Integrating Aspects of Gamification in the Classroom: Takeaways from a Tentative Experience with Undergraduates

Taciana Pontual Falcão

Departamento de Estatística e Informática (DEINFO)
Universidade Federal Rural de Pernambuco (UFRPE)
Rua Dom Manoel de Medeiros, s/n, Dois Irmãos - 52171-900 - Recife/PE - Brasil
taciana.pontual@ufrpe.br

Abstract. The growing interest for gamification in education is undeniable. However, the pedagogy behind it is overshadowed by promises of seducing students through the appeal of games. Despite the rather optimistic accounts in the literature, integrating gamification into learning processes is in fact much more complex than it may seem. This paper reflects on a tentative experience at undergraduate level, with aspects of gamification integrated to the face-to-face classroom using a methodology that tried to escape from traditional format of lectures and exams.

1. Introduction

The learning process at the university has undergone radical changes with the popularisation of digital technologies and the world wide web. Access to information, previously restricted to libraries and to the words of the teacher, is now ubiquitous. Students have access to any subject, at any time, through the smartphones in their pockets, and consequently the motivation to attend classes has dropped considerably. Contributing to this lack of interest is the fact that teaching methods have not kept up with the changes in society and learning styles, remaining heavily-based on exposition of content through lectures [Weller and Gould 2015].

A number of initiatives emerge as efforts to reshape teaching and increase students' motivation and interest. Gamification is a technique that is gaining a lot of attention in education for integrating aspects of games into learning processes [Kapp 2012]. Digital games are among the favourite hobbies of children and young people nowadays, triggering an intrinsic motivation sought by educators. The idea is that bringing learning contexts closer to the gaming world will make learning fun and engaging for the current generation of students, and have a positive impact on their achievement and knowledge construction. A lot has been reported about isolated initiatives using gamification at different levels of depth [Borges et al. 2013] [Dicheva et al. 2015] [Klock et al. 2015], but the discussion about the actual learning benefits of the technique is still little (e.g. [Nogueira Neto et al. 2015]). Furthermore, most authors discuss gamification implemented in virtual environments [Brazil and Baruque 2015] and few in face-to-face classroom activities [Laster 2010]. The ‘hype’ over gamification urges for deeper investigations about its pedagogical basis, how and in which contexts it should be implemented, and the pros and cons of the technique. This paper aims to contribute to this discussion by presenting students' opinions and teacher's reflections on a face-to-face course with integrated elements of gamification.
2. Context and Methodology

The initiative described in this paper took place in the course of Technologies for Learning, in a Computer Science department of a Brazilian public university. The 32 students who initially enrolled to the course were taking three different degrees: Bachelor in Computer Science (20), Bachelor in Information Systems (6), and Educational degree in Computer Science\(^1\) (6). As the course is optional and has no pre-requisites for attendance, students were from varied university years. There were 9 females and 23 males, mainly around 20 years-old. Classes took place twice a week, in the evening, and lasted for 1 hour and 40 minutes. The course was planned and mediated by the author of this paper.

The idea to 'gamify' the course on Technologies for Learning was driven by a permanent personal pursuit to innovate in teaching - with a special focus on students' motivation - and by the belief that lectures are each year less effective as a teaching method. At the same time, the course itself aims to discuss how technologies can improve the learning process, making it almost 'compulsory', from a pedagogical point of view, to experiment new methods, showing students in practice some of the techniques discussed during the course. As the teacher, I performed a small piece of interpretivist action research, with a qualitative, flexible design, where I analysed my own work and my students' feedback. According to interpretivism, apprehension of the world goes through selection and interpretation, linked to people’s values, context and cultural background [Rubin and Rubin 2005]. More specifically, this means that researchers’ previous knowledge, even if subconsciously, affect and inform the research, and research findings represent a combination of the understanding of the researcher and of those being researched. The present work is a descriptive and exploratory kind of research that aims to investigate the ‘how’ and ‘why’ of phenomena (here, gamification in the classroom). Throughout the course, data was constructed to enable me to reflect on the experience in the light of learning theories and innovative teaching methods.

2.1. Starting Off: Planning and Expectations

The methodology was presented to the students at the beginning of the course. They were told the course would follow a constructivist approach (statement which I myself later questioned, see section 4.1), with a lot of practical activities and very few (or maybe none) lectures. They were expected to work in groups to design and prototype a system (mobile application, digital game, web site, or other) to support the learning of a content of their choice, aligned with a learning theory (of their choice). They were given flexibility in their choices, as long as they were able to convince me with reasonable and theory-grounded arguments.

Gamification was initially presented as follows: the course would be divided in phases, unknown to all until each group unblocked them by completing the previous phase. So, groups could progress independently and compete for the lead. Both groups and individuals could collect points throughout the course according to some criteria (Table 1). These points would be then transformed into students' grades (somehow). The decision of giving points separately to groups and individuals was due to difficulties, in

\(^1\) In Brazil, this is a degree for those aiming at teaching Computer Science in schools
previous classes, in evaluating work of individuals within groups. Repeatedly, problems arise involving students who do no work at all and those who feel unfairly evaluated. Individual points, thus, were used as an opportunity for good students to stand out, 'despite' their group, if this was the case.

Table 1. Initial criteria for distributing points

<table>
<thead>
<tr>
<th>Individual criteria</th>
<th>Group criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creativity</td>
<td>Creativity</td>
</tr>
<tr>
<td>Interesting question</td>
<td>Adherence to deadlines</td>
</tr>
<tr>
<td>Initiative</td>
<td>Quality of deliverables</td>
</tr>
<tr>
<td>Participation</td>
<td></td>
</tr>
<tr>
<td>Peer collaboration</td>
<td></td>
</tr>
</tbody>
</table>

Students were summoned to build the course with me, giving suggestions, being autonomous, surprising me, going beyond expectations. However, they were also aware that this was my first attempt at gamification, and therefore not everything was defined at the outset, decisions would be made on-the-go, and changes were to be expected. Initially, the course was loosely planned into phases that fit the user-centred design process, i.e. research, ideation, prototyping and evaluation. However, more fine-grained phases emerged during the course, as presented in the next section.

2.2. From Theory to Practice: Definitions Along the Way

2.2.1. Phases

Within the user-centred design overarching process, phases were defined according to the teacher's insights as the groups progressed (Figure 1). Formal techniques were introduced in the process incrementally, so that students had theoretical constructs they could adopt.

![Figure 1. Phases of the course](image)

Some phases had deliverables to help students organise their ideas and make them more concrete. Besides the final presentation, students were asked to present their on-
going work at checkpoints, as moments of sharing with all groups. As part of the gamification techniques, groups only knew of the following phase once they had completed the previous one.

2.2.2. Visualisation of Progress

A 'ranking-like' method was needed for publicising students' progress through the phases, feedback on deliverables and activities, and points earned. Rankings are common elements of games that engage through competition and therefore are expected to motivate students to pursue better achievement. With this aim, a spreadsheet was created online and shared with all students (with read-only permission). The spreadsheet had the following information:

- Progress sheet: the progress of all groups through the phases;
- Group points sheet: points earned by each group, per criterion;
- Individual points sheet: points earned by each student, per criterion;
- Feedback sheet: an evaluative feedback per group, per activity and deliverable.

The spreadsheet was updated by myself after each class (twice a week), and students were encouraged to access it as often as possible.

2.2.3. Additional Assessment Criteria

As the course progressed, I perceived that other assessment criteria could be included for a more comprehensive student evaluation, besides those in Table 1. Table 2 shows the additional criteria that emerged from my observation in class.

<table>
<thead>
<tr>
<th>Table 2. Additional criteria for distributing points</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Individual points</strong></td>
</tr>
<tr>
<td>Motivation</td>
</tr>
<tr>
<td>Suggestions given</td>
</tr>
<tr>
<td>Content knowledge</td>
</tr>
<tr>
<td>Attendance to class</td>
</tr>
</tbody>
</table>

2.2.4. Grading

The university system where this work is contextualised demands that the teacher formally evaluates students through grades at mid-term and end of term. For this reason, the first grades were calculated after Phase 5, using all group and individual criteria, in the following manner:

1. Weighed averages were calculated for group points of adherence to deadline and quality of deliverables, for each phase.

2. All other group points were summed per criterion, and a normalised grade was given, i.e. the greatest number of points for each criterion became the maximum grade (the group obtained grade 10) and the other groups were graded proportionally.

3. A final weighed average was calculated considering the three averages described above, to define the group average grade.
4. An individual average grade was calculated following the same procedure as (2), considering the individual criteria.

5. Finally, the grade of the student was the arithmetic average between the group average grade and the individual average grade.

For the grades at the end of the term, a different process was adopted:

- A weighed average was calculated for group points of adherence to deadline, for each phase.

- Quality of deliverables was an arithmetic average between the teacher's evaluation and peer evaluation. Peer evaluation was performed through an online form prepared by the teacher and accessed by the students during group presentations. The whole class had to evaluate the group's prototype (quality, interactivity and visual interface); user evaluation (quantity of participants, method and results); and the presentation itself (collaboration within the group, clarity and ability to hold the class attention). All answers were collected through a 1-5 Likert scale. The teacher used the same form for evaluation.

- Other group points were calculated in the same way described in item (2) above, and the final group average was calculated as described in (3).

- The individual average consisted of a weighed average between a guided self-assessment and the participation in the evaluation of groups' presentation. The guided self-assessment was performed through an online form distributed by the teacher (results in section 3.1).

- Finally, the grade of the student was calculated as in (5).

The modification in the assessment process is discussed in section 4.

3. The Journey

Six groups were formed according to students' preferences. All groups successfully completed the course, producing prototypes in varied levels of quality and fidelity. Two male students from the Bachelor in Computer Science degree quit: one moved cities and the other decided to take a different degree. Three other male students, one from each degree, needed extra activities in the end (according to the university regulations) to reach the passing grade (7/10), due to low individual achievement during the course. Figure 2 shows that one third of the class (10 students) was below the passing grade at mid-term, while the end of term grades were generally higher. This mainly reflects the two procedures used for assessment and grade calculation, discussed in section 4.
3.1. Students' Voices

Students were encouraged to give me feedback at any time, but during the course this only happened with two students, one male and one female, from the Educational degree in Computer Science, who felt their mid-term grade (below 7) was unfair. The female student came to me in person, while the male student made a complaint via e-mail. I answered both in the same way, asking them to justify with concrete arguments (related to the assessment criteria and the work they had done) why their grade was not fair. Both had their grades adjusted, although I did not accept all arguments presented. The male student, in particular, argued that a lot of his work had been done extra-class (for example in organising group work using repositories and management software tools), and was not considered in the evaluation. At the end of the course, an online form was shared with all students to evaluate the course. The students did not have to identify themselves (but could do), and were encouraged to give their sincere opinions to help me improve the methodology. The form contained: Likert scale (1-5) questions; yes/no questions; other closed questions with three options; and open fields for justifications, suggestions and opinions. Results are presented next.

3.1.1. Information Spreadsheet

Regarding the form of visualisation of progress and feedback, 13 students declared they accessed the information spreadsheet a few times only. The same number of students accessed it several times, while only 2 students declared having accessed it many times. Only one student said the spreadsheet did not contain all necessary information, but this answer (as justified by the student) was based on the fact that all grades at the end of the term were only made available after the last group made their presentation. Two students used the open field to praise the organisation and level of detail of the spreadsheet. One of them declared: "the individual grades for each aspect demonstrates care and concern with the development of each student". Only 3 students declared they felt embarrassed with the fact that all peers could see their grades in the spreadsheet, which was a personal concern of mine.

3.1.2. Assessment and Grading

Twelve students thought their assessment (and consequent grades) was very fair, and the same number thought it was fair. Four students opted for the middle Likert scale (1-5) regarding the fairness of their assessment. The same male student who formally complained via e-mail reinstated that the assessment did not take into account extra-class work. This student gave a couple of suggestions to adjust this methodological issue: use online Google drive tools to evaluate group work, including using the 'History' tool to
identify which students contributed and how much; interview group members about their peers work (this was also suggested by another student, who said that "maybe informal questions to the group members would help to know what is really going on"). Another problem pointed out was the fact that a group had a low grade for adherence to deadline because some members of the group did not finish their work in time. A student also asked for more detailed individual assessment.

The assessment criteria were considered very adequate by 12 students, while 15 considered them adequate, and one student opted for the middle Likert scale regarding the adequacy of criteria. Students argued that several criteria were subjective, and it was hard for them to know how the teacher reached the evaluation and why (e.g. students disagreed that their work was not creative). Criteria related to participation were criticised for penalising shy students, who are not keen on taking part in debates or asking questions in presentations. One student declared that the assessment helped him/her to improve their performance and most of them would not exclude any of the criteria. Very few suggestions were given for including criteria; they were: development (related to efforts made to improve as the course progressed); and more criteria related to initiative and motivation.

3.1.3. Learning

Fourteen students said they learned a lot in the course, and 11 declared they learned reasonably well. Three students classified their learning as average. Eighteen students said they learn more doing practical activities, 3 said they learn more attending lectures, and 8 said their learning is similar for both methods. Four students declared they missed having lectures in the course.

3.1.4. Motivation

Twenty-four students classified the methodology as good, and said they were keen to take another course using gamification. Two students found it interesting, but in need of improvements, and two others said they did not adapt to the methodology. Nine students classified their motivation in the course as very high, 17 as high, and 2 as average. These two students justified their choice by tiredness caused by evening classes and problems with the group.

3.1.5. Overall Evaluation

Twenty-one students classified the course as great, and 6 as good. One student classified it as average. Comments in the open field included suggestions ("integration with other courses"); "include some lectures, to present theoretical concepts, and decrease the number of practical activities"; "have lectures followed by debates") and praising ("this course lived up to the expectations. Thank you for the effort, dedication and organisation (very important). I want to be a teacher, and I have learned valuable lessons in this course"; "dynamic classes, different from the routine of lectures, very interesting"; "the course was very well structured, with a special concern to give students feedback, in general and individually"; "classes were fun and motivating, with excellent feedback from the teacher. I looked forward to the classes"; "loved it. It wasn't tiresome, and the teacher knew how to stimulate students to get involved"; "great methodology, with incremental prototype building").
4. Reflections and Takeaways

The main aim of this experience was to try gamification out at the undergraduate level in face-to-face classes, particularly as a way of engaging students. This final session summarises the lessons learned and my impressions as a teacher.

4.1. Gamification: Behaviourism Disguised?

Being an enthusiast of constructivism and students' autonomy, and having established to the students that the course would follow a constructivist approach, I found myself in contradiction as I began to identify the similarities between gamification and behaviourism [Skinner 1976]. I was using gamification to: give students rewards in the form of points (positive reinforcement) for every aspect I considered important, or penalise students when failing to achieve in any of these aspects (punishment) - which ultimately translated into their grades; and to encourage competition through rankings, which showed students' progress and achievement to all, enabling comparison. This realisation made me go beyond the hype of gamification to get to its pedagogical basics. Truth be told, I have used a very basic form of gamification, which lacked more sophisticated elements like avatars with different powers and capabilities, a themed general background to the activities, and the use of specific software tools for gamification [Dicheva et al. 2015]. Furthermore, it could be argued that I fell into what Challco et al. (2015) call "pointsification", a simplistic type of gamification that relies a lot on using some form of coercion, even if 'disguised'. The authors suggest that well-thought-out gamified scenarios should apply game elements to model learners’ attitudes and behaviours through persuasion and social influence. Nevertheless, most applications of gamification remain at the basic level, using points, rewards, levels, rankings and feedback, as shown by two recent systematic literature reviews [Borges et al. 2013] [Klock et al. 2015]. The PBL approach (Points, Badges and Leaderboards) [Werbach and Hunter 2012] is considered to present the basic elements of gamification, said to compose the most usual approach in education [Brazil and Baruque 2015]. Klock et al.'s (2015) systematic literature review only found two in seven papers analysed that implemented personalisation, and one that implemented narrative and challenges. In addition, even in sophisticated forms of gamification (often within virtual collaborative environments on digital platforms with customised tools and sometimes based on ontologies [Challco et al. 2015] [Dicheva et al. 2015]), the core concepts of gamification remain strongly linked to behaviourism, as can be noted when Skinner's reinforcement theory [Skinner 1976] is quoted by Challco et al. (2015, p. 502): "(...) the change in the learners’ attitudes and behaviors is learned by operant conditioning, where the consequences of humans’ actions modify the tendency to repeat a behavior. (...) The game actions taken by these game components follow the learners’ actions to reinforce the intended learning behavior defined by the script.". Borges et al.'s (2013) systematic literature review identified the goal of "promoting some kind of behavioural change in students" (p. 237, our translation) in 19 out of 26 selected publications on gamification in education. Other expressions used in papers on this theme reinforce the association between behaviourism and gamification (our translation and stress): "users engage more in developing some activity when they perceive that they are rewarded for that" [Ferreira et al. 2015, p. 512]; "(...) we can remember a typical case of gamification that many of us have lived as children, when we were rewarded with objects like stars for good performance or good behaviour." [Nogueira Neto et al. 2015, p. 667]; "The narratives make the user present the expected
Behaviour control through elements of gamification did not work well in the context where this work took place. Students presented a low level of competition against peers, which was shown by the lack of interest in following closely and frequently the information spreadsheet. The idea of having 'blocked' phases (i.e. unknown to the students) that could be reached at different times by each group did not work either. On the one hand, students were not comfortable with not knowing how the course would progress, mainly as they felt the need to plan their university timetable and activities in the near future considering other courses they were taking. Although this is not shown in the data collected, I was constantly asked about the next activities and the demands for passing during the course. The 'surprise' factor was thus more a source of anxiety than of mystery and fun for the students. On the other hand, groups usually reached the phases at the same class, or, if they did not, the next phase ended up being revealed to all anyway, due to the dynamics of the classes and the information spreadsheet.

So, although the course went well and I received encouraging feedback, the most positive students' reactions and impressions were not related to the aspects of gamification per se, but to other methodological aspects used, namely: practical activities and detailed and constant feedback (for groups and individuals) through the written evaluations provided in the spreadsheet. The combination of these two seemed to be much more important for students than gamification. On the other hand, this finding also reinforces the need to take into account students' preferences and psychological, anthropological and pedagogical factors involved, as described by theories of motivation, human behaviour and game design, all of which matter for the educational outcomes of gamification [Werbach and Hunter 2012].

4.2. The Quest for Ideal Student Assessment

The goal of the assessment criteria established for this course was to cover the broadest possible range of aspects that play a part in learning processes, going beyond traditional evaluation through written exams, merely based on correct answers about the content, many times memorised to be then forgotten but not understood. However, four issues arose. Firstly, the subjectivity of some criteria (e.g. creativity and motivation) led to students' dissatisfaction and frustration for not understanding the reason for receiving few or no points. In some cases, for example, they truly believed they were being creative, while I did not think so - but judging creativity objectively proved to be quite hard.

Secondly, assessment was based on class activities (and deliverables), but did not take into consideration the extra-class group dynamics. As pointed by students, this proved to be a serious flaw in the assessment, as some students did a lot of 'backstage' work that was not of my knowledge. This also penalised shy / introspect students who did not participate very actively in class. I now realise that the criteria chosen reflected my expectations of my favourite student profile, the type who shows their motivation by active, loud and joyful participation, initiative, creative ideas and provocative questions. In this small piece of research, introvert and quiet hard-working students reminded me that the set of evaluation criteria absolutely must contemplate all learning styles. In this regard, literature on gamification suggests models to help instructional designers choosing adequate game elements based on learners' preference and individual
characteristics [Domínguez et al. 2013] [Simões et al. 2013]. However, this is still very hard to implement. Klock et al.'s (2015) systematic literature review recently found that characteristics most taken into consideration are gender, age and type of gamer, which would not account for the difficulties reported here.

Thirdly, there were 9 individual criteria and 7 group criteria. Most of them were to be evaluated during class activities. This meant I had to observe the behaviour of 30 students every class and distribute points accordingly. It was as overwhelming as it sounds, exhausting and most probably, not fair. At mid-term, transforming the points distributed through the quite large set of assessment criteria into grades was extremely challenging, as can be told from the complicated procedure describe in section 2.2.4. In addition, subjectivity of criteria and bias towards a particular student profile led to many grades below 7 at mid-term. Students' feedback and my own reflections made me adopt a different approach for the second half of the course, giving the chance for students to take part in their own evaluation and their peers', while also keeping the group assessment criteria. This approach was better received by the class.

Finally, I have continued to struggle to find optimal ways to fairly assess group work. As in previous classes, there were groups of good friends who divided the work of the several courses they were taking and were very corporative; there were groups of people who had just met in the course and could barely establish proper extra-class communication - they got upset with peers that did not work and told them off to me, especially when they received low group grades as a consequence; and finally there were groups in which some did the work and just ignored those who did not, but I was never explicitly told about it. I made the choice, in the present course, not to interfere in group work, i.e. I avoided the role of 'investigator' who discovers and punishes students who do not contribute, and assumed the position of blindly trusting everyone. Not all students were satisfied with such approach. The idea that individual assessment criteria would compensate for deficiencies in evaluating the work of group members did not hold, as revealed by peers telling others off, and suggestions of "interviewing" group members to find out who did the work, or what was "really going on". Such 'investigative role' is not what I look for as a teacher, although this experience seems to indicate that a combination of peer assessment and guided self-assessment might be a good way forward.

4.3. Lectures: to Keep or Not to Keep?

Giving lectures has been increasingly frustrating in the courses I teach, in the context I work. Teaching in the evening means having tired students, who in many cases spent the day at work and fall asleep during lectures, the content of which more often than not will be heard and forgotten by the students. In addition, a lot of what one can put into lectures on educational technologies and human-computer interaction is one or two clicks away in the Internet, and thus I truly believe that the teacher must go beyond. Although planning practical activities that enable learning for all meetings with the students is hard and time-consuming, it is rewarding, from my point of view. Surprisingly, however, several students missed having lectures in the course, and suggested a more balanced approach between lectures and practical activities. Their feedback showed that they feel more confident when they hear some content delivered by the teacher, than when they are exclusively given instruments to construct this knowledge themselves. Even though this is probably a reflect of the format of the educational system they are used to, and could
be changed to a more constructivist approach, such students' feelings should not be ignored presently.

4.4. Where to Go from Here

The experience described in this paper was a - maybe naïve - attempt at gamifying the classroom within a constructivist philosophy, and reaching an innovative course format where students would feel motivated and empowered. It is probably fair to say that the (behaviourist) aspects of gamification integrated into this specific course did not increase students' motivation, but more constructivist methods like practical activities and constructive feedback were taken in high account by the students and mentioned as the main highlights of the course. Gamification would have to be applied in a richer and more elaborated manner, but, more importantly, more adherent to the particular context and students' profile. More questions than answers arose from the experience, but which lead to an exciting starting point for reflections and methodological refinements:

- Re-introduce a few lectures with the main theoretical concepts and balance them with closely related practical activities;
- Include peer assessment and guided self-assessment in students' evaluation;
- Have assessment criteria that contemplate extra-class work and the diversity of learning styles;
- Reduce the subjectivity of assessment criteria and plan assessment in a feasible and scalable way;
- Give students personalised, constructive feedback as often as possible.

Last but not least, the great challenge that arises is: can gamification and constructivism dialogue at all? And, more specifically, is it possible to combine gamification with learners' empowerment and autonomy? Seixas et al. (2016) go in this direction when they argue that the "use of gamification in education should not be restricted to giving points (...) the use of other gaming strategies, allows the student to awaken creativity (...) and build learning situations in which they are free to make choices" (p. 50) and "it is necessary to think about the student not only as a “player” who will receive a reward for his effort, but he should be responsible for building his knowledge and gamification is an opportunity to make this process funnier and more challenging according to his skills” (p. 59). However, Seixas et al. have also employed a method heavily based on behaviour rewarding and it is not clear how freedom and empowerment fit into the model. Such combination, if possible, will probably lead to a different model for learning, where motivation from games would somehow be reached in learning environments without so much emphasis on competition and on behaviour control and manipulation, but instead valuing empowerment and autonomy.

Acknowledgements

I deeply thank all students of the 2015.2 class, who had to bear with my experimental ideas, helped me with their feedback, and motivated me with their good humour.
References


Laster, J. (2010). At Indiana U., a class on game design has students playing to win. Available at: http://chronicle.com/blogs/wiredcampus/at-indiana-u-a-class-on-game-design-has-students-playing-to-win/21981.


