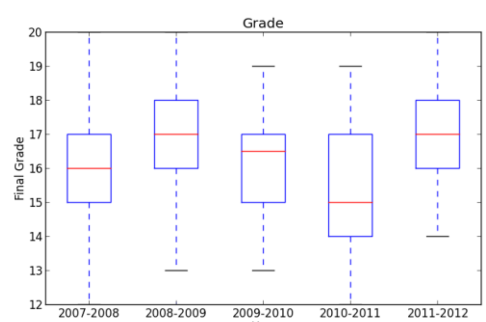

Figure 7. Student final grade boxplot, showing the first and third quartiles (box), and the lowest datum still within 1.5 times the interquartile range (whiskers).

amount of students reaching the top grade ever seen, as depicted by the yearly progression of 2%, 6%, 0%, 0% and 11.5%. This suggests that we were able to homogenize learning outcomes by reducing grade variance and minimizing the gap between the lowest and the top graders, as see in Figure 7. While trying to identify which grading components were responsible for the final grade growth between gamified years, we saw similar improvements over quiz, lab, multimedia presentation and final exam grades, but with no statistically significant differences.

There was a ***strong correlation between the number of posts per students and their final grade*** in 2009-2010 (Spearman's coefficient, $\rho = 0.71$, $p < 0.001$), 2010-2011 ($\rho = 0.68$, $p < 0.001$) and 2011-2012 ($\rho = 0.67$, $p < 0.001$), while it was moderate in 2007-2008 ($\rho = 0.53$, $p < 0.001$) and weak in 2008-2009 ($\rho = 0.35$, $p < 0.001$). Although the effect on the two gamified years can easily be explained by posts becoming a substantial part of the evaluation, the same effect in 2009-2010 is hard to explain, as it greatly differs from the rest of the non-gamified years. We also found a ***growing strong correlation between the amount of challenge posts and a student's final grade*** in 2010-2011 ($\rho = 0.58$, $p < 0.001$) and 2011-2012 ($\rho = 0.60$, $p < 0.001$). In particular, ***theoretical challenge posts seem to be strongly correlated with the final grade*** in the first ($\rho = 0.61$, $p < 0.001$) and second gamified years ($\rho = 0.58$, $p < 0.001$), and also moderately correlated with quiz grades in both 2010-2011 ($\rho = 0.50$, $p < 0.01$) and 2011-2012 ($\rho = 0.45$, $p < 0.001$). This trend suggests that students that participate more in challenges also have better grades, and those that undertake theoretical challenges more might also score better on quizzes, which might be related to both having a continuous nature over the term. Yet, we could not find any correlation between theoretical challenge posts and the exam grade, or between lab challenge posts and the lab evaluations grades, contrarily to what we would expect.

Student Feedback

By the end of both gamified semesters, we carried out a questionnaire to gather qualitative student feedback about the new learning experience, where students had to rate statements using a five-point Likert scale. We got 28 answers in the first gamified year and 46 in the second. Taking a look at the answers' mode, students considered that the gamification experiment applied to the MCP performed very well (4 in both years) [1 – terrible; 5 – excellent]. When compared to other courses, they considered the MCP course to be more motivating (5 in the first year, 4 in the second) and interesting (4) [1-much less; 5 - much more]. They also considered that the course required more work (4, 5) but was neither more difficult (3) nor harder to learn from (3) [1-much less; 5 - much more]. They considered the study to have the same quality (3) of other courses, but also to be more continuous (4) [1 - far less; 5 - far more]. Students mildly felt that they were playing a game instead of just attending a regular course (3) [1 - not at all; 5 - a lot] and while in the first year they had not a clear opinion on whether achievements should account for more of the course grade or not (3), most of them though it should on the second year (4) [1-definitely not; 5 - definitely yes]. Also, students deemed achievements that required extra work, like the challenges and posting related material on the forums, to have contributed to their learning (4) [1-not at all; 5 - definitely] and agreed that gamification should be extended to other courses (5, 4) [1-definitely not; 5 - definitely yes].

The New Achievements

The new achievements had limited success. Those targeted at promoting cooperation were underused. Groups with good performance often blamed those with lower for the XP that had not been awarded. As compiling challenge results was too much trouble for only 100 XP, only one student undertook this task. The achievement for timely responding to challenges was highly criticized for promoting fast responses over meaningful posts. The oral participation badges were earned by 23 students and we had a few critiques concerning students feeling pressured to talk in class and resenting others being rewarded for doing it.

DISCUSSION

With our thorough analysis completed, we can now return to the research questions described in the introduction.

Does our new data support our previous findings?

Contrarily to what we have previously found, we could not observe a consistent increase on student attendance in both gamified years. We saw a significant increase in the first year but the same did not occur in the second, which presented attendance levels similar to non-gamified years. Our approach seems to have no effect on attendance, which might be explained by most gamification occurring over online content. This subject requires further study.

The number of normalized downloads per lecture grew consistently on both gamified years, presenting an increase of 1.5 to 3 times the number of downloads observed on those non-gamified. This suggests that students might have been more motivated to download lecture slides, as seen in our previous study. However, the number of downloads per student had a significant increase on the first gamified year, but in the second it went back to values similar to the non-gamified years. This might be due to the variation of course materials, which could have affected student interest and the amount of downloadable items. It is hard to draw conclusions here but, ultimately, it suggests that students can be engaged to pay attention to course material as long as it is rewarded, like many other aspects of the experience.

Compared to non-gamified years, the number of posts per student grew significantly 4 to 6 times on the first gamified year, and 6 to 10 times in the second. This derives from significant increases in both reply posts and initiated threads. Results support our previous findings, suggesting that the gamified course can engage students into participating and being more proactive in forums.

Student feedback was consistent across experiments, with them finding the gamified course to be more motivating and interesting than other regular courses, and that the gamified experience should be extended to other courses. MCP was perceived as being as easy to learn from as it is from other courses, and to have a study with the same quality but more continuous. Students mildly felt they were playing a game, which suggests that the game-like feeling has still room for improvement. They also felt achievements contributed for their learning but should be better rewarded, especially in the second year, which is in pair with their increased perception that the course required more work.

How did the gamified experiment affect the grades?

Although our second gamified year presented the highest final grade to date, we lack statistical evidence to support a significant increase. However, we observed both the highest minimum grade and the most students reaching the top grade ever, which suggests an improvement of learning outcomes. This effect implies a decreased disparity of grade distribution, which seems to have benefited all students, including those that would score poorly. We hypothesize that this happened because there were more and better rewarded challenges, and these gave students more opportunities to succeed. The significant increase in terms of both challenge posts per student and posts per challenge thread, between gamified years, and the growing strong correlation between challenge posts and final grade, seem to support this hypothesis. It may also explain, in part, the differences in the learning outcomes between experiments.

We found that in both gamified years there was a strong correlation between theoretical challenges and final grades, and also a moderate correlation with the quiz grades. Given that both quizzes and theoretical challenges are forms of continuous assessment, that both cover broadly the same topics, and that quiz grades slightly improved from the first to the second experiments, we hypothesize that theoretical challenges might help students to study and get better grades on continuous assessment components, like the quizzes. In the first gamified year, due to the small number of challenges and their uneven distribution, students felt that the second half of the term had fewer activities to do. This seems to support the benefits of using challenges to boost both student engagement and learning outcomes.

How was engagement affected by Gamified MCP 2.0?

Both experiences suggest that students were more proactive and participative in our gamified course, and that attention to reference materials might be positively influenced, which suggests a deeper engagement. This is corroborated by the students' opinion that the course was more motivating and interesting than other non-gamified courses. The second experiment, however, brought additional evidence that gamification can indeed enhance student engagement.

In the second year, we increased the amount of challenges, improved their distribution over time, and increased their reward. As a result, we saw a significant increase of 66% on student posts over the previous year, with 73% of these being made in challenge threads, which reflects a huge improvement in student participation. One might argue that students posted more because they had more challenges to attend, but data shows they posted 55% more per challenge thread. The question of whether they have posted more because challenges were more rewarding might arise, but challenges were worth less than 5% of the maximum grade. Also, we found a moderate correlation between the number of challenge posts and non-rewarded posts, and this correlation grew from the first to the second experiment. This might suggest that the more students participate on challenges, the more they participate on other non-rewarding activities, and that we might improve overall participation and engagement by creating better challenges. In the first experiment we hypothesized that an uneven distribution of challenges over the semester might have rendered the course less engaging, due to the existence of long periods without interesting activities to perform. Albeit we do not have strong evidence, our data leads us to believe that by evening out challenge distribution we can not only improve student engagement but also improve other forms of continuous evaluation and, therefore, improve their final grade. This is an interesting topic of future research.

Study limitations

Our study has four major limitations, most related to variables we could not control. First, the student population was different from one year to another, which might have influenced the results. A between-subjects experiment on the same year, comparing non-gamified and gamified approaches, would solve this problem. However, the two entail two very distinct workloads and grading systems, and it would not be ethical to grade students taking the same course in two different ways. Secondly, topics covered by the course changed over the years, which might have affected the number and nature of support materials. In order to keep the course updated, as enforced by the school, we had to accept this limitation. Third, although the faculty staff remained the same for the theoretical lectures, for lab classes it changed from one year to the next. Given that lab instructors only dealt with students directly once a week, were not liable, and had limited intervention in the course, we argue that the inherent effect might be neglectable. The forth limitation refers to the usage of informal measures of engagement, which lack validation. Given the significance penalty these limitations entail, our results must be taken with caution. In the future we would like to study the possibility of reproducing the experiments in a controlled environment and perform formal engagement validation.

CONCLUSION

In this paper we presented a gamified course and discussed how gamification can be used to improve student engagement. Although lecture attendance seems to be unaffected, results showed that with our gamified learning experience students participated more and were more proactive in the forums, and also paid more attention to the lectures' slides, which suggests a deeper engagement. This is in pair with student feedback, with them finding the course to be more motivating and interesting than other regular courses. Our study also suggests that evening out challenge distribution over the term and making them fairly rewarded might significantly improve student participation and performance. Students seem to score better with the gamified version of the course and grade differences between them seem to decrease. For future work we would like to further study the impact of our approach over student outcomes and perform a formal engagement evaluation.

ACKNOWLEDGMENTS

This work was supported by FCT (INESC-ID multiannual funding) under project PEst-OE/EEI/LA0021/2013 and the project PAELife, reference AAL/0014/2009. Gabriel Barata was supported by FCT, grant SFRH/BD/72735/2010.

REFERENCES

1. G. Barata, S. Gama, J. Jorge, and D. Gonçalves. Engaging engeneering students with gamification. In *Proc. VSGAMES 2013*, 2013.
2. Y. Beldarrain. Distance education trends: Integrating new technologies to foster student interaction and collaboration. *Distance Education*, 27(2):139–153, 2006.
3. J. Chen. Flow in games (and everything else). *Commun. ACM*, 50:31–34, 2007.
4. B. Coller and D. Shernoff. Video game-based education in mechanical engineering: A look at student engagement. *International Journal of Engineering Education*, 25(2):308–317, 2009.
5. C. Crumlish and E. Malone. *Designing social interfaces*. O'Reilly, 2009.
6. M. Csikszentmihalyi. *Flow: The psychology of optimal experience*. Harper Perennial, 1991.
7. E. Deci and R. Ryan. *Handbook of self-determination research.* University of Rochester Press, 2004.
8. S. Deterding. Gamification: designing for motivation. *interactions*, 19(4):14–17, 2012.
9. S. Deterding, D. Dixon, R. Khaled, and L. Nacke. From game design elements to gamefulness: defining "gamification". In *Proc. MindTrek '11*, volume Tampere, F, 9–15. ACM, 2011.
10. S. Deterding, M. Sicart, L. Nacke, K. O'Hara, and D. Dixon. Gamification. using game-design elements in non-gaming contexts. In *Proc. CHI EA '11*, 2425–2428, ACM, 2011.
11. D. R. Garrison and H. Kanuka. Blended learning: Uncovering its transformative potential in higher education. *The internet and higher education*, 7(2):95–105, 2004.
12. J. P. Gee. What video games have to teach us about learning and literacy. *Comput. Entertain.*, 1(1):20–20, Oct. 2003.
13. O. Inbar, N. Tractinsky, O. Tsimhoni, and T. Seder. Driving the scoreboard: Motivating eco-driving through in-car gaming. In *Proc. CHI '11 Workshop Gamification: Using Game Design Elements in Non-Game Contexts.* ACM, 2011.
14. A. J. Kim. Putting the fun in functional. http://www.slideshare.net/amyjokim/putting-the-fun-in-functiona, March 2008.
15. W. Li, T. Grossman, and G. Fitzmaurice. Gamicad: a gamified tutorial system for first time autocad users. In *Proc. UIST '12*, 103–112, ACM, 2012.
16. B. J. Mandernach, E. Donnelli-Sallee, and A. Dailey-Hebert. Assessing course student engagement. *Promoting Student Engagement*, 1, 2011.
17. F. G. Martin. Will massive open online courses change how we teach? *Commun. ACM*, 55(8):26–28, 2012.
18. P. McClean, B. Saini-eidukat, D. Schwert, B. Slator, and A. White. Virtual worlds in large enrollment science classes significantly improve authentic learning. In *Proc. of the 12th International Conference on College Teaching and Learning*, 111–118, 2001.
19. J. Moreno. Digital competition game to improve programming skills. *Educational Technology & Society*, 15(3):288–297, 2012.
20. L. Natvig, S. Line, and A. Djupdal. "age of computers"; an innovative combination of history and computer game elements for teaching computer fundamentals. In *Proc. FIE 2004*, S2F – 1–6 vol. 3, 2004.
21. R. M. Ryan and E. L. Deci. *Handbook on motivation at school*, chapter promoting self-determined school engagement, 171–196. Routledge, 2009.
22. L. Sheldon. *The Multiplayer Classroom: Designing Coursework as a Game*. Course Technology PTR, 2011.
23. S. Sheth, J. Bell, and G. Kaiser. Halo (highly addictive, socially optimized) software engineering. In *Proc. of the 1st international workshop on Games and software engineering*, 29–32, ACM, 2011.
24. C. Thompson. How khan academy is changing the rules of education. *Wired Magazine*, 1–5, 2011.
25. K. Werbach and D. Hunter. *For the Win: How Game Thinking Can Revolutionize Your Business*. Wharton Digital Press, 2012.
26. G. Zichermann and C. Cunningham. *Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps*. O'Reilly Media, Inc., 2011.