

RESEARCH

Open Access

Gamification for smarter learning: tales from the trenches

Gabriel Barata*, Sandra Gama, Joaquim Jorge and Daniel Gonçalves

* Correspondence:
gabriel.barata@gmail.com
INESC-ID/Instituto Superior Técnico,
Universidade de Lisboa, Rua Alves
Redol 9, 1000-029 Lisbon, Portugal

Abstract

Gamification has been explored recently as a way to promote content delivery in education, yielding promising results. However, little is known regarding how it helps different students experience learning and acquire knowledge. In this paper we study and analyze data from a gamified engineering course, to search for distinct behavior patterns. We examined data collected from two gamified years, between which game changes took place. By clustering students according to their performance, we identified three distinct student types, common to both years: Achievers, Disheartened, and Underachievers. Interestingly, in the second year a new type of student emerged: the Late Awakeners. In this paper we carefully describe each student type, and explain how gamification can provide for smarter learning by catering to students with different profiles. Furthermore, we discuss how our findings, both in gamification and cluster analysis can be used to develop adaptive and smart learning environments.

Keywords: Gamification; Education; Cluster analysis; Student participation; Adaptive learning

Introduction

Good games are natural learning machines that, unlike traditional educational materials, can deliver information both on demand and within context (Gee, 2003). By design, they prevent players from becoming either bored or frustrated, while allowing them to experience flow (Chen, 2007; Csikszentmihalyi, 1991). Thus, it is not surprising that videogames have long been used to transmit knowledge (de Aguilera and Mendiz, 2003; Squire, 2003), given the remarkable educational possibilities they unveil (Bennett et al., 2008; O'Neil et al. 2005; Prensky, 2001).

Gamification draws on the motivational qualities of games to entice users to adopt specific behaviors. To this end, it includes game design elements in non-game processes (Deterding et al., 2011a; Deterding et al., 2011b), to make them both more fun and engaging (Reeves and Read, 2009; Shneiderman, 2004). Gamification has been used for many different purposes, ranging from marketing (Zichermann and Cunningham, 2011), to fitness and health (Brauner et al., 2013), helping in driving instruction (Fitz-Walter et al., 2013), improving productivity (Sheth et al., 2011) and promoting eco-friendly driving (Inbar et al., 2011).

Gamification in education is a growing topic and a variety of studies already show that gamification has the potential to significantly increase student activity

(Denny, 2013) and performance (Domínguez et al., 2013). However, little is known about how different students experience gamified learning and what different strategies and traits they present. Therefore, understanding how different students learn in gamified environments may provide a stepping-stone to develop new and adaptive gamified courses, capable of adjusting to different students' particular needs and thus improving their learning experience.

In a previous work (Barata et al., 2013) we described a two-year study to gamify a college course, Multimedia Content Production (MCP), in order to engage students and encourage them to be more participative. We found that students were both more participative and proactive as compared to previous non-gamified editions of the course. Furthermore, they also considered it as being more motivating and interesting than other "regular" courses. Yet, we knew nothing about how different students performed under our gamified experience, and how it could be designed to adapt to their different traits, needs and wants.

In this paper we describe how we used clustering algorithms to analyze student progression on two consecutive years and identify different student profiles. While we found three distinct student types on the first year, a fourth type emerged in the second year. We will present a detailed description of each student type and propose cluster analysis in our gamified approach as a tool to support adaptive gamified learning environments.

Related work

Several studies reported that using serious games to teach resulted in significant improvements in student understanding, diligence and motivation, at different academic levels, from grade school (Lee et al., 2004), through high school (Kebritchi et al., 2008), to college (Coller and Shernoff, 2009), and in diverse fields of knowledge, such as biology (McClellan et al., 2001), programming (Moreno, 2012), numerical methods (Coller and Shernoff, 2009) or electromagnetism (Squire et al., 2004). This suggests that games can indeed provide smart ways to improve people's learning experience and outcomes.

Defining gamification

Gamification is a novel concept that might easily be confused with games in general or other game-based approaches used in learning, namely serious games. Game is commonly defined as "a physical or mental activity or contest that has rules and that people do for pleasure" (Merriam-Webster.com, 2015), although its definition has been widely debated and has assumed different forms (Caillois, 1991; Crawford, 2003; Feyerabend, 1955). While games are usually designed to entertain, serious games are games conceived to educate and not necessarily to entertain, although they might be entertaining as well (Michael and Chen 2006). Gamification differs from (serious) games in that the latter consist of using full-fledged games (without a purpose beyond entertainment), which usually simulate real-world or fantasy scenarios and events, while the former adopts game design elements (only) in contexts or processes that are not games (Deterding et al., 2011a), usually encompassing real-world activities. However, both allow for gameful interactions (McGonigal, 2011). Although there is not a formal list of game elements, the most consensual ones used in gamification seem to be (Crumlish and

Malone, 2009; Kim, 2008; Zichermann and Cunningham, 2011): a) Points, used to reward user actions; b) Levels, which together with points indicate progress and serve as a means of comparison; c) Leaderboards, to spur competitiveness and allow users to compare themselves to others; d) Challenges, which are tasks users must undertake to progress in the gaming experience (that should be meaningful and fulfilling); and e) Badges, to serve as rewards for accomplishing challenges, which grant bragging rights and appeal to our need to collect.

Gamified learning

Gamification has a great potential to help people acquire and improve skills. For example, Microsoft Ribbon Hero (Microsoft, 2015) is an add-on that uses points, badges and levels to motivate people to explore Microsoft Office tools. Jigsaw (Dong et al., 2012) teaches Photoshop by using a jigsaw puzzle to challenge players to match a target image. Users reported that they were able to better explore the application and discover new techniques. GamiCAD (Li et al., 2012), a tutorial system for AutoCAD, allows users to perform line and trimming operations, to help NASA build a spacecraft. Using the gamified system, people completed tasks faster and found the experience to be both more engaging and enjoyable.

Although game elements in the aforementioned gamified systems were integrated by hand, several third-party gamification services exist that offer a variety of solutions. Badgeville (2015) is a commercial service that gamifies other services and offer means to measure and influence user behavior. Mozilla Open Badges (Mozilla, 2015), on the other hand, is a free online program focused on learning, which issues digital badges to recognize skills and achievements. Prominent online learning services also use visual elements to track student progress. Examples of these are Khan Academy (2015), which allows students to learn by reading and watching videos and then performing exercises, and Codecademy (2015), which uses interactive exercises to teach. Both use points and badges to give feedback and transmit progress.

Gamifying education is a growing topic, with some studies presenting beneficial effects for the students' learning experience. Sheldon (2011) showed how a conventional course could be turned into an exciting game, where students start with an F grade and go all the way up to an A+, by completing challenges and gaining experience points. Domínguez et al. (2013) made a comparative study of an e-learning ICT course, during which students could take optional exercises, either by reading a PDF document or via a gamified system. In the latter, students were awarded with badges and medals on completion. While students who completed the gamified experience performed better in practical assignments and had higher overall score, they appear to have performed poorly on written assignments and participated less on class activities. Cheong et al. (2013) used a gamified quiz to evaluate IT undergrad students, where they earned points for answering questions and used a leaderboard to compare scores with other students. Students reported that the quiz helped improving both their learning effectiveness and their grades, and also their enjoyment and engagement. However, no empirical results other than self-reports were presented. On another study, Hakulinen et al. (2013) added achievement badges to an online learning environment where students completed automatically assessed interactive exercises about data structures and algorithms. Results suggest

achievement badges had a significant impact on student behavior, with more of them getting more perfect scores. Yet, only a small group of students was particularly motivated to collect badges.

Player classification

People that play games enjoy doing it for a variety of reasons. Thus, several models have been proposed to categorize players according to their behavior and preferences. One of the first models was proposed by Bartle (1996, 2004), which classifies Multi User Dungeon (MUD) users into four types: Socializers, who prefer to interact with other players; Killers, who enjoy imposing and causing distress upon others; Achievers, that like to assign themselves game goals and set out to achieve them; and Explorers, players that try to explore and find out as much about the world as possible. Other modes emerged, such as the Demographic Game Design 1 (DGD1), proposed by Chris Bateman (Bateman and Boon, 2005) and primarily based on the Myers-Biggs personality model (Myers and McCaulley, 1988), which allowed players to be classified with combined playing styles, in opposition to the Bartle types. A posterior model was proposed, named BrainHex (Nacke et al., 2011), which categorizes players into seven types based on neurobiological responses inherent to playing games (Bateman and Nacke, 2010). Player classification models encode preferences and even bodily responses from people that play games.

Assessing student achievement

Research has also been performed on what differentiates students, regarding their behavior and achievement. Particularly, distinction has been made between Achievers and Underachievers on gifted students (Diener, 1960; Farquhar and Payne, 1964; McCoach and Siegle, 2003). While highly achieving students are not considered at risk of failing to accomplish their academic potential, underachievers present a discrepancy between ability and performance (Dowdall and Colangelo, 1982; Whitmore, 1980). Research shows that factors affecting underachievement are manifold, which include attention-deficit/hyperactivity disorder, hearing impairment, nontraditional learning styles, and emotional problems (Moon and Hall, 1998). Underachievers usually present low academic self-perception, negative attitudes towards school, teachers and classes, low goal valuation and low motivation and self-regulation (Dowdall and Colangelo, 1982; McCoach and Siegle, 2003; Whitmore, 1980).

Taylor (2002) proposed a model for students of e-learning courses, who participated mostly via posts on a discussion board. Students were classified into three types: Workers, Lurkers and Shirkers. Workers are characterized by an above average number of posts and large amount of visits to the board. "These students were continuously involved in discussions and were often among the first to post a message, and to respond quickly to other messages, thereby creating "threads" of ongoing dialogue between students". Lurkers made a below average number of posts but read the discussion board regularly. Lastly, the Shirkers "contributed only one third of the average number of postings or less to the Discussion Board, and similarly visited this part of the site on less than fifty percent of the group average". While Workers and Lurkers appear to have similar levels of

performance, even though their interaction patterns differed, Shirkers perform poorly and have higher risk of failure.

Fleming and Mills (2001) proposed the VARK model that encodes different learning styles. The model comprises four modalities: Visual, which covers preferences for learning using graphic material, such as maps, charts, diagrams, symbolic arrows, hierarchies, and so on, but excludes still pictures, movies and slides; Aural covers preferences for “information that is ‘heard or spoken.’”, which include lectures, group discussions, email, among other mediums; Read/write, which covers preferences for “information displayed as words”, such as manuals, reports, essays, assignments, PowerPoint slides, Google and Wikipedia; and Kinesthetic, which refers to the “perceptual preference related to the use of experience and practice (simulated or real)”. This modality encodes a preference to learn through experience, demonstrations, practice and simulations. Because each modality is assigned an independent score, this model supports multimodality. Even if the student types found in the literature apply, we don’t know if they are the most relevant within a gamified learning experience, given the lack of studies to assess its applicability in this field.

Despite the broad body of knowledge on categorizing players and student achievement, performance and learning styles, there are no studies to our knowledge that explored how different students react to gamified learning, which we believe is key to designing smarter and adaptive learning environments.

Playing the mcp game

Multimedia Content Production is a Master of Science (MSc) level gamified course on Information Systems and Computer Engineering, at *Instituto Superior Técnico*, the School of Engineering of the University of Lisbon. The course follows a blended learning method, through which students attend both theoretical lectures and lab classes, but also engage in discussions and complete assignments online using the Moodle (2015) platform. Theoretical lectures cover multimedia concepts such as capture, editing and production techniques, multimedia standards, copyright and Digital Rights Management among others. In lab classes, they learn concepts and tools on image, audio and video manipulation, and exercise them via regular assignments. In the first year we had 35 students, two professors and two lab assistants, and in the second we had 52 students, the same two professors and one lab assistant.

Course design

Instead of collecting traditional grades at the end of the course, students started with zero experience points (XP) and acquired additional XP by completing course activities, which included quizzes (20 % of total course XP in the first year and 10 % in second year), a multimedia presentation (20 %), lab classes (15 %), a final exam (35 %) and a set of collectible achievements (10 % on the first year and 20 % on the second, plus 5 % extra). Achievements required students to complete course activities such as attending classes, finding bugs in course materials or completing challenges and quests, and they earned students XP and badges on completion. Badges could either require a single iteration or up to three to be completed, with each iteration being worth a specific amount of XP and a badge. There were a total of 25 and 28 achievements to complete in the first and second years, and up to 61 and 71 badges respectively could be earned

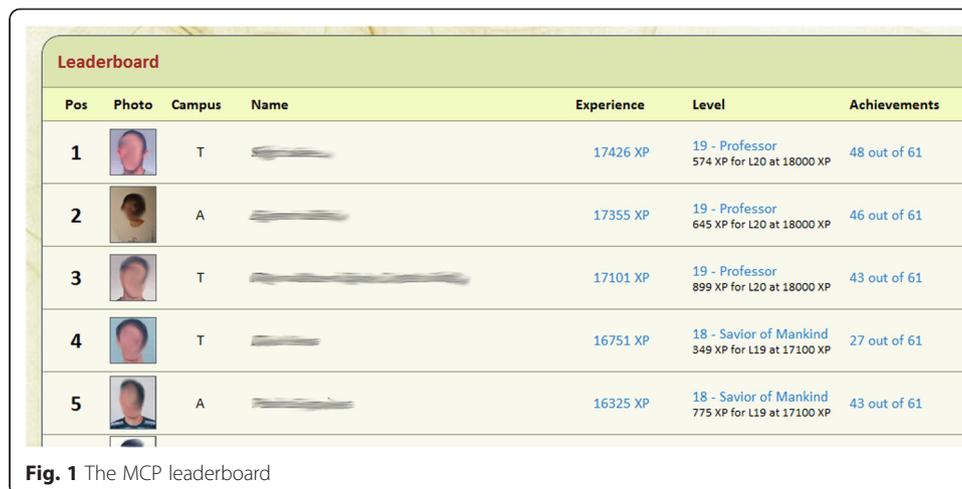
by students. Challenges and Online Quests were time-limited activities that required students to complete specific tasks. In the first year there were six Theoretical challenges, three Lab challenges and three quests, and in the second year there were 11 Theoretical challenges, five Lab Challenges and two quests. Challenges and Quests required students to post materials to course fora in response to challenges posted by faculty. When students accumulated enough XP their experience level would increase. Experience levels maxed out at 20, where reaching level 10 was required to pass the course.

The entry point of the gamified experience was the leaderboard, a public webpage linked from the forums, which allowed students to track both theirs and others' progress and where scores were sorted in descending XP order, as shown in Fig. 1. By clicking a row, a dedicated page for the respective player was displayed, which included a rich dashboard with several charts portraying student progress (see Fig. 2), and a list of the completed achievements and badges earned so far (see Fig. 3). Extra experience points were allocated throughout some of the challenges, which would earn XP above the maximum grade, although the final grade would always be clamped to 100 %. The main goal of this decision is to allow students to get the same amount of grade through different paths, thus enabling them to do more of what they like or to earn the XP lost on failed assignments, thus potentiating learning from trial and error.

The leaderboard's webpage was developed using HTML and Javascript, and it was responsible for displaying all game design elements. Students interacted with the system via posts on the Moodle forums and some class activity was manually logged by faculty, using a dedicated spreadsheet. A script was then run several times a day on our server, which processed the spreadsheet and the Moodle logs to update the leaderboard's web page with the new student data.

Student feedback and courser evolution

By analyzing data collected at the end of the first year, we could see that our gamified course had a few problems. Indeed, challenges were not distributed evenly over the term, and the course lacked interesting goals for students to pursue during the second half of the term. Furthermore, students stated that there should be more ways to



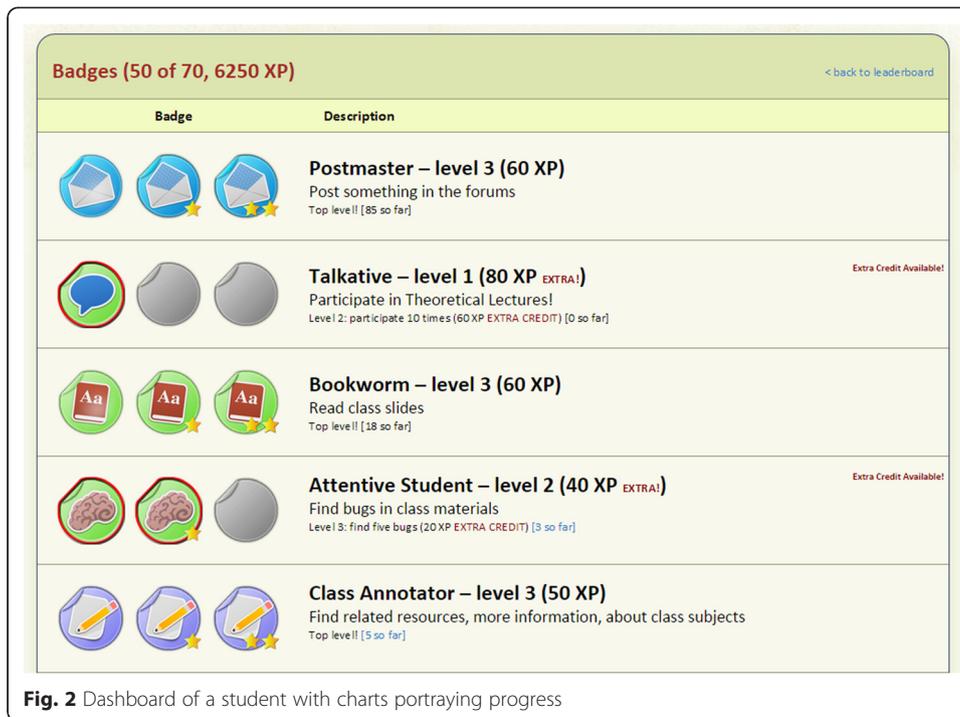


Fig. 2 Dashboard of a student with charts portraying progress

promote cooperation, oral participation in class should be rewarded, and the achievements were in general not well rewarded. Students also complained about the lack of autonomy and room for creativity. To make the gamified approach more flexible and cater to different students, we performed a few changes in the second year, via the removal of 2 redundant achievements which were ignored by the students, and by adding five new achievements. Two rewarded cooperation in labs. One – Guild Warrior – was awarded when all students in a lab class performed above 80 %. Another – Guild Master – recognized students whose lab class was the best. We created a Talkative Badge to reward oral participation. Another

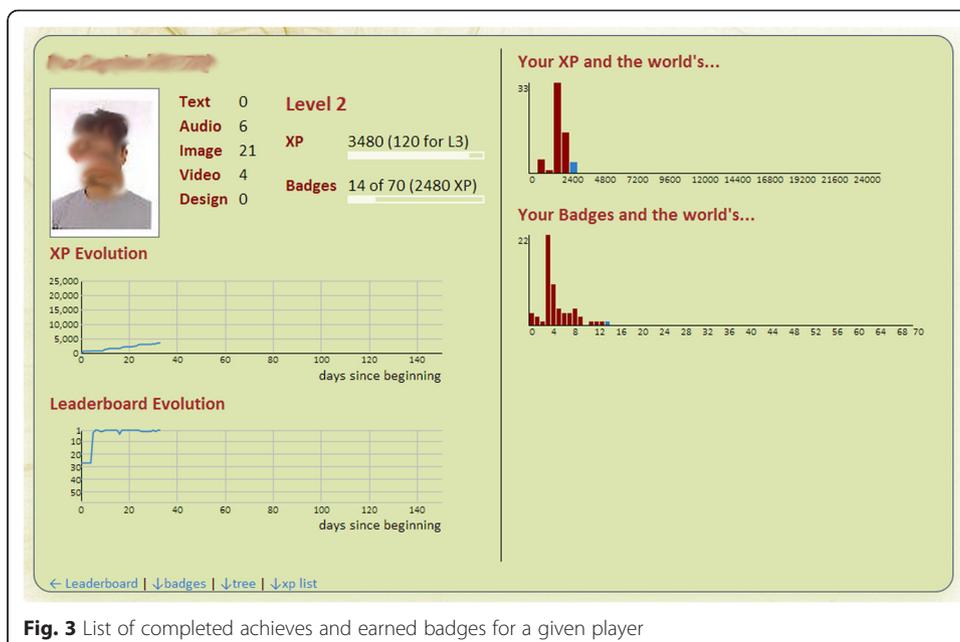


Fig. 3 List of completed achieves and earned badges for a given player

rewarded students for timely responses to challenges (Proactive badge). Finally, the Archivist Badge rewarded students for compiling challenge results. To make achievements more appealing, we re-graded the course so that achievements would be worth 20 % of total XP and quizzes only 10 %. To create more participation opportunities and reach out to more students, we added five theoretical and two lab challenges and removed one quest (which was very similar to one of the other quests), and carefully distributed them over the semester. With these changes, in the second year we saw significant improvements in student participation and marginal gains in terms of final grade (Barata et al., 2013).

Student clustering

Our approach aims to provide a flexible experience that can adapt to students in real-time, by providing rich feedback tailored according to the particularities of every student. To this end, we decided to study how different students experience the gamified course and what strategies they adopt towards learning in this context. To profile students into types, we needed to devise a measure, which would be both adequate to track progress over time and usable with a clustering algorithm. We considered both accumulated XP and rank as candidate measures, but rejected the latter since students with equal performance could never be at the same rank. Accumulated XP plotted over time revealed a few distinctive patterns, which seemed to support the goodness of this measure. Therefore, we used cluster analysis to mine these patterns and identify different student categories.

We performed cluster analysis using Weka (2015), a collection of machine learning algorithms for data mining tasks in Java. Several algorithms were available to perform the analysis, such as the K-Means (MacQueen, 1967), the COBWEB (Fisher, 1987), and the DBSCAN (Ester et al., 1996). However, these algorithms present several limitations. The first requires the number of clusters to be provided as a parameter, which we did not know. The second assumes that that probability distributions on separate attributes are statistically independent of one another (Sharma et al., 2012), but correlation between attributes often occur. The third does not work well with high-dimensional data, such as ours (one dimension for each day).

We opted for the Expectation-Maximization (EM) algorithm (Dempster et al., 1977), which does not require the number of clusters to be specified beforehand and works well with small datasets (Sharma et al., 2012). This algorithm assigns a probability distribution to each instance, which indicates the probability of it belonging to each of the clusters. It then decides how many clusters will be created by cross validation. As parameters we used 100 as the maximum number of iterations, 1.0E-6 as the minimum standard deviation, 100 as the random number seed and did not specify the number of cluster, and these were the default parameters. EM received as attributes the amount of accumulated XP per each day and students (the instances) were assigned to the cluster with the highest probability, thus grouping them by similarity or dissimilarity of XP acquisition.

During the first few days, most students had zero XP, either because they were not fully enrolled in the course or because there was still no significant activity. This makes their activities to look alike, which might mislead the clustering algorithm to group all students into a single cluster. To avoid this, we excluded from the analysis the initial days that satisfied either one of the following criteria: a) there were still students with

incomplete enrollment and b) students were tied up at zero XP due to lack of activity. As a result of this filtering, four and 15 days were excluded from the first and second year samples respectively. The second year, saw a larger number of days excluded since the professors were attending a conference, which limited student participation during the first few days.

In this section we characterize the clusters found in both years, by taking into account student performance, online participation, and qualitative feedback from satisfaction questionnaires. Given that we could not verify the normality of our samples, owing to the clusters' small size, we checked for inter-cluster differences using a non-parametric test, the Kruskal–Wallis one-way analysis of variance, with post-hoc Mann-Whitney's U tests and Bonferroni correction. The resulting significant level is given by $\alpha = 0.05/3 = 0.016$ for the first year and $\alpha = 0.05/6 = 0.0083$, which was the one used in our tests and reported in Tables 1 and 2.

First year

The EM algorithm discovered three distinct clusters on the first-year sample. The first consisted of 12 students that strived to be the best. They were characterized by a distinctively higher average XP accumulation curve, which suggests that these students were most of the time ahead of others, as shown by Fig. 4. Here, every abrupt change corresponds to a major XP acquisition, and these students seemingly grabbed every opportunity to acquire XP. We named them the Achievers, and they were typically

Table 1 Student performance metrics per cluster, on the first year

Property	Achievers	Disheartened	Underachievers	All	Significant differences (p < 0.016)
Quizzes Grade (%)	82.22	77.50	66.00	74.19	(A, C)
Labs Grade (%)	95.96	85.45	81.32	87.28	(A, B), (A, C)
Presentation Grade (%)	91.08	78.88	76.27	81.94	(A, B), (A, C)
Exam Grade (%)	76.62	63.37	60.82	66.82	(A, B), (A, C)
Final Grade (%)	90.16	77.82	71.88	79.51	(A, B), (A, C), (B, C)
Posts (#)	36.58	22.88	16.67	24.91	(A, C)
First Posts (#)	4.92	1.63	2.00	2.91	None
Reply Posts (#)	31.67	21.25	14.67	22.00	(A, C)
Challenge Posts (#)	18.08	12.00	9.47	13.00	(A, C)
XP from Challenges (%)	97.37	89.47	72.46	84.89	(A, C)
Theoretical Challenge Posts (#)	12.42	6.75	6.27	8.49	(A, B), (A, C)
XP from Theoretical Challenges (%)	98.15	87.50	75.93	86.19	None
Lab Challenge Posts (#)	5.67	5.25	3.20	4.51	None
XP from Lab Challenges (%)	96.67	91.25	69.33	83.71	(A, C)
Quest Posts (#)	5.25	1.50	1.33	2.71	(A, B), (A, C)
XP from Quests (%)	98.00	24.50	36.53	54.86	(A, B), (A, C)
Badges (#)	38.17	29.63	21.67	29.14	(A, C)
XP from Achievements (%)	78.24	60.32	49.63	61.88	(A, B), (A, C)
Completed Achievements (#)	11.75	8.50	5.20	8.20	(A, C)
Explored Achievements (#)	17.67	14.00	12.07	14.43	(A, B), (A, C)

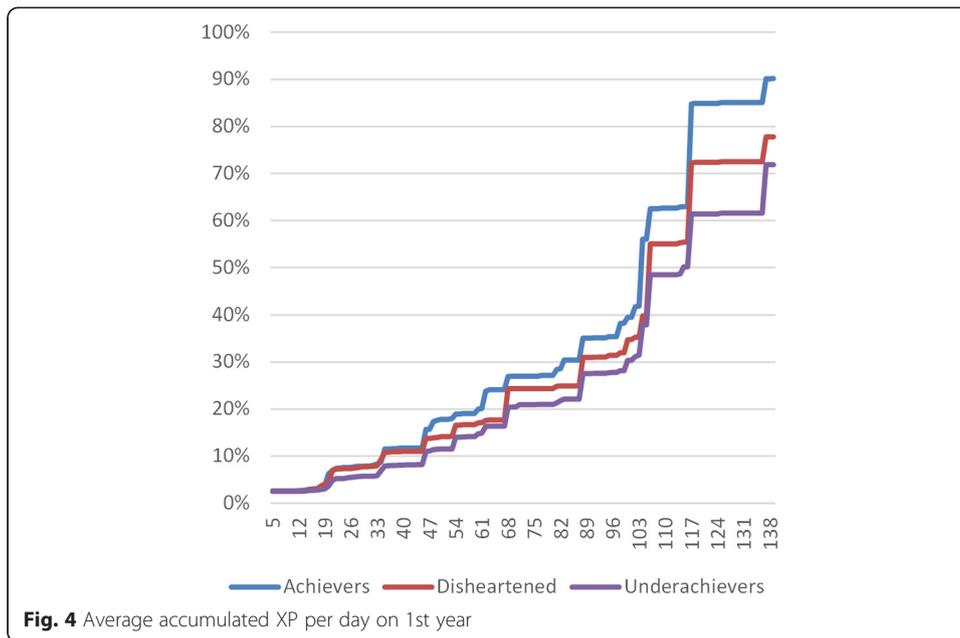
Table 2 Student performance metrics per cluster, on the second year

Property	Achievers	Late awakers	Disheartened	Underachievers	All	Significant differences (p < 0.016)
Quizzes Grade (%)	93.67	89.78	81.61	77.95	85.37	(A, C), (A, D), (B, D)
Labs Grade (%)	98.90	96.10	95.54	91.65	95.38	(A, C), (A, D)
Presentation Grade (%)	91.10	89.57	84.82	85.54	87.58	None
Exam Grade (%)	80.75	74.52	70.09	65.65	72.31	None
Final Grade (%)	98.74	91.31	86.00	79.22	88.29	(A, B), (A, C), (A, D), (B, C), (B, D), (C, D)
Posts (#)	73.60	50.40	33.36	15.38	41.52	(A, C), (A, D), (B, D), (C, D)
First Posts (#)	4.70	2.20	1.14	0.77	2.04	(A, C), (A, D)
Reply Posts (#)	68.90	48.20	32.21	14.62	39.48	(A, C), (A, D), (B, D), (C, D)
Challenge Posts (#)	46.40	37.33	27.93	12.38	30.31	(A, C), (A, D), (B, D), (C, D)
XP from Challenges (%)	100.00	94.67	95.51	72.75	90.44	(A, D), (B, D), (C, D)
Theoretical Challenge Posts (#)	29.50	21.93	20.14	7.38	19.27	(A, D), (B, D), (C, D)
XP from Theoretical Challenges (%)	100.00	92.73	97.40	69.93	89.69	(A, D), (B, D), (C, D)
Lab Challenge Posts (#)	16.90	15.40	7.79	5.00	11.04	(A, C), (A, D), (B, C), (B, D)
XP from Lab Challenges (%)	100.00	97.95	92.31	77.51	91.72	(A, D), (B, D)
Quest Posts (#)	3.50	4.33	2.93	1.77	3.15	(B, D)
XP from Quests (%)	100.00	94.44	78.57	35.90	76.60	(A, D), (B, D)
Badges (#)	47.30	37.53	32.71	26.00	35.23	(A, B), (A, C), (A, D), (B, C), (B, D), (C, D)
XP from Achievements (%)	97.07	85.68	77.56	60.37	79.36	(A, B), (A, C), (A, D), (B, C), (B, D), (C, D)
Completed Achievements (#)	14.90	9.60	8.57	6.08	9.46	(A, B), (A, C), (A, D), (B, D), (C, D)
Explored Achievements (#)	20.20	17.20	15.43	13.08	16.27	(A, B), (A, C), (A, D), (B, C), (B, D), (C, D)

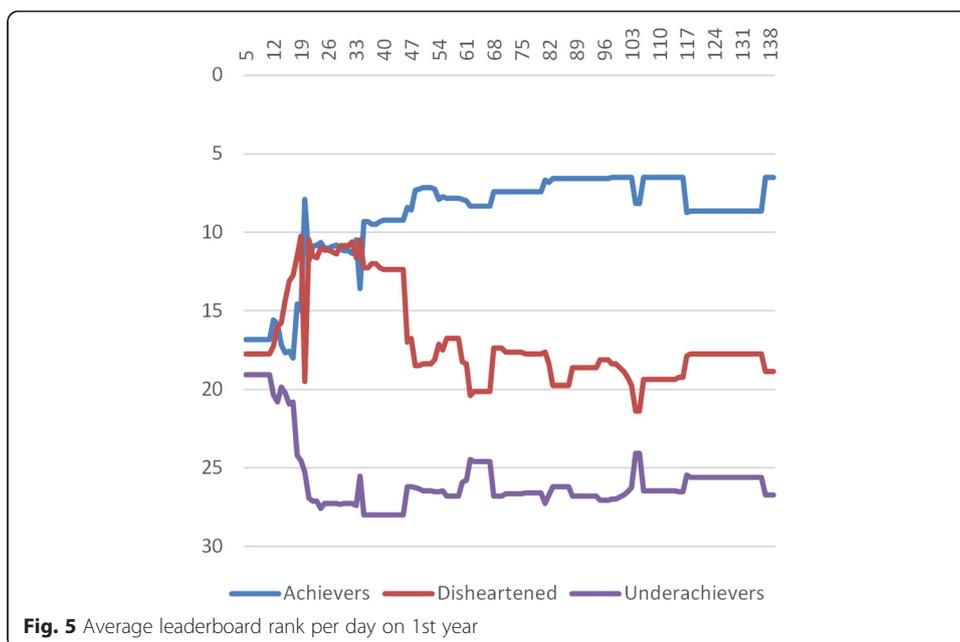
found at the top of the leaderboard (see Fig. 5). A second cluster gathered 14 students, which shared the lowest XP accumulation curve and had fewer pronounced slopes. These people spent most of the time at the bottom of the leaderboard. We named them the Underachievers since they seemingly showed both lower interest and lesser engagement with the course. The third cluster included eight students who started out at a pace similar to that of the Achievers, but soon started to lose ground and stabilized on XP acquisition rates between those of the Achievers and the Underachievers. Their average leaderboard positions, which started close to the Achievers, dropped to the Underachievers' level as the semester evolved. For this reason we named these students the *Disheartened*.

Cluster performance

As shown in Table 1, Achievers had the highest quiz, lab, presentation, exam and final grades, which reflects a high level of performance. They also performed the most posts, both first and reply posts, which reveals high levels of participation and proactivity,



even though significant differences were only observed in comparison with Underachievers and not for the number of first posts. These students also performed the most Challenge posts and earned the most XP from them, although no significant differences were observed for the amount of XP received from Theoretical Challenges and for the number of Lab Challenge posts. Achievers also participated and earned the most from Quests and the Achievements in general, presenting the highest number of accumulated XP from Quests and respective number of posts, number of acquired badges, number of completed and explored achievements and the most XP from Achievements.



Satisfaction questionnaire responses

The Disheartened students had below average performance and participation, with values between those of the Achievers and Underachievers. Exceptions were the number of first posts and the amount of XP acquired from the Quest, which were the lowest observed, although these differences as compared to other clusters were not significant. The Underachievers presented the worst final grade, a reflex of the worst performance and participation level on almost every aspect.

We asked students to complete a satisfaction questionnaire by the end of the term, where they had to rate statements using a 5-point Likert scale. Achievers considered that the course was more motivating (mode: 5) and interesting (mode: 4) than regular courses [1-Much less; 5 - Much more], and felt they were playing a game instead of just taking a course (mode: 4) [1 - Not at all; 5 - A lot]. Disheartened students and Underachievers also found our course to be more motivating (mode: 4) than other non-gamified courses [1-Much less; 5 - Much more], but mildly felt they were play a game (mode: 3) [1 - Not at all; 5 - A lot]. Disheartened students also considered the course to be more interesting than other courses (mode: 4, 5), and so did Underachievers (mode: 4) [1-Much less; 5 - Much more].

Second year

In the second year we discovered four student clusters. The first cluster exhibited an XP accumulation curve with a similar behavior to that of the previous year. This cluster included ten students that were ahead of all others, and showed the steepest slopes (see Fig. 6). These people also struggled for the top positions on the leaderboard (see Fig. 7). Given these similarities, they retained the name *Achievers*. A second cluster, composed by 13 students, shared many similarities with the Underachievers of the first year. These students were characterized by lower XP accumulation curves, small slopes, and occupied typically low positions on the leaderboard. However, in the second year these

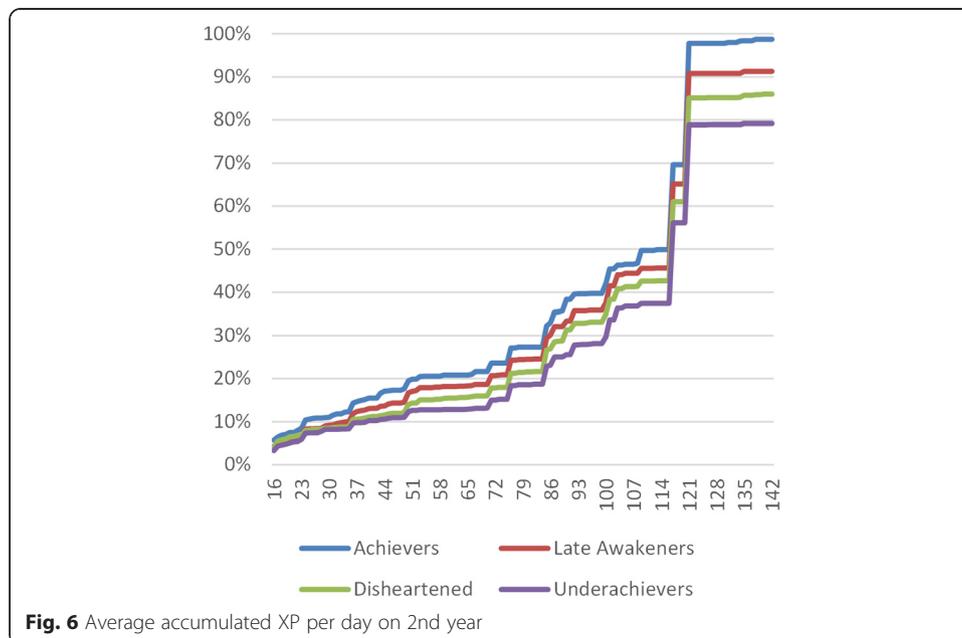
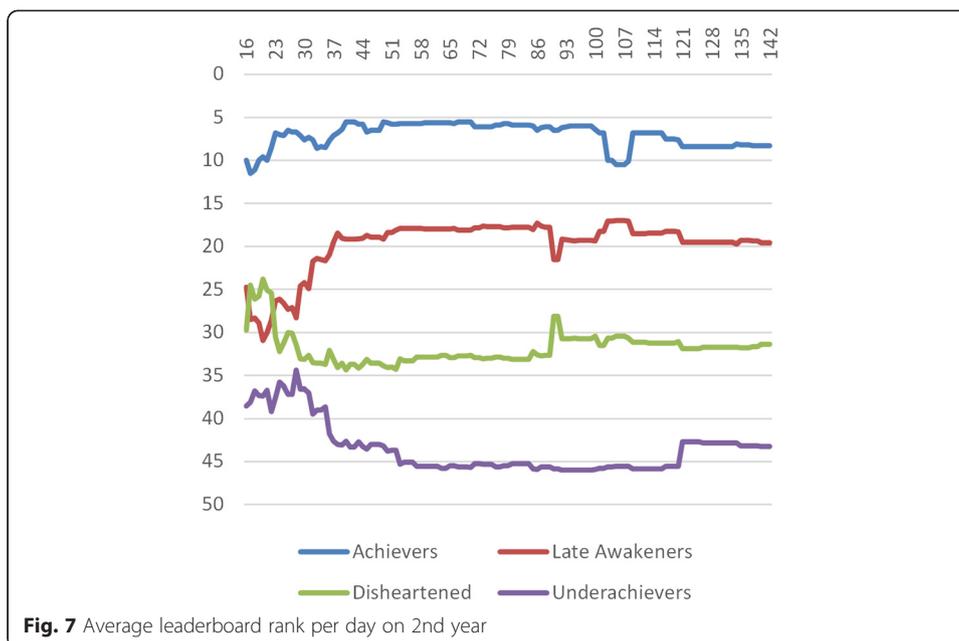


Fig. 6 Average accumulated XP per day on 2nd year



students presented a cumulative behavior closer to that of the other clusters during the first four weeks, and only then major discrepancies started to appear. These students kept the name *Underachievers*. A third cluster, which included 14 learners, displayed some similarities to the Disheartened students of the first year, with an XP accumulation curve close to that of the Achievers during the first weeks, and then occupying a middle-ground between Achievers and Underachievers. Leaderboard progression over time was also similar to that of the corresponding first year cluster. Given these similarities, we named these the *Disheartened* students.

Surprisingly, we found a fourth cluster composed by 15 students, characterized by an XP accumulation curve with some similarities to those of Achievers and Disheartened students. However, these students accumulated on average more XP than the latter student type, but less than the former. The average ranking progression of learners in this cluster resembles the inverse of Disheartened students, resulting in what seems to be like a decline in rank at the beginning of the course followed by a recovery. Based on this behavior we labeled these people the *Late Awakeners*.

Cluster performance

Just as in the previous year, Achievers presented the highest quizzes, labs, presentation, exam and final grade (see Table 2). However, these differences were not significant for the presentation and exam grades. These students made the most posts, both first and reply posts, which again denotes a high level of participation and proactivity. They earned the most from challenges, quests and the achievements, actually achieving 100 % of the grade allocated to challenges and quests. They also participated the most on challenges and the achievements in general, but made the second highest number of posts on the Quests.

Late Awakeners presented above average performance and participation levels, situated between those of the Achievers and the Disheartened students. These students had the second highest final grade, earned the second most from achievements and made the second most posts. They participated the most on the Quests, and the second most on the

Challenges, although they have only earned more XP than the Disheartened students on the Lab challenges and the Quests. Similar to the previous year, the Disheartened students presented below average performance and participation levels, on almost every aspect of the course. Exceptions were the lab grades and the amount of XP earned from challenges and quests, which were slightly above average. These students also had the worst presentation grade, although differences were not significant as compared to any other cluster. Underachievers presented the worst final grade, again a result of the worst average performance and participation levels.

Satisfaction questionnaire responses and feedback

When filling in the end-term satisfaction questionnaire, Achievers considered the course as being more motivating (mode: 5) and interesting (mode: 4) than regular courses [1-Much less; 5 - Much more]. While these students felt they were playing a game instead of just taking a course (mode: 4), other students did not share that opinion (mode: 3) [1 - Not at all; 5 - A lot]. Late Awakeners also found the course as being more motivating than other courses (mode: 4, 5), and so did Underachievers (mode: 4), and less the Disheartened students (mode: 3, 4) [1-Much less; 5 - Much more]. Late Awakeners and Underachievers also deemed the course as being more interesting than regular courses (mode: 4), while Disheartened students shared a neutral view on this (mode: 3) [1 - Not at all; 5 - A lot].

Students reported mixed feelings towards the newly introduced achievements. Regarding the Guild Warrior achievement, Achievers collected on average 2.2 badges, Late Awakeners 2.3, Disheartened students 2.4 and Underachievers 2.6, although the differences were not significant. The Guild Master achievement was harder to get and thus Achievers were able to collect 1.6 badges on average, Late Awakeners 1.3, Disheartened students 1 and Underachievers 0.8. Although students profited from these achievements, at the end of some classes, some students often complained with the faculty staff and blamed others for not obtaining the Guild Warrior badge, although no apparent effort was made to help them instead, which was the main intent. This idea is supported by comments made by four students in the satisfaction questionnaire, which claimed that the Guild achievements made them feel pressured to help others, would raise everybody's grades, and that it was not up to them to complete those achievements (they were only responsible for their part). These views were posteriorly confirmed by surveys made by email to these students.

Achievers collected on average 1.6 Talkative badges to reward their in-class participation, while Late Awakeners, Disheartened and Underachievers received 0.9, 0.2 and 0.2 respectively. However, these differences were significant between Achievers and Underachievers ($p < 0.001$) only. This achievement received some criticism because some students felt pressured to talk in class against their will, even though this was an optional item that counted for extra credit only, and some people even resented others talking in class. The Proactive achievement yielded 1.1 badges to the average Achiever, while 0.5, 0.5 and 0.2 went to average Late Awakeners, Disheartened and Underachievers, respectively. This achievement was highly criticized for encouraging students to post as fast as they could in detriment of quality. The Archivist was the least popular of the newly introduced achievements, and only one Disheartened student participated. Learners considered this achievement to be poorly rewarded and thus unappealing.

Discussion

Our analysis suggests that our learning environment can cater to the different student groups observed in our course. This effect seems more predominant in the second year, with a new student cluster emerging, while we observed a significant reduction of the Underachiever population from 40 % down to 25 %. In this section we describe each cluster, analyze how they evolved from one year to the next, and discuss how the improvements made in the second year seemed to affect learner attitudes and behavior. We finalize by exploring how the results of this study could be used to further improve our gamified learning approach, making it smarter and more adaptive to students.

Student type overview

We found different student types, three of which appear recurrently in both years, and one arose in the second year only. We'll now describe what characterizes students of each type.

Achievers, disheartened and underachievers

The Achievers strived to be the best and participated on most aspects of the course. As a result, they hoarded many XP and were often found at the top of the leaderboard. These students presented the highest mean final grade (1st year = 90.16 %, 2nd year = 98.74 %) and acquired the most XP (1st year = 78.24 %, 2nd year = 97.07 %) and made the most posts on all evaluation components on either year, and they also considered the course to be more motivating than others. Achievers were the most participative and proactive and they were the only student type that seemingly felt that they were actually playing a game. This suggests that these students were highly engaged in the course.

Achievers are the student cluster whose behavior appears to remain more constant between years, with the most remarkable changes being a significant improvement both in terms of final grade, as well as the 100 % grade obtained from Quests and Challenges. We hypothesize that this happened because the game component was better rewarded and more appealing in the second year, and there were more challenges and achievements to attain. As a consequence, more opportunities to participate and acquire XP emerged, which may have positively impacted student engagement, as compared to the first year. This leads us to believe that the gamified learning experience was better adapted to these students' needs in the second year.

Underachievers were typically the worst performing group concerning almost every aspect of the course. They were characterized by having acquired the fewest XP (1st year = 49.63 %, 2nd year = 77.56 %) and occupying the lowest leaderboard ranks. They also had the lowest final grade (1st year = 71.88 %, 2nd year = 79.22 %) and earned the fewest XP and made the fewest posts on all evaluation components, which denotes a lack of participation. Underachievers did not particularly feel they were playing a game, although they found the course to be quite motivating and interesting, as compared to regular offerings. This suggests that these students were not very engaged by the course and that they probably saw it as just another hurdle they needed to clear. We believe that our experience did not satisfy many of these students' needs, which results in exclusion.

Underachiever performance saw little change from the first to the second year, the most notable being the resemblance of its XP accumulation curve to that of the other

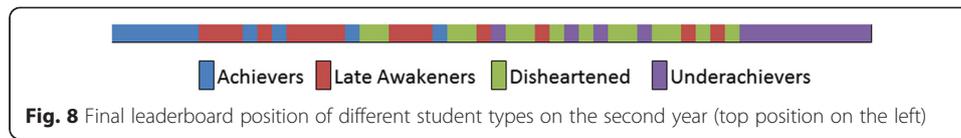
clusters during the first weeks of class, which suggests that they might have been more interested in the course during this period, as compared to the previous year. We hypothesize that these students may have been initially lured by the increased value of the game component.

Disheartened students performed similarly to Achievers during the first weeks on either year, but then fell into a tier of their own, with performance levels between those of Achievers and Underachievers. These students presented below average XP gains and participation levels, which led them to occupy the middle of the leaderboard. They had below average final grade (1st year = 77.82 %, 2nd year = 85.99 %) and amount of XP earned from achievements (1st year = 60.32 %, 2nd year = 77.56 %), and amount of XP and posts made on all other evaluation aspects. These students usually found the course to be both more motivating and engaging than other offerings, but did not really feel they were playing a game. This suggests that they were only moderately engaged with the course. In the second year, however, this cluster underwent a few changes. The initial period during which they performed closer to Achievers shrank from around 45 days in the first year, to only 25 in the second, meaning that their performance drop occurred earlier in the second year. Furthermore, they reported the lowest score regarding how interesting and motivating they found the course to be, which together with the previous finding suggests that Disheartened students were less engaged by the second gamified instance of the course. We hypothesize that the additional competition or the perceived heavier workload may have demotivated these students. Thus, our gamified approach seems not to be aligned with the needs of these students.

The emergence of late awakeners

The Late Awakeners, discovered in the second year, shared some traits with both Achievers and Disheartened students, presenting slightly above average performance and participation levels. Members of this cluster exhibited average XP accumulation curves, situated between those of Achievers and Disheartened students, and shared the middle of the leaderboard with the latter student type. Late Awakeners presented above average final grade (mean = 91.31 %) and amount of XP earned from achievements (mean = 85.68 %) and performed close to the Achievers, regarding the amount of both XP and posts made on almost every evaluation component, and they also participated the most on the Quests (mean = 4.33 posts). However, they performed much closer to the Disheartened students on the Theoretical Challenges (mean = 21.93 posts). Just like Disheartened students, Late Awakeners did not feel they were playing a game. However, they reported the course as being both more interesting and motivating than the other cluster.

Both Late Awakeners and Disheartened students displayed similar XP accumulation curves (see Fig. 6). However, looking closely at how their leaderboard rank evolved over time (see Fig. 7), we can see that the progression curve of Late Awakeners resembles the inverse of Disheartened students, suggesting that both student groups were directly competing for the same ranks. This finding is corroborated by their final position on the leaderboard. Both groups ended scattered throughout the middle of the leaderboard (see Fig. 8), even though Late Awakeners tended to occupy higher ranks. Between days 21 and 37 Late Awakeners started to grab more XP and Disheartened students seemed to slow down, which caused a swap in ranking between the two student types.



Understanding the emergence of this new cluster is not an easy task. The sheer existence of an additional 17 students in the second year might have potentiated the appearance of a fourth cluster. However, we believe that in the second year our course was better adapted to the needs of more students, which might have catalyzed the Late Awakeners cluster. We hypothesize that the additional challenges and their evenner distribution along the term, as well as the additional grade value of the game component, allowed them to do more of what they liked while creating further opportunities for them to recover lost XP. This might have engaged (or “recovered”) some of the students that would otherwise fall either into the Disheartened or the Underachiever clusters. We also hypothesize that the introduction of these new participation opportunities might be partly responsible for the changes observed on the Disheartened students’ behavior. We believe that many Late Awakeners came from Disheartened students that took advantage of the additional chances to turn the game around in the second year.

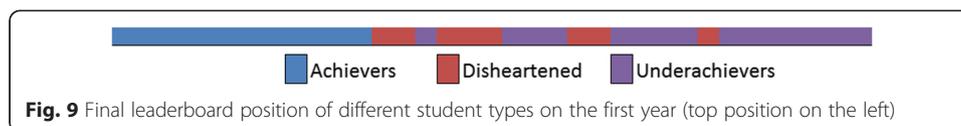
Implications to engagement

Overall, students seem to have enjoyed the gamified learning experience, with all student types reporting higher levels of interest and motivation with our course, as compared to other regular courses. There were a few changes in the second year, where we added five new achievements, with moderate success. While the Talkative badge encouraged students to participate, adding value to the learning experience, some students felt pressed to talk in class against their will, which made them feel penalized. We believe that providing alternative achievements to this one would mitigate its harmful effect. We already had a few achievements that provided bonus XP, which could be completed in alternative to mandatory achievements. The notion that students do have alternatives should be further emphasized throughout the term. On the other hand, the cooperative achievements proved a mild success. We believe that a new achievement requiring students to collaborate towards a common goal should be added instead of trying to get them to help each other completing lab work, as this turned out to be an unappealing task.

Cluster dynamics also suffered some changes. The XP accumulation curves of all student types were much closer apart during the first weeks of class on the second year, as compared to the first. We believe the new participation opportunities and the increased XP value ascribed to achievements made the experience more attractive to all student types, thus increasing their participation. This might have made some students within each cluster stand out from the crowd, increasing cluster overlap. This is supported by the final leaderboard disposition on the second year, which is less homogeneous than that of the first year (see Figs. 9 and 8).

Connection with existing research

It is possible to observe some level concordance between our findings and those of previous research. We observed a parallelism between our cluster structure and the model proposed



by Taylor (2002). Our Achievers resemble Taylor's Workers, characterized by high performance, participation and proactivity. Taylor's Shirkers are similar to our Underachievers, which are students with poor performance and proactivity levels, and with low participation as well. Taylor's Lurkers are similar to our Late Awakeners performance-wise. Both present an above average performance, close to the best performing student type (i.e. Taylor's Worker and our Achiever). Regarding participation metrics, our Disheartened students also post below average, just like Taylor's Lurkers.

Both Taylor's and our models diverge regarding the nature of the data used to split the students into different types. Taylor's model is focused mainly on how students perform on assignments and how much they post and visit the discussion boards. Our model distinguish students based on XP accrual, which encompasses temporal fluctuations on XP acquisition. Although there was indeed a direct matching between Taylor's Workers and Shirkers and our Achievers and Underachievers respectively, Taylor's Lurkers were covered by two of our clusters, the Late Awakeners and the Disheartened students. We have two hypotheses that may explain this divergence: 1) our model covers something that Taylor's did not, which are the temporal XP acquisition differences between students; or 2) our gamified learning experience provides several mechanisms to allow students to learn from trial and error-this gives them room to recover lost XP, which splits our middle cluster in those that take the most advantage of this an those that do not, the Late-Awakeners and the Disheartened. We lack data to support any of these hypotheses, and we believe that they are not necessarily mutually exclusive. Additional research is required to verify them and further compare both models.

Fleming and Mills (2001) proposed the VARK model, which encodes four different learning modalities: Visual, Aural, Read/Write and Kinesthetic. We consider that some of our performance and behavioral measures might portray a preference towards some of those modalities. For example, it is possible that students that went after the Talkative badge had a preference for the VARK's Aural modality. It is also tempting to believe that students performing better on the quizzes and the exam might have a preference for the Read/Write modality, whereas students that do better on the Lab assignments might prefer the Kinesthetic. However, we must not forget that students received XP for doing these activities, which clouds whether or not preference played a major role here. Therefore, it is hard to associate a specific cluster with any set of VARK modalities. Indeed, this is also an interesting topic for future research.

Although we lack the data to associate our clusters to specific player types, we can derive a few insights on this subject based on how our course was designed. A great part of the feedback we provide is focused on badge acquisition and achievement completion, which requires students to plan ahead to achieve these goals. Thus, we hypothesize that our course might appeal more to the BrainHex's Achiever and Mastermind types (Nacke et al., 2011). While the former enjoys long term goals and completing collections, the latter likes to make efficient decisions. Similarly, our experience should appeal more to Bartle's Achievers (Bartle, 1996; Bartle, 2004), who like to beat the game to collect points,

equipment objects and make a major effort to get rewards that will not necessarily benefit gameplay, simply for the prestige of having them. Our hypothesis is that these player types, the BrainHex's Achiever and Mastermind and the Bartle's Achiever, might be predominant in students that participate more and show greater signs of engagement, such as the Achievers and the Late Awakeners. We would like to further investigate the relationship between our cluster structure and different player classifications and gaming preferences in future research.

Towards adaptive gamified learning

We believe that gamified learning shows great potential to reach out to more students and to reduce exclusion. However, just like games, gamified approaches need to be fine-tuned and adequate rules must be created to captivate all students. In the second year, we observed a reduction in the percentage of Underachievers, in comparison to the first year, which dropped from 40 % to 25 %. This suggests that more students participated in and took advantage of this novel learning experience. We hypothesize this was a consequence of making the game component more rewarding while creating new opportunities for students to participate and recover lost XP later in the term.

We believe this study is a steppingstone to the creation of adaptive gamified learning environments. It revealed a large amount of student data that can be extracted within a gamified learning experience, which can be used to describe student behavior and performance, and characterize them into different categories. We believe this descriptive power has the potential to be converted into predictive power, by using the features that best characterize each student type to train a classifier and then predict what kind of behavior a student might develop, as early as possible.

Student classification has been addressed in previous research. A few studies aimed at classifying gifted students regarding their achievement and underachievement (Diener, 1960; Farquhar and Payne, 1964; McCoach and Siegle, 2003), or to distinguish between different learning styles (Fleming and Mills, 2001; Kolb et al., 1984). Machine-learning techniques have also been explored as a means to predict student performance and comprehension. Larkey (1998) trained Naïve Bayes (Hand and Yu, 2001) and k-nearest-neighbors classifiers "to assign scores to manually-graded essays". These scores were then combined with several text features using linear regression. The author reported that the generated models worked as well as if classified by human graders. Pattanasri et al. (2012) used Support Vector Machines (Cortes and Vapnik, 1995) to predict student comprehension of slides displayed in class, based on self-reported comprehension levels. On another work, Minaei-Bidgoli et al. (2003) used several algorithms to classify students using logged data in an online learning system and predict their final grade.

For future research we would like to make use of classification algorithms to predict a student's type as early as possible in the course, so that it could adapt in real time to students in need. This would be the most useful to detect Disheartened students and Underachievers as soon as possible and adapt content delivery to re-engage them with the course. For example, these students could be cheerfully reminded of easy tasks that they could undertake to accumulate XP and get back on track. However, before this can be achieved, two preparatory steps must be taken: 1) cluster consistency must be

verified, exploring if their existence and composition is replicable on upcoming years; 2) a thorough study has to be done to identify a consistent set of features that best describe each student type, as early as possible, which would be fed to the classification algorithm.

Study limitations

Our study has several limitations that need to be taken into account, most of which relate to the existence of uncontrolled variables. We could not control neither the number of students enrolled on both years nor the nature of the students' population. Furthermore, some of the course materials had to change in order to keep the course updated, according to the university's policy, and the composition of the faculty staff also differed on both years. These differences might have affected the students' engagement differently from one year to the next. A within-subjects study on a single year might help mitigate these limitations. However, a totally gamified course involves a different and demanding evaluation method, and having two groups evaluated according to different criteria would neither be fair nor ethical.

We also consider that differences in the number of students and the number of first days excluded from the cluster analysis may have influenced the nature of the identified clusters. While we could not control the first variable, we had some freedom to tweak the latter. We opted for excluding the first days where either there was no activity or some students were not yet playing the game. This would lead the clustering algorithm to consider that there were a lot more students with 0 XP for a longer time than there actually were, because these people were not playing the game yet. We assumed this compromise and we recognize that it might have affected the results.

We acknowledge the limitations of this study and consider that our results must be interpreted with caution. However, this exploratory research effort was valuable as it allowed us to identify different student types in a gamified learning experience, and understand how gamification and cluster analysis may work together to create adaptive learning environments.

Conclusion

In a previous experiment we have gamified a college course on Multimedia Content Production. This included adding game elements such as experience points and levels, badges and a leaderboard to the learning experience. Additionally, course activities were shaped into meaningful endeavors, such as achievements and time-limited tasks, namely Challenges and Online Quests. Results show that students participate more in activities with gamification and also report the course as being both more motivating and interesting than non-gamified courses.

In this paper we presented an exploratory study where we tried to understand how different students experienced our gamified learning experience. We analyzed how they acquired XP over the term, during two years of gamified delivery, and adopted a clustering algorithm to group them according to different XP acquisition patterns. We identified three clusters common to both years: 1) Achievers, which participated the most and strived to be the best, 2) Underachievers, which had the lowest performance and participation levels and seemed to lack interest for the course, and 3) Disheartened

students, whose initial performance was closer to that of Achievers but soon dropped to average levels. A fourth cluster, which we named Late Awakeners, emerged in the second year, where a few changes to improve game rewards and to create more participation opportunities took place. These students were characterized by an XP accumulation pattern opposite to that of Disheartened students: their performance dropped during the first weeks, but then recovered to levels closer to Achievers'.

Still, our learning experience has yet to appeal to all students. Indeed, Disheartened students and Underachievers seemingly are not taking as much advantage of the gamified experience as the others do, and might require specific help. We believe that our clustering mechanism based on student XP acquisition patterns can be used to train a classifier and develop an adaptive gamified learning experience, which might prove useful in improving performance by dynamically tailoring student learning experiences and courseware delivery in real-time.

Abbreviations

EM: Expectation-Maximization; ICT: Information and communications technology; IT: Information technology; MCP: Multimedia Content Production; MUD: Multi User Dungeon; XP: Experience points.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

GB conceived the study and performed statistical analysis. SG helped drafting the manuscript. JJ and DG coordinated and administered the study in class, and revised manuscript critically for important intellectual content. All authors read and approved the final manuscript.

Acknowledgement

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

Received: 8 August 2014 Accepted: 7 May 2015

Published online: 21 May 2015

References

- Badgeville (2015). <http://badgeville.com/>. (Accessed March 2015)
- G Barata, S Gama, J Jorge, D Gonçalves, Improving participation and learning with gamification, in *Proceedings of the First International Conference on Gameful Design, Research, and Applications* (Gamification '13, New York, NY, USA: ACM, 2013), pp. 10–17
- R Bartle, Hearts, clubs, diamonds, spades: players who suit muds. *J MUD Res* 1(1), 19 (1996)
- RA Bartle, *Designing virtual worlds* (New Riders, San Francisco, CA, US, 2004)
- C Bateman, R Boon, *21st Century Game Design (Game Development Series)* (Charles River Media, Inc., Rockland, MA, USA, 2005)
- C Bateman, LE Nacke, The neurobiology of play, in *Proceedings of the International Academic Conference on the Future of Game Design and Technology* (Futureplay '10, New York, NY, USA: ACM, 2010), pp. 1–8
- S Bennett, K Maton, L Kervin, The 'digital natives' debate: a critical review of the evidence. *Br J Educ Technol* 39(5), 775–786 (2008)
- P Brauner, A Calero Valdez, U Schroeder, M Ziefle, Increase physical fitness and create health awareness through exergames and gamification, in *Human Factors in Computing and Informatics, volume 7946 of Lecture Notes in Computer Science*, ed. by A Holzinger, M Ziefle, M Hitz, M Debevc (Springer, Berlin Heidelberg, 2013), pp. 349–362
- R Caillois, *Les jeux et les hommes: le masque et le vertige*, vol. 184 (Editions Gallimard, Paris, France, 1991)
- J Chen, Flow in games (and everything else). *Commun ACM* 50, 31–34 (2007)
- C Cheong, F Cheong, J Filippou, Quick quiz: a gamified approach for enhancing learning, in *Pacific Asia Conference on Information Systems* (AIS, Jeju Island, Korea, 2013), p. 206
- Codecademy (2015). <http://www.codecademy.com/>. (Accessed March 2015)
- B Collier, D Shernoff, Video game-based education in mechanical engineering: a look at student engagement. *Int J Eng Educ* 25(2), 308–317 (2009)
- C Cortes, V Vapnik, Support-vector networks. *Mach Learn* 20(3), 273–297 (1995)
- C Crawford, *Chris Crawford on game design* (New Riders, San Francisco, CA, US, 2003)
- C Crumlish, E Malone, *Designing social interfaces* (O'Reilly Media, Inc., Sebastopol, CA, US, 2009)
- M Csikszentmihalyi, *Flow: The psychology of optimal experience* (Harper Perennial, New York, NY, USA, 1991)
- M de Aguilera, A Mendiz, Video games and education: (education in the face of a "parallel school"). *Comput Entertainment* 1(1), 1:1–1:10 (2003)
- AP Dempster, NM Laird, DB Rubin, Maximum likelihood from incomplete data via the em algorithm. *J R Stat Soc Ser B Methodol* 39(1), 1–38 (1977)

- P Denny, The effect of virtual achievements on student engagement, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '13, New York, NY, USA: ACM, 2013), pp. 763–772
- S Deterding, D Dixon, R Khaled, L Nacke, From game design elements to gamefulness: defining "gamification", in *Proceedings of the 15th International Academic MindTrek Conference Envisioning Future Media Environments* (ACM, New York, NY, USA, 2011a), pp. 9–15
- S Deterding, M Sicart, L Nacke, K O'Hara, D Dixon, Gamification. using game-design elements in non-gaming contexts, in *Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems* (CHI EA '11, New York, NY, USA: ACM, 2011b), pp. 2425–2428
- CL Diener, Similarities and differences between over-achieving and under-achieving students. *Guid J* **38**(5), 396–400 (1960)
- A Domínguez, J Saenz-de Navarrete, L de Marcos, L Fernández-Sanz, C Pagés, J-J Martínez-Herráiz, Gamifying learning experiences: practical implications and outcomes. *Comput Educ* **63**, 380–392 (2013)
- T Dong, M Dontcheva, D Joseph, K Karahalios, M Newman, M Ackerman, Discovery-based games for learning software, in *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems* (CHI '12, New York, NY, USA: ACM, 2012), pp. 2083–2086
- CB Dowdall, N Colangelo, Underachieving gifted students: review and implications. *Gifted Child Quarterly* **26**(4), 179–184 (1982)
- M Ester, H-P Kriegel, J Sander, X Xu, A density-based algorithm for discovering clusters in large spatial databases with noise. *Kdd* **96**, 226–231 (1996)
- WW Farquhar, DA Payne, A classification and comparison of techniques used in selecting under-and over-achievers. *Personnel Guid J* **42**(9), 874–884 (1964)
- P Feyerabend, Wittgenstein's philosophical investigations. *Philos Rev* **64**(3), 449–483 (1955)
- DH Fisher, Knowledge acquisition via incremental conceptual clustering. *Mach Learn* **2**(2), 139–172 (1987)
- Z Fitz-Walter, P Wyeth, D Tjondronegoro, B Scott-Parker, Driven to drive: Designing gamification for a learner logbook smartphone application, in *Proceedings of the 2013 Symposium on Gameful Design, Research, and Applications* (Gamification 2013, Stratford, ON, Canada, 2013), pp. 42–49
- N Fleming, C Mills, *Vark: a guide to learning styles*, 2001. <http://vark-learn.com/>. (Accessed March 2015)
- JP Gee, What video games have to teach us about learning and literacy. *Comput Entertain* **1**(1), 20–20 (2003)
- L Hakulinen, T Auvinen, A Korhonen, Empirical study on the effect of achievement badges in traka2 online learning environment, in *Learning and Teaching in Computing and Engineering (LaTICE), 2013* (IEEE Computer Society, Washington, DC, USA, 2013), pp. 47–54
- DJ Hand, K Yu, Idiot's bayes—not so stupid after all? *Int Stat Rev* **69**(3), 385–398 (2001)
- O Inbar, N Tractinsky, O Tsimhoni, T Seder, Driving the scoreboard: Motivating eco-driving through in-car gaming, in *Proceedings of the CHI 2011 Workshop Gamification: Using Game Design Elements in Non-Game Contexts.: ACM, 2011*
- M Kebritchi, A Hirumi, H Bai, The effects of modern math computer games on learners' math achievement and math course motivation in a public high school setting. *Br J Educ Technol* **38**(2), 49–259 (2008)
- Khan Academy (2015). <https://www.khanacademy.org/>. (Accessed March 2015)
- AJ Kim, *Putting the fun in functional*, 2008. <http://www.slideshare.net/amyjokim/putting-the-fun-in-functiona>
- DA Kolb, *Experiential learning: Experience as the source of learning and development*, vol. 1 (Prentice-Hall, Englewood Cliffs, NJ, 1984)
- LS Larkey, Automatic essay grading using text categorization techniques, in *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval* (ACM, New York, NY, USA, 1998), pp. 90–95
- J Lee, K Luchini, B Michael, C Norris, E Soloway, More than just fun and games: assessing the value of educational video games in the classroom, in *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (CHI EA '04, New York, NY, USA: ACM, 2004), pp. 1375–1378
- W Li, T Grossman, G Fitzmaurice, Gamicad: a gamified tutorial system for first time autocad users, in *Proceedings of the 25th annual ACM symposium on user interface software and technology* (UIST '12, New York, NY, USA: ACM, 2012), pp. 103–112
- J MacQueen, Some methods for classification and analysis of multivariate observations, in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, volume 1* (University of California Press, California, USA, 1967), pp. 281–297
- P McClean, B Saini-eidukat, D Schwert, B Slator, A White, *Virtual worlds in large enrollment science classes significantly improve authentic learning*. In *Proceedings of the 12th International Conference on College Teaching and Learning*, Center for the Advancement of Teaching and Learning (Jacksonville, FL, US, 2001), pp. 111–118
- DB McCoach, D Siegle, Factors that differentiate underachieving gifted students from high-achieving gifted students. *Gifted Child Quarterly* **47**(2), 144–154 (2003)
- J McGonigal, *Reality is broken: Why games make us better and how they can change the world* (Penguin Publishing Group, New York, NY, US, 2011)
- Merriam-Webster.com, "game", 2015. Accessed March 2015
- DR Michael, SL Chen, *Serious Games: Games That Educate, Train, and Inform* (Thomson Course Technology, Boston, MA, US, 2006)
- Microsoft, *Microsoft ribbon hero*, 2015. <http://www.ribbonhero.com/>. (Accessed March 2015)
- B Minaei-Bidgoli, DA Kashy, G Kortmeyer, WF Punch, Predicting student performance: an application of data mining methods with an educational web-based system, in *Frontiers in Education, 2003. FIE 2003 33rd Annual, volume 1* (IEEE Computer Society, Los Alamitos, CA, USA, 2003), pp. T2A–13
- Moodle, Moodle, 2015. <http://www.moodle.org/>. (Accessed March 2015)
- SM Moon, AS Hall, Family therapy with intellectually and creatively gifted children. *J Marital Fam Ther* **24**(1), 59–80 (1998)
- J Moreno, Digital competition game to improve programming skills. *Educ Technol Soc* **15**(3), 288–297 (2012)
- Mozilla, *Open badges*, 2015. <http://openbadges.org/>. (Accessed March 2015)
- IB Myers, MH McCaulley, *Manual: A guide to the development and use of the MyersBriggs type indicator* (Consulting Psychologists Press, Palo Alto, CA, US, 1988)
- L Nacke, C Bateman, R Mandryk, Brainhex: Preliminary results from a neurobiological gamer typology survey, in *Proceedings of 10th International Conference on Entertainment Computing (ICEC'11)* (Springer, Vancouver, BC, 2011), pp. 288–293

- HF O'Neil, R Wainess, EL Baker, Classification of learning outcomes: evidence from the computer games literature. *Curric J* **16**(4), 455–474 (2005)
- N Pattanasri, M Mukunoki, M Minoh, Learning to estimate slide comprehension in classrooms with support vector machines. *Learn Technol, IEEE Trans* **5**(1), 52–61 (2012)
- M Prensky, Digital natives, digital immigrants part 1. *Horizon* **9**(5), 1–6 (2001)
- B Reeves, J Read, *Total Engagement: How Games and Virtual Worlds Are Changing the Way People Work and Businesses Compete* (Harvard Business Press, Boston, MA, US, 2009)
- N Sharma, A Bajpai, MR Litoriya, Comparison the various clustering algorithms of weka tools. *Facilities* **4**, 7 (2012)
- L Sheldon, *The Multiplayer Classroom: Designing Coursework as a Game*, 2011. Course Technology PTR, 1 edition
- S Sheth, J Bell, G Kaiser, Halo (highly addictive, socially optimized) software engineering, in *Proceeding of the 1st international workshop on Games and software engineering, volume 11 of GAS* (ACM, New York, NY, USA, 2011), pp. 29–32
- B Shneiderman, Designing for fun: how can we design user interfaces to be more fun? *Interactions* **11**(5), 48–50 (2004)
- KD Squire, Video games in education. *Int J Intell Games Simul* **2**(1), 49–62 (2003)
- K Squire, M Barnett, JM Grant, T Higginbotham, Electromagnetism supercharged!: learning physics with digital simulation games, in *Proceedings of the 6th international conference on Learning sciences, ICLS '04* (International Society of the Learning Sciences, Los Angeles, CA, US, 2004), pp. 513–520
- JC Taylor, Teaching and learning online: the workers, the lurkers and the shirkers. *J Chin Distance Educ* **9**, 31–37 (2002)
- Weka (2015). <http://www.cs.waikato.ac.nz/ml/weka/>. (Accessed March 2015).
- JR Whitmore, *Giftedness, conflict, and underachievement* (Allyn and Bacon, Boston, 1980)
- G Zichermann, C Cunningham, *Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps* (O'Reilly Media, Inc., Sebastopol, CA, US, 2011)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
