

Exploring the Synergies between Gamification and Data Collection in Higher Education

Martín Liz-Domínguez ¹, Manuel Caeiro-Rodríguez ¹, Martín Llamas-Nistal ¹ and Fernando Mikic-Fonte ¹

¹ *AtlantTic Research Center, University of Vigo, Spain*

Abstract

In recent years, gamification techniques have been gaining popularity in all kind of educational scenarios, helping students improve their learning process by fostering engagement and attention. Implementing gamification aspects in a course can also provide an opportunity to gather student data that would not have been available otherwise. This paper describes a data gathering process in the context of a university course, as a work-in-progress. Among these data there is information regarding the participation of students in quizzes presented as games in the classroom. These quizzes combined questions covering course contents, as well as some regarding self-regulated learning habits. The main advantage observed was a high student participation in the quizzes. As a result, this gamification approach proved to be a more effective way to gather student data compared to other methods applied in previous academic years, which often failed due to many students ignoring optional activities.

Keywords

Data collection, gamification, learning analytics, self-regulated learning.

1. Introduction

In recent years, the influence of modern digital media has led to the rise of new technological trends, subsequently originating transformations in many fields of human activity. The concept of “gamification” is one of such trends. Originating in the late 2000s and gaining mainstream popularity in the decade of 2010, gamification is often defined as “the use of game design elements in non-game contexts” [1]. Since its inception, gamification aspects have been applied in many different kinds of contexts. For example, nowadays there are a multitude of fitness applications that use game elements to engage the user in the practice of physical activity [2]. In healthcare, there exist gamified applications and serious games for rehabilitation and management of chronic diseases [3].

One of the fields with most prominent implementations of gamification elements is education. Researchers in this field have devised many different approaches to applying gamified elements to educational contexts, with the ultimate goal of encouraging and engaging students in their learning process by making it more interesting and enjoyable.

Some examples of game mechanics that have been applied in educational contexts throughout the years are leaderboards [4], achievements [5] or even role-playing game elements such as levels and experience points [6]. These studies often find that the introduction of gamification in education succeeds in increasing student motivation, even ultimately leading to an overall improvement in grades. However, while gamified mechanics are usually well received by most students, some of them may feel uncomfortable with some of the added competitive aspects [5].

The application of gamification in educational context has often gone hand in hand with learning analytics, another recent trend enabled by the progressive digitalization of education. The participation of students in gamified course activities is a source of data that can be processed for visualization

Learning Analytics Summer Institute Spain (LASI Spain) 2022, June 20–21, 2022, Salamanca, Spain
EMAIL: mliz@gist.uvigo.es (A. 1); mcaeiro@gist.uvigo.es (A. 2); martin@gist.uvigo.es (A. 3); mikic@gist.uvigo.es (A. 4)
© 2022 Copyright for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

purposes [7] or combined with other educational data sources to build more complex learning analytics applications, such as those involving predictive analysis [8].

This paper focuses on how integrating gamification elements in a university course affects the nature of data that is collected for the purpose of future analysis. The experiment targets a first-year engineering course which, in previous academic years, has served as the target context for data collection and analysis tasks, aiming to detect students at risk of failing and assess proficiency in several self-regulated learning (SRL) aspects. During the academic year 2021/2022, gamified activities were introduced in the course, and alongside them, new types of data that can be obtained and transformations to some of the data that could already be acquired in previous years. It is important to note that the target course is still underway as of the time of writing this article, which makes this a work-in-progress.

The rest of this document is structured as follows: Section 2 describes the analyzed learning scenario in the 2020/2021 academic year, and the types of data that were collected back then. Section 3 summarized how the course was updated going into the 2021/2022 academic year, highlighting the new types of data introduced because of this. Section 4 explains the advantages and disadvantages of the new scenario compared to the previous one in terms of data collection for analysis. Finally, Section 5 provides some closing thoughts and details the expected future work to be performed in this line of research.

2. Old educational scenario

The course that serves as the base for the data collection and analysis tasks described in this paper is a first-year Computer Architecture course at University of Vigo's Telecommunications Engineering school. This course implements a blended learning methodology in the form of the flipped classroom system: lectures are provided online to students in the form of videos to watch at home. Additionally, the option of following a continuous assessment system is provided, in which students' grades are determined by the performance in several short exams scheduled throughout the duration of the course [9].

During the 2020/2021 academic year, there were 212 students enrolled in the subject, and 123 of them followed the continuous assessment system. Students were split into six groups, each one of which having a weekly one-hour-long in-classroom session for problem solving and performing exams.

This course makes use of two different online tools to manage all its resources: Moovi, which is University of Vigo's Moodle-based institutional LMS, and BeA, an e-assessment platform for the digital management of traditional on-paper exams [9]. They are used for the following specific purposes:

- Moovi is used as a hub for students to access educational resources, which include video lectures, course notes and self-assessment tests. These resources are progressively uploaded and made available to students as the course progresses. This platform is also used as a communication channel, implementing forums and a board of announcements in which teachers can post messages directed to all students.
- BeA handles every aspect related to designing, assessing and reviewing exams. Additionally, it has a feature that allows instructors to create non-summative questionnaires or surveys for students to complete within the platform.

The combined use of both platforms grants the opportunity to collect various types of educational data. Moovi — and, generally speaking, Moodle — provides the possibility of exporting activity logs, which provide information regarding the times at which students log into the platform and the sequence of resources that they access while logged in. On the other hand, BeA provides rich assessment information, including exam grades, the types of mistakes that students make in specific exercises and communications between students and instructors during the review process.

Additionally, as a way to collect complementary data to estimate students' self-regulated learning proficiency, several SRL questionnaires were created and made available for students to complete at different stages of the course [10]. Specifically, a 20-item questionnaire was performed on paper at the beginning of the course during an in-classroom session, and three 7-item questionnaires were uploaded to BeA throughout the course, roughly one every four weeks. Participation in these questionnaires was not mandatory, and their results had no impact on the course grade.

The main problem observed regarding these questionnaires was low student participation. While the initial questionnaire, performed in the classroom, was completed by 113 students, the three subsequent questionnaires on BeA were only submitted by 22, 22 and 17 students respectively. This prompted us to re-design the way in which these regularly scheduled questionnaires were presented to students moving onto the following academic year.

3. New educational scenario

During the 2021/2022 iteration of the Computer Architecture course — which, as aforementioned, is not over as of the date of writing this document —, there were a total of 199 enrolled students, 120 of which following the continuous assessment option. Moovi is being used for the exact same purposes as in the previous academic year. However, a new BeA feature is being tested in the course: the possibility of performing games (quizzes) during in-classroom sessions [11].

Figure 1 shows a simple representation of how running a quiz on BeA works. This feature is similar in concept and functionality to that of Kahoot! (<https://kahoot.com>), a popular gamification tool used in many kinds of courses and at all academic levels [12]. These kinds of programs are rooted in the concept of audience response systems (ARS), allowing live interaction between a speaker and an audience by providing each person with a device to answer polls or questions.

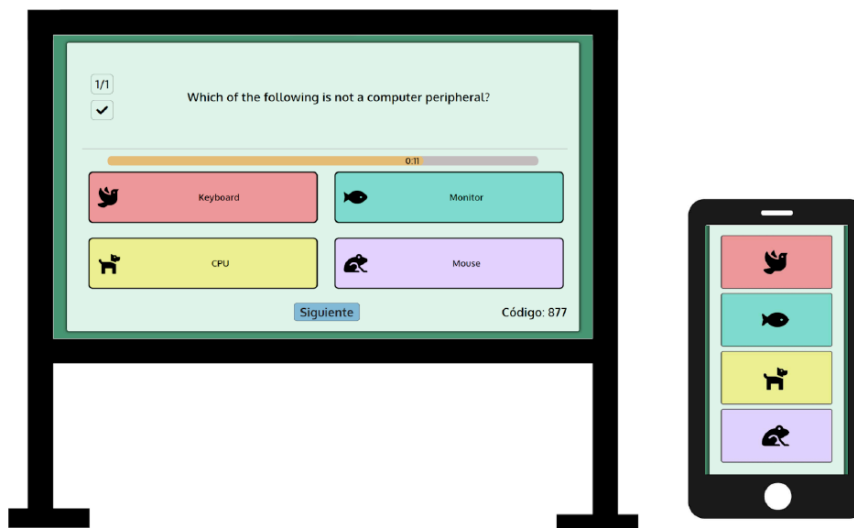


Figure 1: Performing a quiz in the classroom with BeA.

In BeA quizzes, the teacher prepares a set of multiple-choice questions before the in-classroom session. Whenever a quiz is being started in the classroom, a numeric code is shown, allowing attending students to join in using their own computers or mobile devices. During the game, questions are shown in succession, ideally on a big screen in the classroom, and students earn points for correct answers depending on how long they take to respond, with bonus points for streaks of consecutive correct answers. At the end of the game, a leaderboard is shown highlighting the highest-scoring students.

BeA quizzes can be seen as similar to self-assessment tests: both consist of sequences of questions that serve as formative activities for students, helping them review and apply course contents. However, they have different use cases in this course: BeA quizzes include short questions meant to be answered in just a few seconds, while self-assessment tests usually include longer, exam-like problems to be solved in a more extended period of time. Regarding the data collection aspect, these two types of activities are also similar, generating data regarding student performance, in terms of wrong and right answers. Additionally, in the case of BeA quizzes, it is possible to monitor exactly how long students take to answer each question.

The introduction of BeA quizzes also provided another way to ask SRL-related questions to students. This was done by mixing some SRL questionnaire items together with course-related questions in each quiz, although designed in such a way so that answers to SRL questions do not affect students' score in

the quiz. This way, student participation in the SRL quizzes is tied to class attendance, resulting in more students answering them, as will be detailed in Section 4. However, a drawback of this method is that it may be harder to fit as many SRL-related questions in a quiz compared to a traditional questionnaire, because teachers often cannot afford to spend too much time in an in-classroom session with activities not directly related to course contents. In this case, just 2 SRL questionnaire items were included in each quiz, compared to the 7 that were present in each periodic SRL questionnaire during the 2020/2021 academic year.

In summary, BeA quizzes took over the role of the regularly scheduled SRL questionnaires from the previous academic years. The only SRL questionnaire that was still performed was the 20-item one at the beginning of the course, in the same conditions: students answered it on paper during an in-classroom session. The items included in this questionnaire were, however, different from the ones present in the initial questionnaire from the 2020/2021 academic year.

4. Results: impact on data collection

As mentioned in the previous sections, the biggest change in the data collection process in the context of this course is caused by the introduction of BeA quizzes, which in turn changed the way in which SRL-related questions were asked to students. Table 1 summarizes the outcomes of 2020/2021 SRL questionnaires compared to 2021/2022 BeA quizzes in terms of student participation. It is important to note that for the 2021/2022 academic year, only data for the three first quizzes were available at the time of writing this document, with at least one more remaining scheduled to be performed during the final weeks of the course.

The first questionnaire listed in Table 1 is the longer, 20-item one that was carried out in a similar way in both academic years: on paper, during an in-classroom session in the second week of the course. These registered high participation levels in both iterations, as class attendance is at its highest point during the first few weeks of the course.

In subsequent questionnaires, substantial differences are observed in terms of student participation. In 2020/2021, questionnaires 2, 3 and 4 registered between 17 and 22 responses, which represents a drop in participation of about 80% compared to the first questionnaire. On the other hand, participation in BeA quizzes during the 2021/2022 academic year matched class attendance, as these are activities performed in the classroom. This is the critical aspect that explains the improvement in participation compared to the online questionnaires. Even if there is a noticeable decline in participation as the course advances, from 60 students in quiz 2 to 37 in quiz 4, these are still much more representative samples of students compared to what was achieved in the previous year.

Table 1
Student participation in 2020/2021 SRL questionnaires and 2021/2022 BeA quizzes.

Questionnaire number	2020/2021			2021/2022		
	Week nr.	Students	Cron α	Week nr.	Students	Cron α
1 (20-item)	Week 2	113	0.7017	Week 2	100	0.5951
2	Week 5	22	0.7701	Week 4	60	0.6069
3	Week 9	22	0.7835	Week 7	57	0.6299
4	Week 13	17	0.8237	Week 9	37	0.6465

As explained in Section 3, the main drawback of performing SRL questions as part of quizzes is the reduction in the number of questions that can be asked in each one of them: from 7 in the 2020/2021 questionnaires, to 2 in the 2021/2022 quizzes. This is somewhat compensated by the fact that BeA quizzes are being scheduled twice as frequently as SRL questionnaires were: from one every four weeks in 2020/2021, to biweekly quizzes in 2021/2022. As aforementioned, more BeA quizzes will be performed until the end of the course.

Also listed in Table 1 are the Cronbach's alpha values, a popular internal consistency metric that estimates the extent to which items in a questionnaire actually measure the same concept, and are thus related to each other [13]. In this study, Cronbach's alpha values for each questionnaire were calculated

in a cumulative way, considering all previously performed questionnaires as well. For example, when calculating Cronbach's alpha for the third questionnaire in the 2021/2022 academic year, the answers for items in the first and second questionnaires of that same academic year were also included, as if they were all part of a single instrument. Regarding the treatment of missing values, pairwise deletion was applied, which just discards missing answers on a question-by-question level, as opposed to disregarding the whole questionnaire from a student if even a single answer is missing [14].

It can be seen that Cronbach's alpha values for the 2020/2021 academic year start at a value of 0.7, while in 2021/2022 start at 0.6. The former is considered an acceptable internal consistency value, while the latter may imply the existence of some flaws in the design of the questionnaire. This difference between both years exists mostly due to the fact that questionnaire items were changed going from one academic year to the next. It can also be observed that Cronbach's alpha values increase as more questionnaires are performed — although this may be partly explained by the use of pairwise deletion, which tends to produce small overestimations in the metric [14]. This is particularly noticeable in the case of the 2020/2021 academic year, in which missing values are widely present due to poor student participation, tending to reduce answer variance, and consequently increasing the obtained alpha value.

There are no significant changes in the way other types of data are collected going from one academic year to another. Moovi activity logs are obtained and interpreted the same way, as are other BeA data types, such as exam grades and student errors.

5. Conclusion and future work

Overall, this paper highlights how gamification has transformed several aspects in academic contexts that implement it, not just in terms of how course activities are designed and carried out, but also in the types of student data that can be collected. In the case of our studied Computer Architecture course, the introduction of quizzes that are playable in the classroom provided an opportunity to include SRL-related questions as part of these activities. As a result, we were able to collect a significantly higher amount of answers to these questions compared to the previous academic year, in which optional SRL questionnaires were mostly ignored by students. The main drawback of this system is asking fewer questions in total throughout the course.

The data collection process described in this paper is part of a bigger data analytics project focused on predictive analysis. Collected data from both studied academic years are being analyzed with the goal of identifying patterns that define students at risk of failing the course. Ultimately, the main idea is building an early warning system (EWS) and applying it in future academic years, which teachers could use to help them focus efforts in helping struggling students. We consider SRL data to be a significantly relevant factor in this endeavor, as proficiency in SRL skills is directly related to student performance. Thus, once the 2021/2022 academic year comes to an end, we will be compiling all the data that we can obtain in the course and use it as input for a machine learning model, aiming to predict from these data which students fail the course and which ones succeed.

6. Acknowledgements

This work is partially financed by public funds granted by the Galician regional government, with the purpose of supporting research activities carried out by PhD students. (“Programa de axudas ‘a etapa predoutoral da Xunta de Galicia — Consellería de Educación, Universidade e Formación Profesional”.) Additionally, this work has received financial support from the Xunta de Galicia (Centro singular de investigación de Galicia accreditation 2019-2022) and the European Union (European Regional Development Fund - ERDF), and from the Galician Regional Government under project ED431B 2020/33.

7. References

- [1] Deterding, S., Dixon, D., Khaled, R., Nacke, L.: From game design elements to gamefulness: defining "gamification". In: Proceedings of the 15th International Academic MindTrek Conference on Envisioning Future Media Environments - MindTrek '11. p. 9. ACM Press, New York, New York, USA (2011), <http://dl.acm.org/citation.cfm?doid=2181037.2181040>
- [2] Lister, C., West, J.H., Cannon, B., Sax, T., Brodegard, D.: Just a Fad? Gamification in Health and Fitness Apps. *JMIR Serious Games* 2(2), e9 (aug 2014), <http://games.jmir.org/2014/2/e9/>
- [3] Sardi, L., Idri, A., Fernández-Alemán, J.L.: A systematic review of gamification in e-Health. *Journal of Biomedical Informatics* 71, 31–48 (jul 2017), <http://www.ncbi.nlm.nih.gov/pubmed/28536062>
- [4] Ortiz-Rojas, M., Chiluíza, K., Valcke, M.: Gamification through leaderboards: An empirical study in engineering education. *Computer Applications in Engineering Education* 27(4), 777–788 (jul 2019), <https://onlinelibrary.wiley.com/doi/10.1002/cae.12116>
- [5] Domínguez, A., Saenz-de Navarrete, J., De-Marcos, L., Fernández-Sanz, L., Pagés, C., Martínez-Herráiz, J.J.: Gamifying learning experiences: Practical implications and outcomes. *Computers & Education* 63, 380–392 (apr 2013), <https://linkinghub.elsevier.com/retrieve/pii/S0360131513000031>
- [6] Stansbury, J.A., Earnest, D.R.: Meaningful Gamification in an Industrial/Organizational Psychology Course. *Teaching of Psychology* 44(1), 38–45 (jan 2017), <http://journals.sagepub.com/doi/10.1177/0098628316677645>
- [7] Cassano, F., Piccinno, A., Roselli, T., Rossano, V.: Gamification and Learning Analytics to Improve Engagement in University Courses. In: *Advances in Intelligent Systems and Computing*, vol. 804, pp. 156–163. Springer, Cham (2019), http://link.springer.com/10.1007/978-3-319-98872-6_19
- [8] Pastushenko, O., Oliveira, W., Isotani, S., Hruška, T.: A Methodology for Multimodal Learning Analytics and Flow Experience Identification within Gamified Assignments. In: *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. pp. 1–9. ACM, New York, NY, USA (apr 2020), <https://dl.acm.org/doi/10.1145/3334480.3383060>
- [9] Llamas-Nistal, M., Mikic-Fonte, F.A., Caeiro-Rodríguez, M., Liz-Domínguez, M.: Supporting Intensive Continuous Assessment With BeA in a Flipped Classroom Experience. *IEEE Access* 7, 150022–150036 (2019), <https://ieeexplore.ieee.org/document/8865067/>
- [10] Liz-Domínguez, M., Caeiro-Rodríguez, M., Llamas-Nistal, M., Mikic-Fonte, F.: Monitoring students' self-regulation as a basis for an early warning system. In: *Learning Analytics Summer Institute Spain 2021: Learning analytics in times of COVID-19: Opportunity from crisis*. vol. 3029, pp. 38–51. CEUR-WS, Barcelona, Spain (2021), <http://ceur-ws.org/Vol-3029/paper04.pdf>
- [11] Mikic-Fonte, F., Llamas-Nistal, M., Caeiro-Rodríguez, M., Liz-Domínguez, M.: A Gamification Module for BeA Platform. In: *2020 IEEE Frontiers in Education Conference (FIE)*. vol. 2020-Octob, pp. 1–5. IEEE (oct 2020), <https://ieeexplore.ieee.org/document/9274180/>
- [12] Wang, A.I., Tahir, R.: The effect of using Kahoot! for learning – A literature review. *Computers & Education* 149, 103818 (may 2020), <https://linkinghub.elsevier.com/retrieve/pii/S0360131520300208>
- [13] Tavakol, M., Dennick, R.: Making sense of Cronbach's alpha. *International Journal of Medical Education* 2, 53–55 (jun 2011), <http://www.ncbi.nlm.nih.gov/pubmed/28029643>
- [14] Béland, S., Pichette, F., Jolani, S.: Impact on Cronbach's alpha of simple treatment methods for missing data. *The Quantitative Methods for Psychology* 12