

To Reward and Beyond: Analyzing the Effect of Reward-Based Strategies in a MOOC

Alejandro Ortega-Arranz^{a,*}, Miguel L. Bote-Lorenzo^a, Juan I. Asensio-Pérez^a, Alejandra Martínez-Monés^b, Eduardo Gómez-Sánchez^a, Yannis Dimitriadis^a

^a*School of Telecomm. Engineering, Universidad de Valladolid, p.º de Belén 15, 47011 Valladolid, Spain*

^b*School of Computer Engineering, Universidad de Valladolid, p.º de Belén 15, 47011 Valladolid, Spain*

Abstract

Despite the benefits of MOOCs (*e.g.*, open access to education offered by prestigious universities), the low level of student engagement remains as an important issue causing massive dropouts in such courses. The use of reward-based gamification strategies is one approach to promote student engagement and prevent dropout. However, there is a lack of solid empirical studies analyzing the effects of rewards in MOOC environments. This paper reports a between-subjects design study conducted in a MOOC to analyze the effects of badges and redeemable rewards on student retention and engagement. Results show that the implemented reward strategies had not significant effect on student retention and behavioral engagement measured through the number of pageviews, task submissions, and student activity time. However, it was found that learners able to earn badges and redeemable rewards participated more in gamified tasks than those learners in the control group. Additionally, results reveal that the participants in the redeemable reward condition requested and earned earlier the rewards than those participants in the badge condition. The potential implications of these findings in the instructional design of future gamified MOOCs are also discussed.

Keywords: Gamification, MOOC, Engagement, Retention, Between-Subjects Design

1. Introduction

The importance and presence of MOOCs in digital education has been growing up since their appearance in 2008 (Siemens, 2013), providing multiple benefits such as access to high quality open learning, accreditation from prestigious universities, or the creation of communities around a shared topic (Siemens, 2013; Ferguson & Sharples, 2014; Deng et al., 2019). The increasing number of courses and enrollments (Shah, 2017) indicates the growing interest of society in MOOCs. Nevertheless, MOOCs still fail to motivate and engage

*Corresponding author

Email addresses: alex@gsic.uva.es (Alejandro Ortega-Arranz), migbot@tel.uva.es (Miguel L. Bote-Lorenzo), juaase@tel.uva.es (Juan I. Asensio-Pérez), amartine@infor.uva.es (Alejandra Martínez-Monés), edugom@tel.uva.es (Eduardo Gómez-Sánchez), yannis@tel.uva.es (Yannis Dimitriadis)

learners with course contents and learning activities (Khalil & Ebner, 2014) leading to low participation and high dropout rates (Jordan, 2014; Alario-Hoyos et al., 2014).

10 Some dropouts can derive from the different profiles enrolling in MOOCs (Kizilcec et al., 2013; Ferguson & Clow, 2015). However, many dropouts are produced as a side effect of the pedagogical models and instructional designs used (Margaryan et al., 2015; Henderikx et al., 2017). Some research works have shown the benefits of strategies promoting active learning to improve engagement (Ferguson & Sharples, 2014; Hew, 2016), thus trying to decrease the
15 dropout rates and the low levels of participation. Gamification is one of these strategies that has attracted the attention of MOOC practitioners during the last years (Davis et al., 2018), due to the benefits already shown in other educational contexts (De Sousa Borges et al., 2014; Dicheva et al., 2015).

Gamification is the application of elements and structures that frequently appear in
20 games (*e.g.*, narrative, rewards) into non-game contexts (Deterding et al., 2011; De Sousa Borges et al., 2014) such as those involving online learning. According to previous literature reviews (Hamari et al., 2014; Dicheva et al., 2015; Ortega-Arranz et al., 2017), one of the most used game elements in online educational contexts are rewards, thus generating the so-called *reward-based gamification* (Nicholson, 2015) or *incentive systems* (Kyewski & Krämer, 2018).
25 In this type of gamification, students are awarded with *game elements* (*rewards*) integrating a *signifier* (*e.g.*, name, visual, description) when a *completion logic* (*i.e.*, relevant actions defined by the teachers beforehand) is satisfied (Hamari & Eranti, 2011; Hamari, 2017).

Previous studies integrating reward-based gamifications in online and blended learning courses have shown positive benefits on student retention (Khalil et al., 2017; Krause et al.,
30 2015) and engagement (O'Donovan et al., 2013; Ding et al., 2017; Ibáñez et al., 2014; Barata et al., 2013; Anderson et al., 2014). However, none of such studies was performed in real MOOC environments with a significantly heterogeneous set of participants, and therefore extrapolating their results to MOOCs is venturesome. Additionally, most of these studies analyze the effects of several game elements as a whole instead of isolating the effects of
35 specific reward types, and do not compare the effects of different types of rewards on learners. Finally, previous literature reviews on gamification in MOOCs reflect the lack of empirical evidence regarding their effects (Ortega-Arranz et al., 2017; Antonaci et al., 2017; Khalil et al., 2018).

Evidence-informed design could be greatly useful for instructional designers and MOOC
40 instructors to better understand the effects of concrete rewards on student engagement and retention, and to better align their MOOC learning designs with the gamification goals. Moreover, researchers can advance in the understanding of gamification benefits for MOOC learners. Thus, the generic underlying research question that leads this study is *To what extent reward-based gamification strategies can foster student retention and engagement in*
45 *MOOCs?*

To address this question, we have conducted a between-subjects study (Charness et al.,
2012) considering three different conditions to which participants were randomly assigned (two experimental conditions, each with a different type of reward, and one control group), in an 8-week MOOC with 866 enrolled students. In this study, students belonging to ex-
50 perimental conditions had to explicitly claim the rewards to earn them once they satisfied

the gamified tasks. Therefore, we could better analyze the effects and motivation behind rewards and avoid influencing those students who were not interested in them.

The structure of the paper is as follows. Section 2 provides a brief overview of similar research works studying with the effects of rewards in MOOCs and formulates the hypotheses of this study. Section 3 explains the design of the study including the sample, the context, the gamification design and the data sources. Then, the results are presented (Section 4) and the findings are discussed regarding the implications of rewards in the instructional design of gamified MOOCs (Section 5). Finally, conclusions, limitations and ideas for future research are introduced in the last section.

2. Reward-Based Strategies in MOOCs

This section first describes previous studies investigating the effects of reward-based strategies on student engagement and retention. Then, the theoretical background supporting the formulation of the hypotheses addressed in this study is introduced.

2.1. Related Work

There exist several previous studies dealing with the effect of gamification in online and blended learning environments (Dicheva et al., 2015; Khalil et al., 2018). However, to the best of our knowledge, only five of these studies have been performed in a MOOC-like context¹ and isolate the effect of a single type of reward (typically, badges) instead of studying the effect of multiple game elements as a whole. These similar works are described below.

Cross et al. (2014) implemented a series of badges associated with a variety of activities in two MOOCs. Badges were requested by learners and manually issued by instructors and other peers. Although results showed positive opinions and attitudes about badges, the study focused on understanding the reasons behind badge acquisition rather than in the effects on engagement and retention.

Ortega-Arranz et al. (2019) carried out a mixed method analysis to understand the correlation between learners' actions towards earning badges and their behavioral engagement in a MOOC. Results revealed a positive correlation between such parameters but causality was not analyzed and the effects of badges on student retention were not examined.

Anderson et al. (2014) performed a quantitative analysis of the effects of badges applied to discussion forums in a MOOC with 112.897 enrolled students. Although results showed that forum participation was higher than in a previous run of the same MOOC, and that the participation increased when the next available badges were clearly visible, no further analysis was performed on student retention and engagement.

Khalil et al. (2017) studied the effects of implementing a meter in form of a battery bar whose charge increases when predefined conditions are met (*e.g.*, attempt a quiz, reading

¹MOOCs present different features with respect to traditional environments which can have an impact in the expected outcomes caused by gamification strategies (*e.g.*, learning design activities, massive and heterogeneous set of participants, lack of time to complete the course) (Ortega-Arranz et al., 2019).

and posting in forums) in a MOOC with 284 enrolled students. The study presented a higher percentage of active and certified students compared with two previous runs of the same MOOC without gamification. Nevertheless, the game element used was a ‘battery bar’ (non-collectable reward) and the analysis was restricted to the student retention level without isolating the effect caused by the context (*i.e.*, different versions of the course).

Similarly, [Kyewski & Krämer \(2018\)](#) divided course participants into three different conditions regarding the visibility of course badges to understand the effects of badges on learners’ motivation, performance (grades) and participation (logins, quiz participations and attempts, resource access). No matter the condition, student’s motivation decreased over time, and learners in the experimental conditions did not participate more actively in the course nor got higher scores than those in the control group (no badges). However, the study was not performed in a real MOOC context. Instead, the context of the study was an online one-semester seminar in a higher education setting.

In contrast with previous research works, the between-subjects study presented in this paper analyzes and compares the effects of two reward types (*i.e.*, badges and redeemable rewards) on student retention and engagement. Additionally, this study was carried out in a real MOOC environment with an heterogeneous set of participants (*e.g.*, age, location), where learners had to explicitly claim the rewards, which provided more variables to complement the analysis of the effects on students’ engagement in the experimental conditions (*e.g.*, the number of students claiming rewards, the claiming time stamp).

2.2. Theoretical Background

Although there is some controversy about the effects of rewards in educational environments (*e.g.*, [Deci et al. \(2001\)](#); [Seaborn & Fels \(2015\)](#)), researchers have identified several reasons behind the interest of achieving rewards in educational learning environments such as intrinsic and extrinsic motivation ([Nicholson, 2015](#)), sense of progression and goal accomplishment ([Hamari, 2017](#)) or simply fun ([Codish & Ravid, 2014](#)). Reward strategies in online environments are frequently implemented through different game elements such as points, badges or levels ([Dicheva et al., 2015](#); [Ortega-Arranz et al., 2017](#)) aiming to increase the student motivation and engagement and to mitigate the high dropout rates. To this end, some authors have proposed the use of specific game elements depending on the *human desires* that the designers want to promote ([Zichermann & Cunningham, 2011](#); [Bunchball, 2010](#)). [Chang & Wei \(2016\)](#) identified badges and redeemable rewards as two of the most engaging game elements to be used in MOOC environments.

Badges are optional rewards, represented with graphical icons and issued when users satisfy predefined requirements typically associated with non-compulsory activities ([Domínguez et al., 2013](#); [Hamari, 2017](#)). Redeemable Rewards are rewards which provide students with a certain privilege during course runtime (*e.g.*, extra attempts in quizzes, access to extra content, join a queue to receive feedback from teachers) ([O’Donovan et al., 2013](#); [Ortega-Arranz et al., 2018](#)). These two reward types have been also positively evaluated by students in other MOOC studies ([Cross et al., 2014](#); [Rizzardini et al., 2016](#)). In this study, we will focus on these two types of rewards to understand and compare their effects on student retention and engagement.

2.2.1. Rewards and Retention in MOOCs

130 Several authors have identified factors that are important to help increase student retention in MOOCs (Adamopoulos, 2013; Hone & El Said, 2016), among which gamification is considered as a potential strategy to this end. For example, Rizzardini et al. (2016) proposed the addition of gamification elements in a conceptual model for MOOCs to decrease their attrition rates. Similarly, Borrás-Gené et al. (2016) created a gamified cooperative MOOC
135 model for the design of engineering education MOOCs to improve the student motivation, learning level and completion rate.

Previous studies have shown that reward strategies can serve students as elements of progression and goal accomplishment, encouraging them to keep track of their learning and performance (Hamari, 2017). Empirically, Krause et al. (2015) performed a between-
140 subjects design study where experimental conditions implementing reward-based strategies in a SPOC reduced the dropout rate in a 25%. Also, the study performed by Khalil et al. (2017) showed that the attrition rates were much lower in a MOOC with a reward element than in two previous versions of the same MOOC without gamification.

Therefore, according to the previous studies and assuming students will try to earn the
145 rewards associated with course activities as a sign of progression and goal accomplishment, the first hypothesis for this study is:

- **H1:** *A higher retention level will occur for participants under the Experimental conditions.*

2.2.2. Rewards and Engagement in MOOCs

150 One of the indirect causes of student retention is the student engagement with the course content and activities (Khalil & Ebner, 2014). Several authors have proposed the inclusion of gamification strategies in online environments to promote student engagement, mainly based on two theories: the Flow Theory (Csikszentmihalyi, 1991) and the Self-Determination Theory (Ryan & Deci, 2000). Adapting from the former theory, when teachers design gamified
155 activities that challenge students and keep them in the *flow*² zone, students will maintain their engagement throughout the course (e.g., Zhu et al. (2017); Antonaci et al. (2018)). The Self-Determination Theory proposes the satisfaction of three human psychological needs to promote student motivation: competence, relatedness and autonomy. Satisfying these three needs with the gamification elements and their associated tasks is likely to increase student
160 engagement in the course activities (e.g., Borrás-Gené et al. (2014); Seaborn & Fels (2015)).

Therefore, if reward-based strategies are aligned with the activity difficulty level and target the previous psychological needs, they can be potentially used as motivators to engage students in MOOCs. According to Fredricks et al. (2004), there are three types of engagement: cognitive, emotional and behavioral. In this study, we will focus on behavioral
165 engagement as *the observable behaviors that represent the student progress and learning in the course*. Previous studies have empirically tested the positive effects of reward strategies

²Flow is defined as a state of absorption characterized by intense concentration, loss of self-awareness, a feeling of being perfectly challenged and a sense that time is flying (Csikszentmihalyi, 1991).

on behavioral engagement in online and blended educational environments (Barata et al., 2013; Ibáñez et al., 2014; Ding et al., 2017; Ruipérez-Valiente et al., 2017; Ortega-Arranz et al., 2019). Therefore, the second hypothesis of this study is:

- 170 • **H2:** *Participants in the Experimental conditions will show a higher behavioral engagement level than participants in the Control group.*

In the current context, the use of reward-based strategies provides researchers with more variables defining the student behavioral engagement within the course. Some of these additional variables are the completion of the task associated with rewards, the time needed to complete the task or the time needed to claim the reward. This additional set of measures of behavioral engagement corresponds to the so-called *reward-derived student engagement*. Some previous studies support the use of rewards as a mean to promote student specific actions (*i.e.*, actions associated with reward conditions) (Anderson et al., 2014; Hakulinen et al., 2013; Hamari, 2017). Learners interested in earning rewards would be expected to show higher engagement through the early completion of the reward conditions. Therefore, according to previous studies, the third and fourth hypothesis of this study are:

- 175 • **H3:** *A higher percentage of students in the Experimental conditions will perform the instructor predefined actions associated with reward conditions, as compared to the Control group.*
- 180 • **H4:** *Students in the Experimental conditions will satisfy the gamified conditions sooner than the students in the Control group.*

Another variable representing the student engagement when rewards are implemented, is the time elapsed from the moment that a student satisfies the reward conditions to the moment s/he claims and receives such reward. This variable can potentially inform us about the student motivation and interest on earning such reward in a massive online learning environment (Berger et al., 2016). The evidence that a student claims the reward right after completing the conditions suggests that s/he was aware of the existence of this reward and wanted to earn it. In this regard, redeemable rewards have been ranked in the 2nd position as the most engaging game element in MOOCs over badges and trophies in the 5th position (Chang & Wei, 2016). Additionally, although redeemable rewards can be collected after the expiration date of the associated privilege, they are expected to be earned and used before (*e.g.*, students are expected to earn a privilege associated to the second week of the course before the end of such week). We believe that students will try to earn redeemable rewards sooner than badges due to expiration of the associated privileges in the different course modules. Therefore, the fifth hypothesis of this study is:

- 185 • **H5:** *Students in the Redeemable Rewards condition will request the rewards sooner than students in the Badge condition.*

Condition	Enrolled Students	Participants of the Study
BADGE	290	223
REDEEM	287	205
CTRL	289	220
Total	866	648

Table 1: Number of MOOC enrolled students and participants of the study per condition.

3. The Study

3.1. Sample

205 The students of the course were randomly assigned to one of the following conditions once they enrolled in the course:

- **BADGE:** Students involved in this condition were able to obtain up to 8 badges throughout the course.
- 210 • **REDEEM:** Students involved in this condition were able to obtain up to 8 redeemable rewards whose requirements were the same as the badges.
- **CTRL:** Students involved in this condition had neither rewards nor game elements implemented in the course. This condition was considered as the control group of the study.

215 This group assignment by enrollment date avoided bias caused by those students registering late who are more prone to disengage with the course (Gurantz, 2015). Among 866 learners enrolled in the course³, 648 submitted the initial questionnaire, allowing them to access the course contents and activities. Additionally, by submitting the initial questionnaire, learners provided their consent to analyze their data with research purposes, thus allowing us to profile the participants of this study (see Table 1).

220 According to the data reported in the initial questionnaire, the students of this study were mostly female (83.02%), between 20-30 years old (64.17%), from Latin America (53.86%), with an undergraduate background (56.17%), and medium knowledge level about the topic of the MOOC (39.81%), planning to actively participate in the course (58.64%), without previous MOOC and gamification experience (69.60% and 60.49% respectively), and with positive 225 beliefs about the benefits of using gamification in educational environments (64.20%).

Before testing the hypotheses of this study, we checked the homogeneity of the three conditions regarding the variables of the initial questionnaire including the student gender, age, background knowledge level and type of participation in the course. If groups are homogeneously distributed, the differences in the composition of the groups are not likely to 230 influence the results of the study. To this end, we have conducted a *Chi-square homogeneity*

Variable	<i>p-value</i>	Exclusions (number of answers)
Age	0.954	DK/NA (2): R (1), C (1)
Gender	0.914	DK/NA (3): B (1), R (1), C (1)
Background	0.121	DK/NA (2): R (1), C (1)
Location	0.875	Asia (2), DK/NA (2): R (2), C (2)
Knowledge level	0.531	None (13), DK/NA (2): B (6), R (3), C (6)
Participant type	0.836	-
MOOC experience	0.928	DK/NA (2): R (1), C (1)
Gamification experience	0.573	-
Gamification beliefs	0.249	-

Table 2: Chi-square test for homogeneity *p-values* regarding the variables of the initial questionnaire (DK/NA=Don't Know/No Answer; B=BADGE, R=REDEEM, C=CTRL). *P-value* is significant at <.05 level (two-tailed).

test (Kirch, 2008) for every variable considered in the initial questionnaire⁴.

According to the results (see Table 2), the *p-values* are much higher than the significance level (.05) for every variable. Therefore, the degree of similarity among groups regarding the variables measured in the initial questionnaire is high and the results obtained in this study are unlikely to be caused by the composition of the groups.

3.2. Course Context

The study was conducted within an 8-week instructor-led MOOC offered by a Spanish university from March, 12th to May, 6th, 2018. The course was published and launched in the Canvas Network platform⁵. The topic of the course was related to translation from English to Spanish in the business and economic field, and the content was divided into 7 weekly modules plus one extra week to complete the activities. The modules included videos, content pages, recommended readings, discussion forums and individual and collaborative activities (see Fig. 1). The activities were classified into compulsory and optional. Students had to submit all the compulsory activities in order to receive the course completion certificate. For all activities, the submission was due eight days after the release of the activity although in some activities the due date was extended a few days according to the instructor criteria. The course team was composed of one instructor and two teacher assistants. Furthermore, the enrollment was closed in the second week of the course to avoid group management problems in the collaborative activities. Teaching and technical support was offered by the course team and the researchers respectively, through private messages and posts in forums.

³A simple random process was applied to assign participants to the different conditions but the number of enrolled students in every condition is different due to the removal of test users and duplicated accounts.

⁴The Chi-homogeneity test assumes a minimum number of frequencies of every multiple choice option (freq.>5). In our case, some questions were answered with a frequency under this value. As a consequence, these values have been removed or grouped as *other answer* representing a maximum number of 15 excluded answers among the three groups (*i.e.*, 2.31% from the total number of answers) as described in Table 2.

⁵Canvas Network: <https://www.canvas.net>, last access: December, 2018.

Week 0	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8
Course Info			Module Video-Introduction					Certificate Req.
Platform Info			Discussion Forum					Goodbye Forum
Twitter Account			Video Contents + Recommended Readings					
Facebook Page			Module Video-Summary				Self-Review: Text Translation	Course Experience Questionnaire
Social Forum	Content Questionnaire	Content Questionnaire	Text Analysis	Term Extraction (individual)	Text Translation	Text Selection (individual)	Content Questionnaire	
Resource Forum			Parallel Text Search	Term Extraction (group)	Peer Review: Text Analysis	Text Selection (group)	Peer Review: Text Translation	
Welcoming Questionnaire		Glossary		Text Translation	Peer Review: Text Translation	Text Translation	Peer Review: Text Translation	

Compulsory Activities Resources and Optional Activities

Figure 1: Learning design of the MOOC under study.

Group BADGE








						Group REDEEM
Week Cond.	Image	Name	Condition	Privilege	Week Priv.	
0		<i>Welcome!</i>	Update your profile picture and introduce yourself in the Social Forum	Get 3 more attempts in Quiz 1 and Quiz 2	1, 2	
1, 2		<i>Quiz Master!</i>	Get a score, equal or higher than 90% in Quiz 1 and Quiz 2	Get access to extra content in week 2	2	
2		<i>Glossary Master!</i>	Contribute with at least 3 terms in the Glossary activity	Extend the due date of the compulsory task at week 3	3	
3		<i>Text Provider!</i>	Share a text in the Parallel Text Search activity and receive 5 likes from other participants	Extend the due date of the text translation task at week 5	5	
5		<i>Expert Reviewer!</i>	Review 2 more submissions from your colleagues (4 in total) in week 5	Join the queue so that the instructors evaluate your work and provide feedback	5	
5		<i>Smartie!</i>	Get a score, equal or higher than 70% in the reviews performed by other peers regarding your submission	Get 20 more minutes in Quiz 7	7	
4, 6		<i>Translation Master!</i>	Submit the optional translations: Public Descriptive text and Expositive Private text	Get 3 more attempts in Quiz 7	7	
6		<i>Graduated!</i>	Watch the "summary videos" in weeks 1 to 6	Get access to an exclusive video-session with the teacher and other students	7	

Figure 2: Gamification design implemented in the MOOC of the study.

3.3. Gamification Design

The gamification was co-designed (Penuel et al., 2007) between the instructor of the course and the researchers. Since completion of compulsory assignments is a prerequisite for the course completion certificate, the research team wanted to avoid the motivational effect of getting this certificate on the fulfilment of reward conditions. Thus, the main instructor and the research team decided to gamify only optional tasks that the main instructor considered beneficial for learning, and whose fulfilment was not directly associated with the course completion certificate. Also, the teacher agreed on the course privileges available for the REDEEM Group according to the capabilities of the MOOC platform. The resulting gamification design (including both the badges and redeemable rewards decisions) is presented in Figure 2.

3.4. Gamification Implementation

The gamification was digitally represented and implemented with an extended version of the GamiTool portal⁶. GamiTool is a web-based system developed by the authors that allows instructional designers and teachers to design, and semi-automatically deploy and enact reward-based gamifications in multiple MOOC platforms such as Canvas Network (Ortega-Arranz et al., 2018). GamiTool allows to include a “Rewards” tab in the course interface where students could check the reward conditions and request the badges and redeemable rewards once the conditions were fulfilled. The IMS LTI compliant interface⁷ allows Canvas Network to provide GamiTool with student information to distinguish the learner condition and to display a different interface for each group. Figure 3 shows the different learner interfaces for the REDEEM and CTRL groups. Thus, the course content (pages, forum posts, etc.) was common for all participants, and only the gamification tab was different for learners belonging to different groups. Additionally, the system allowed automatic handling of the student requests for rewards, checking whether the conditions were satisfied and issuing the badges and the course privileges.

In this study, participants had to explicitly claim the rewards, which served a two-fold purpose. First, this claiming procedure allowed to distinguish which learners were really interested in the rewards and provided extra variables to measure their behavior towards gamification (e.g., variables measuring whether the student claimed the rewards right after satisfying the conditions). Second, learners who were not interested in nor attracted by rewards can avoid being bothered with these extra elements. Furthermore, rewards were only visible for each specific learner (i.e., students could not share with other students the rewards earned) to isolate the possible effects caused by social comparison (Festinger, 1954).

Finally, by midweek a message was sent to the groups, encouraging learners to perform the tasks associated with rewards and to claim them. In order to ensure that all the participants were prompted to perform the same actions, the control group was also advised to perform such optional actions with a similar message (without mentioning course rewards).

⁶GamiTool: <https://gamitool.gsic.uva.es>, last access: December, 2018.

⁷Learning Tools Interoperability: <https://www.imsglobal.org/activity/learning-tools-interoperability>, last access: December, 2018.

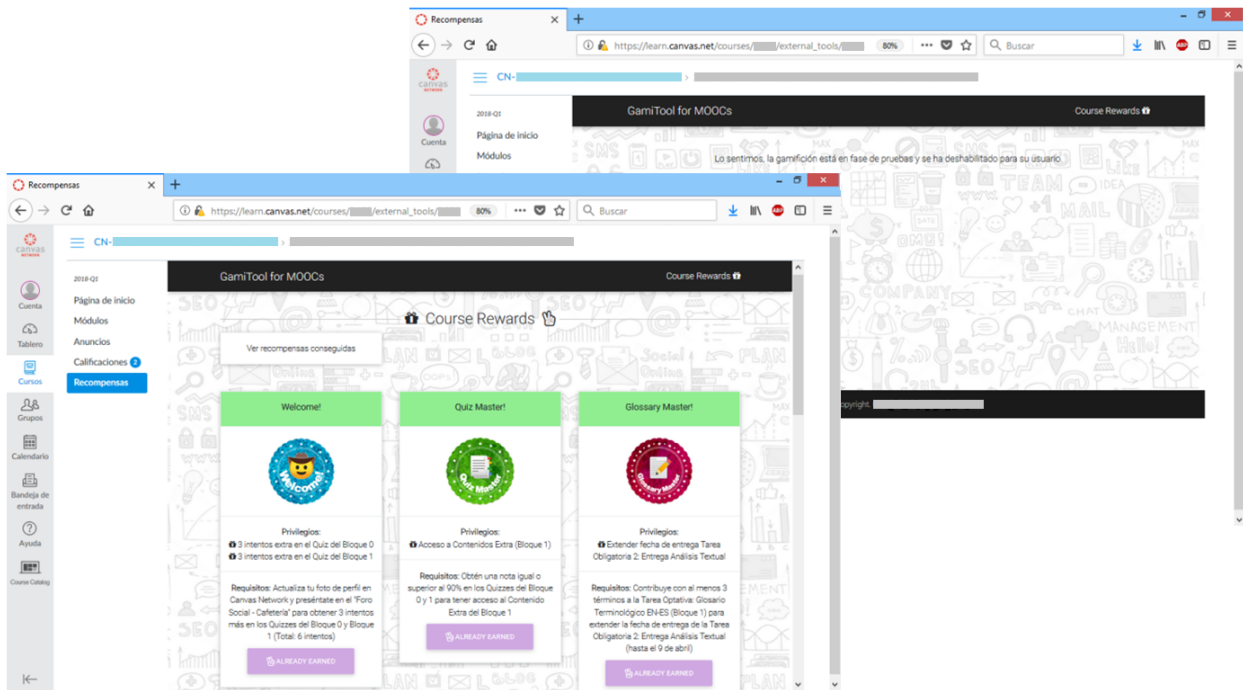


Figure 3: “Rewards” tab for the students belonging to the REDEEM (left) and CTRL (right) groups.

290 3.5. Data Sources and Processing

Quantitative data was collected from the Canvas Network and GamiTool platforms as described in Table 3. The student retention (Hypothesis H1) was measured through three different variables: (i) the number of learners that submitted the compulsory weekly task in the intermediate and last weeks of the course, (ii) the number of learners that visited at least one page in the intermediate and last weeks of the course, and (iii) the number of learners that obtained the course certificate (Khalil et al., 2017). A *Chi-square goodness of fit* test (Navidi, 2008) was performed to identify significant differences among conditions in the proportion of learners satisfying the previous variables in relation with the number of learners that submitted the initial questionnaire.

300 The general behavioral engagement (H2) was calculated through four variables typically used to this end in the literature (Henrie et al., 2015; Ferguson & Clow, 2015): the number of pageviews, the number of submitted tasks, the number of forum posts and the total activity time registered at the end of the course. Pairwise *z-tests* were calculated to analyze the mean differences of the previous variables across conditions due to the large sample sizes (Navidi, 2008).

305 The reward-derived student engagement (H3, H4, H5) was analyzed considering the percentage of active learners satisfying the conditions associated with rewards, the percentage of learners claiming the rewards, and the timestamps⁸ of both actions. In this context,

⁸In order to perform the statistical analysis, timestamp information was first retrieved from the different data sources, homogenized to the same time zone format (CET), and converted to *Unix* timestamp format.

Source Code	Description
Canvas_Log	Registry of learner actions performed in the MOOC platform. This registry includes information such as the number of pageviews, the tasks submitted and the number of forum posts.
Canvas_Pre_Quest	Answers provided by students in the initial questionnaire regarding the age, location, background, previous experience with MOOCs and with gamification, etc.
GamiTool_Log	Registry of student actions performed in the gamification platform such as the number of rewards issued and the timestamps.

Table 3: Description of the Data Sources used in this study.

Hypothesis	Variable	Source	Test
Group Homogeneity	Initial questionnaire variables	Canvas_ Pre_Quest	Chi-Square Test of Homogeneity
H1 Retention	Percentage of learners visiting the course, submitting compulsory tasks and obtaining the course certificate	Canvas_Log	Chi-Square Goodness of Fit Test
H2 Behavioral engagement	Pageviews, compulsory submissions, forum posts, activity time	Canvas_Log	Z-Test
H3 Fulfillment of conditions	Percentage of learners satisfying the reward conditions	Canvas_Log	Wilcoxon Signed-Rank Test
H4 Fulfillment of conditions	Timestamps when conditions associated to rewards are satisfied	Canvas_Log	Wilcoxon Signed-Rank Test
H5 Request of rewards	Timestamps when rewards are claimed and issued	GamiTool_Log	Wilcoxon Signed-Rank Test

Table 4: Summary of the data analysis.

310 *Wilcoxon signed-rank tests* (Navidi, 2008) were used to compare the differences among the three conditions. This test requires that differences between paired samples should be continuous and distributed symmetrically around the median. The presence of outliers on either side makes it reasonable to assume that the population is approximately symmetric for the three variables (Navidi, 2008). Additionally, we have calculated the effect size to isolate the mean difference sizes (Cohen, 1988).

315 In every hypothesis testing, the reasons for using an appropriate test (parametric and non-parametric) were based on the purpose of the analysis (*e.g.*, differences in population mean ranks), the data type (*e.g.*, numerical, categorical), the sample size and the assumptions that can be made on the data distribution of probability (*e.g.*, the application of the Central Limit Theorem). All the previous data (see Table 4) was gathered together in a
320 Microsoft Excel⁹ worksheet and processed with the RStudio software¹⁰.

4. Results

4.1. Student Retention (H1)

Hypothesis 1 states that *student retention would increase in the experimental conditions*. To examine this hypothesis, we have measured, for the BADGE, REDEEM and CTRL

⁹Microsoft Excel: <https://products.office.com/>, last access: January, 2019.

¹⁰RStudio: <https://www.rstudio.com/>, last access: January, 2019.

	BADGE	REDEEM	CTRL	<i>p-value</i>
Initial Questionnaire	223 (34.41%)	205 (31.64%)	220 (33.95%)	
Visited interm.	119 (33.71%)	112 (31.73%)	122 (34.56%)	0.956
Visited last	88 (36.06%)	68 (27.87%)	88 (36.06%)	0.446
Submitted interm.	67 (34.71%)	59 (30.57%)	67 (34.71%)	0.947
Submitted last	67 (36.02%)	55 (29.57%)	64 (34.41%)	0.817
Certified	63 (37.06%)	49 (28.82%)	58 (34.12%)	0.679

Table 5: Descriptive statistics of variables measuring student retention including the *p-value* Chi-square goodness of fit test (*p-value* is significant at the $<.05$ level).

325 groups, the proportion of (a) certified learners (*certified*), (b) learners submitting the in-
intermediate (*submitted interm.*) and last (*submitted last*) compulsory assignment, and (c)
learners visiting the course in the intermediate (*visited interm.*) and last week (*visited last*¹¹)
of the course. In all three cases, the variables are measured in relation with the learners
that submitted the initial questionnaire (see Table 5).

330 Results obtained after performing the statistical test (see Table 5) show that the *p-value* is
much higher than the *alpha* level (.05) for every variable measuring student retention among
the three different conditions. The results suggest that the evolution of active participants
throughout the mid and last week of the course and the proportion of students obtaining
the course certificate is similar for the three conditions of the study. Therefore, data does
335 not support Hypothesis 1.

4.2. Student Behavioral Engagement (H2)

Hypothesis 2 states that *student behavioral engagement would be higher in the Experi-*
mental conditions. To test this hypothesis, we analyzed (i) the number of pageviews, (ii) the
tasks completed (optional and compulsory), (iii) the number of posts in discussion forums
340 (entries and replies), and (iv) the total web browser activity time (in minutes) in the course.
In order to avoid bias caused by students that enrolled in the course without intention to
finish it, two clusters of students were considered: (a) students who submitted the initial
questionnaire excluding *Samplers* and *Strong Starters*¹² (Kizilcec et al., 2013; Ferguson &
Clow, 2015), and (b) students who were active until the end of the course, considering only
345 *Late* and *Keen Completers*¹³ (Kizilcec et al., 2013; Ferguson & Clow, 2015). Results are
presented in Table 6 for both clusters (see left and right columns respectively).

With all the data gathered, we conducted pairwise z-tests (two-tailed, *alpha* = .05)
between the three conditions according to the four variables described before (for the two

¹¹Since the last week of the course was intended to allow students complete remaining activities, this parameter considers those students visiting the course in at least one of the last two weeks of the course.

¹²*Samplers* are those learners who visit course content during a very small number of weeks, in this case, the first two weeks. *Strong Starters* are those learners that completed the first compulsory activity but then dropped out the course.

¹³*Late Completers* are those learners who completed the last compulsory task and submitted most of other assessments, but were either late or missed some out. *Keen Completers* are learners who completed all the compulsory activities, engaging actively throughout the whole course.

	C1: Excluded Samplers and Strong Starters			C2: Late and Keen Completers		
	BADGE (N=127)	REDEEM (N=112)	CTRL (N=123)	BADGE (N=66)	REDEEM (N=54)	CTRL (N=60)
Pageviews						
Median	388	310	299	574.5	590	561.5
Mean	439.4	406.3	386.5	614.7	660.6	615.5
Std	375.91	345.92	320.39	240.12	342.88	316.09
Task Submissions						
Median	10	6	8	12	12	12
Mean	7.945	7.33	7.602	12.35	11.98	12.17
Std	4.80	4.71	4.68	1.36	1.64	1.22
Forum Posts						
Median	1	1	1	1	1	1
Mean	1.803	1.821	1.22	2.727	2.796	1.717
Std	4.24	2.93	1.81	5.67	3.90	2.19
Activity Time*						
Median	731.4	710.59	810.32	1101.3	1306.5	1085.1
Mean	1134.4	1019.93	1255.28	1545.1	1465.2	1775.6
Std	1242.65	888.93	1624.26	1375.93	954.97	1968.55

Table 6: Median, mean and standard deviation values for variables measuring student behavioral engagement regarding the first (left) and the second (right) clusters. *Participants unregistering before the course end were not considered (15, 14, 14 participants in C1 from BADGE, REDEEM and CTRL respectively, and 1 participant in C2 from the BADGE condition).

	C1: Excluded Samplers and Strong Starters			C2: Late and Keen Completers		
	BADGE VS. CTRL	REDEEM VS. CTRL	BADGE VS. REDEEM	BADGE VS. CTRL	REDEEM VS. CTRL	BADGE VS. REDEEM
Pageviews	0.230	0.649	0.479	0.986	0.467	0.405
Task Submissions	0.567	0.658	0.318	0.430	0.499	0.189
Forum Posts	0.155	0.061	0.969	0.18	0.072	0.937
Activity Time	0.535	0.190	0.439	0.451	0.277	0.710

Table 7: Z-test *p-values* of variables measuring student behavioral engagement between conditions for the first (left) and second (right) clusters. *P-value* is significant at the $<.05$ level, two-tailed.

clusters). Results (see Table 7) revealed that there is a tendency towards significance in the number of forum posts between the REDEEM and the CTRL group in both clusters (on average, 0.6 and 1 forum posts higher in C1 and C2 respectively). However, the *p-values* are higher than the threshold level for every variable measuring student engagement in both clusters. Therefore, Hypothesis 2 is not supported.

4.3. Reward-derived Student Engagement: Percentage of Completion (H3)

Hypothesis 3 predicts that *A higher percentage of students in the Experimental conditions will complete the instructor predefined actions associated with reward conditions, as compared to the Control group.* The list of percentage completion for each reward is presented in Table 8.

	BADGE			REDEEM			CTRL		
	N	%	Date (μ)	N	%	Date (μ)	N	%	Date (μ)
<i>Welcome</i>	72	36.73	17/03 01:07:51	70	36.65	15/03 00:38:35	41	20.30	15/03 03:54:19
<i>Quiz</i>	51	30.18	25/03 04:01:14	40	27.78	24/03 23:52:21	35	22.01	24/03 18:39:56
<i>Glossary</i>	28	16.57	26/03 13:48:40	39	27.08	25/03 23:17:24	17	10.69	24/03 21:43:41
<i>Text Prov.*</i>	17	12.32	03/04 05:31:49	22	18.49	03/04 12:43:41	17	12.88	02/04 14:42:26
<i>Reviewer</i>	53	57.61	13/04 00:54:20	45	52.33	13/04 19:12:00	43	43	13/04 09:29:18
<i>Smartie*</i>	69	75	12/04 18:49:49	56	65.12	12/04 21:11:28	67	67	12/04 08:29:51
<i>Translation</i>	12	15.58	23/04 13:38:08	7	9.72	23/04 10:27:10	6	7.69	21/04 20:14:47
<i>Graduated</i>	35	45.45	21/04 17:21:38	30	41.67	22/04 10:39:41	22	28.20	18/04 18:22:48

Table 8: Statistical summary of participants satisfying the reward conditions (number and percentage of participants in relation with the number of active students per week, and mean value for the date and time of completion). *Rewards whose conditions' satisfaction depended on course peers instead on the own learner.

After performing a Wilcoxon signed-rank test, the R -value obtained when comparing the BADGE and REDEEM group with the CTRL group is 1 for both groups, which is under the critical value = 3 (two-tailed, $\alpha=.05$, $n=8$) (Navidi, 2008). We can conclude that the median weights of the percentage of students satisfying the gamified-task conditions in the BADGE and REDEEM groups (36.18% and 34.85% respectively) are significantly different from the median weight in the Control group (26.47%) with a p -value of 0.016. Additionally, the effect size is approximately 0.586 for the BADGE and REDEEM conditions, which is very large according to Cohen's classification of effect sizes for *behavioral sciences* (Cohen, 1988). Results also showed no significant differences between the BADGE and REDEEM groups (p -value = 0.641).

According to the results, both experimental groups present significant differences with the CTRL group regarding the percentage of students satisfying the gamified-task conditions (9.71% and 8.38% difference for the BADGE and REDEEM groups respectively). Therefore, Hypothesis 3 is supported and we can conclude that a significantly higher percentage of students in the experimental groups satisfied the gamified conditions compared with the CTRL group.

4.4. Reward-derived Student Engagement: Completion Dates (H_4)

Hypothesis 4 poses that *students in the Experimental conditions will satisfy the gamified conditions sooner than the students in the Control group*. The list of completion dates is presented in Table 8.

	BADGE			REDEEM		
	<i>N</i>	%	Date (μ)	<i>N</i>	%	Date (μ)
<i>Welcome</i>	56	77.78	23/03 22:08:27	56	80	20/03 03:01:47
<i>Quiz</i>	33	64.71	31/03 00:49:54	31	77.5	27/03 20:13:57
<i>Glossary</i>	25	89.29	27/03 15:47:03	32	82.05	26/03 23:54:53
<i>Text Prov.</i>	14	82.35	07/04 11:15:09	19	86.36	07/04 15:38:37
<i>Reviewer</i>	33	62.26	17/04 12:21:20	23	51.11	15/04 20:07:43
<i>Smartie</i>	41	59.42	20/04 19:51:15	29	51.79	17/04 11:57:25
<i>Translator</i>	10	83.33	25/04 18:09:09	5	71.43	23/04 22:11:50
<i>Graduated</i>	29	82.86	25/04 11:55:54	19	63.33	23/04 04:30:04

Table 9: Statistical summary of participants claiming and earning course rewards.

After performing a Wilcoxon signed-rank test, the R -value obtained when comparing the BADGE and REDEEM group with the CTRL group is 1 in both cases, which is over the critical value = 0 (two-tailed, $\alpha=.05$, $n=6^{14}$) (Navidi, 2008). Therefore, we can conclude that the median weights of the dates when students satisfy the gamified-task conditions in the BADGE and REDEEM groups are close to be significantly different from the median weight in the CTRL group (1 day, 9 hours and 4 minutes, and 1 day, 3 hours and 17 minutes later, respectively) with a p -value of 0.0625 for both groups. Furthermore, there are not significant differences between the median weights of the dates that students satisfied the reward conditions in the BADGE and REDEEM groups (p -value = 0.844).

Surprisingly, results suggest that students in CTRL group performed the tasks associated to rewards earlier than the students in the experimental groups. However, this difference is not significant and therefore, Hypothesis 4 is not supported.

4.5. Reward-derived Student Engagement: Claiming Dates (H5)

Hypothesis 5 states that *students in the Redeemable Rewards condition will request the rewards sooner than students in the Badge group*. The statistical summary of claiming dates is presented in Table 9.

After performing a Wilcoxon signed-rank test, the R -value obtained when comparing the BADGE with the REDEEM group is 1, which is under the critical value = 3 (two-

¹⁴In this hypothesis testing, we have removed the tasks associated to *Text Provider* and *Smartie* rewards because the fulfillment of the conditions depended on peers actions and not in the own student.

tailed, $\alpha=.05$, $n=8$) (Navidi, 2008). Therefore, we can conclude that the median weight of the dates when students claimed and earned the gamified-task conditions in the BADGE group is significantly different from the median weight in the REDEEM group (17 hours, 55 minutes and 4 seconds later) with a *p-value* of 0.016 and a very large effect size ($r = 0.604$) (Cohen, 1988).

According to the results, students from the REDEEM group claimed the rewards significantly earlier than those students from the BADGE group. Therefore, we can confirm the Hypothesis 5 and state that students in the REDEEM condition claimed and earned the rewards sooner (17 hours, 55 minutes and 4 seconds) than the students in the BADGE condition.

5. Discussion

In this study, the gamification was co-designed with the instructor of the course to encourage learners carry out the optional activities which were considered important for their learning, and to indirectly enhance their engagement within the course.

The analysis of the results presented in the previous section showed that the implemented reward strategies: (H1) did not lead to a higher student retention; (H2) did not increase the student behavioral engagement (measured in terms of number of pageviews, number of completed tasks, number of forum posts and activity time in the course); and (H4) did not encourage learners to perform earlier the optional activities. Although these results may contradict the results reported in many gamification studies in online and blended learning, they are in line with some studies performed in massive environments where gamification did not have the expected benefits on student retention and engagement (Rizzardini et al., 2016; Kyewski & Krämer, 2018). These results highlight the importance of the context (*e.g.*, MOOCs) and the individuals (*e.g.*, heterogeneity of learners) in the gamification design (Hamari et al., 2014; Seaborn & Fels, 2015).

On the other hand, the results also showed that the implemented rewards: (H3) significantly encouraged participants to satisfy the conditions associated with rewards; and (H5) affected the time when rewards were claimed (*i.e.*, redeemable rewards were claimed sooner than badges), which could be interpreted as a higher student intentionality to earn such rewards. These results support the idea that rewards helped the instructor achieve her main goal when introducing gamification in the course: to encourage participants to perform optional tasks.

Thus, the overall results suggest that those learners who are unlikely to complete the MOOC due to external reasons (*e.g.*, lack of time, lack of previous knowledge, or lack of interest on the course contents), will neither be motivated or engaged with the reward strategies (regardless of the reward type used). However, it seems that reward strategies can potentially encourage learners who are already motivated to complete the course (*e.g.*, interest on course topic and contents) to perform the optional tasks that would otherwise not be fulfilled.

Finally, (H5) the sooner claiming of rewards by participants in the REDEEM condition suggests a higher extrinsic motivation caused by the associated privileges. In this case,

the privileges associated with deadline extensions of compulsory assignments and instructors' feedback were the most valued by participants, which was also observed in a previous study conducted in a blended course (O'Donovan et al., 2013). This extrinsic motivation can be also used by MOOC designers and instructors to enhance the attainment of some specific pedagogical goals. However, the inclusion of these privileges into MOOC learning designs brings new variables in the alignment of the gamification design with the expected learning goals (*e.g.*, what activities should incorporate privileges, which privileges should be implemented). Therefore, the development of guidelines and tools supporting practitioners to successfully put in practice this kind of gamification strategies in MOOCs appears as a promising line of future research.

6. Conclusions

Despite the increasing number of works proposing the use of gamification strategies in MOOCs, there is a scarcity of empirical works testing the effects of such strategies (Ortega-Arranz et al., 2017; Antonaci et al., 2017; Khalil et al., 2018). In some cases, the literature has proposed the use of gamification as an easy-to-implement strategy to diminish some of the current drawbacks of MOOCs such as the attrition rates, the lack of motivation and engagement, the learner performance or the lack of interaction. This paper provides empirical evidence about whether reward-based gamification increases student engagement and retention in MOOCs.

Although no significant differences were found on student general engagement and retention among the different conditions, participants in the experimental conditions (Badges and Redeemable Rewards) showed higher participation in tasks associated to rewards. This effect can be used by instructors to promote specific learning goals, such as the learners' completion of optional tasks beneficial for their learning (providing a good alignment between the gamification design and pedagogical goals).

This study has some limitations as it is based on a unique MOOC oriented to Spanish-speaking population. It would be interesting to analyze to what extent this MOOC and gamification design (*e.g.*, number of implemented rewards, type of activities associated to rewards, rewards only visible to students themselves, rewards associated with optional tasks) affected the results of this study, and if the adaptation of this design to other topics and contexts would have similar effects (Hamari et al., 2014; Seaborn & Fels, 2015). Further evaluations involving reward-based strategies in other MOOCs with different topics, features, language and target population are needed to generalize the results of this study. In future versions of the same course, we plan to analyze the extent to which the aforementioned gamification parameters would change the effects on retention and engagement reported in this study.

Participants in MOOCs can typically be classified according to specific profiles based on their background, goals and participation (Kizilcec et al., 2013). In order to avoid a bias caused by the different types of learners in our study, we performed a Chi-square homogeneity test, thus checking if there was any significant difference between the experimental and control conditions regarding variables such as the background, knowledge level and the

expected type of participation in the course. Results showed that conditions are homo-
480 geneously distributed according to these variables. As a future work, we plan to analyze
the relationship between the different student types and the reward-derived behavioral en-
gagement (*e.g.*, if students with higher knowledge level of the course topic earned more
rewards).

Nevertheless, despite these apparent benefits, the design and implementation of such
485 gamification strategies can be time-consuming and cognitively-costly for MOOC practition-
ers. When using reward strategies, practitioners are responsible of (*i*) gamifying the learning
design and aligning the pedagogical goals with the gamification intentions (*e.g.*, create the
reward-condition rules); (*ii*) implementing the gamification design in the platform (*e.g.*, learn
490 how to use the gamification platform, make changes in the gamification design according to
the platforms capabilities); and (*iii*) managing the evolution of gamification during course
runtime (*e.g.*, watch over the effect of rewards on student behavior). All these gamification-
related activities imply an extra time and effort (Dicheva et al., 2018) added to the existing
work employed by practitioners to produce a MOOC. Such extra time and effort may hinder
the use and adoption of these gamification strategies. As a future work, we plan to explore
495 the affordability of GamiTool for MOOC practitioners.

Acknowledgements

This research has been partially funded by the European Regional Development Fund
and the National Research Agency of the Spanish Ministry of Science, Innovations and
Universities under project grants TIN2017-85179-C3-2-R and TIN2014-53199-C3-2-R; by the
500 European Regional Development Fund and the Regional Ministry of Education of Castilla
y León under project grant VA257P18; and by the European Commission under project
grant 588438-EPP-1-2017-1-EL- EPPKA2-KA. The authors thank the rest of the GSIC-
EMIC research team for their valuable ideas, and the Canvas Network team for the support
received conducting this research.

505 References

- Adamopoulos, P. (2013). What Makes a Great MOOC? An Interdisciplinary Analysis of Student Retention
in Online Courses. In *Proceedings of the 34th International Conference on Information Systems*.
- Alario-Hoyos, C., Pérez-Sanagustín, M., Delgado Kloos, C., Parada G., H. A., & Muñoz-Organero, M.
(2014). Delving into Participants' Profiles and Use of Social Tools in MOOCs. *IEEE Transactions on*
510 *Learning Technologies*, 3, 260–266.
- Anderson, A., Huttenlocher, D., Kleinberg, J., & Leskovec, J. (2014). Engaging with Massive Online Courses.
In *Proceedings of the 23rd International Conference on World Wide Web* (pp. 687–698). ACM.
- Antonaci, A., Klemke, R., Kreijns, K., & Specht, M. (2018). Get Gamification of MOOC right! How to
Embed the Individual and Social Aspects of MOOCs in Gamification Design. *International Journal of*
515 *Serious Games*, 5, 61–78.
- Antonaci, A., Klemke, R., Stracke, C. M., & Specht, M. (2017). Gamification in MOOCs to enhance users'
goal achievement. In *Proceedings of the 2017 Global Engineering Education Conference* (pp. 1654–1662).
IEEE.

- Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2013). Improving participation and learning with gamification. In *Proceedings of the 1st International Conference on gameful design, research, and applications* (pp. 10–17). ACM.
- Berger, P., Hennig, P., Bocklisch, T., Herold, T., & Meinel, C. (2016). A journey of bounty hunters: Analyzing the influence of reward systems on stackoverflow question response times. In *2016 IEEE/WIC/ACM International Conference on Web Intelligence (WI)* (pp. 644–649). IEEE.
- Borrás-Gené, O., Martínez-Núñez, M., & Fidalgo-Blanco, Á. (2014). Gamification in MOOC: Challenges, Opportunities and Proposals for Advancing MOOC Model. In *Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality* (pp. 215–220). ACM.
- Borrás-Gené, O., Martínez-Núñez, M., & Fidalgo-Blanco, Á. (2016). New Challenges for the Motivation and Learning in Engineering Education Using Gamification in MOOC. *International Journal of Engineering Education*, *32*, 501–512.
- Bunchball, I. (2010). Gamification 101: An Introduction to the Use of Game Dynamics to Influence Behavior. *White paper, available at <http://jndglobal.com/wp-content/uploads/2011/05/gamification1011.pdf>, last access: January, 2019, 9.*
- Chang, J.-W., & Wei, H.-Y. (2016). Exploring Engaging Gamification Mechanics in Massive Online Open courses. *Journal of Educational Technology & Society*, *19*, 177–203.
- Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization*, *81*, 1–8.
- Codish, D., & Ravid, G. (2014). Academic course gamification: The art of perceived playfulness. *Interdisciplinary Journal of E-Learning and Learning Objects*, *10*, 131–152.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edition).
- Cross, S., Whitelock, D., & Galley, R. (2014). The use, role and reception of open badges as a method for formative and summative reward in two Massive Open Online Courses. *International Journal of e-Assessment*, *4*, 1–16.
- Csikszentmihalyi, M. (1991). *Flow: The psychology of optimal experience*. New York: HarperPerennial.
- Davis, D., Chen, G., Hauff, C., & Houben, G.-J. (2018). Activating learning at scale: A review of innovations in online learning strategies. *Computers & Education*, *125*, 327–344.
- De Sousa Borges, S., Durelli, V., Reis, H., & Isotani, S. (2014). A Systematic Mapping on Gamification Applied to Education. In *Proceedings of the 29th Annual ACM Symposium on Applied Computing* (pp. 216–222). ACM.
- Deci, E. L., Koestner, R., & Ryan, R. M. (2001). Extrinsic rewards and intrinsic motivation in education: Reconsidered once again. *Review of educational research*, *71*, 1–27.
- Deng, R., Benckendorff, P., & Gannaway, D. (2019). Progress and new directions for teaching and learning in MOOCs. *Computers & Education*, *129*, 48–60.
- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From Game Design Elements to Gamefulness: Defining Gamification. In *Proceedings of the 15th ACM International Academic MindTrek Conference: Envisioning Future Media Environments* (pp. 9–15).
- Dicheva, D., Dichev, C., Agre, G., & Angelova, G. (2015). Gamification in Education: A Systematic Mapping Study. *Journal of Educational Technology & Society*, *18*, 75–88.
- Dicheva, D., Irwin, K., & Dichev, C. (2018). Oneup: Supporting practical and experimental gamification of learning. *International Journal of Serious Games*, *5*, 5–21.
- Ding, L., Kim, C., & Orey, M. (2017). Studies of student engagement in gamified online discussions. *Computers & Education*, *115*, 126–142.
- Domínguez, A., Saenz-de Navarrete, J., De-Marcos, L., Fernández-Sanz, L., Pagés, C., & Martínez-Herráiz, J.-J. (2013). Gamifying learning experiences: Practical implications and outcomes. *Computers & Education*, *63*, 380–392.
- Ferguson, R., & Clow, D. (2015). Examining engagement: analysing learner subpopulations in massive open online courses (MOOCs). In *Proceedings of the 5th International Conference on Learning Analytics And Knowledge* (pp. 51–58). ACM.
- Ferguson, R., & Sharples, M. (2014). Innovative pedagogy at massive scale: teaching and learning in

- 570 MOOCs. In *Open Learning and Teaching in Educational Communities* (pp. 98–111). Springer.
- Festinger, L. (1954). A theory of social comparison processes. *Human relations*, 7, 117–140.
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School Engagement: Potential of the Concept, State of the Evidence. *Review of educational research*, 74, 59–109.
- Gurantz, O. (2015). Who loses out? Registration order, course availability, and student behaviors in
575 community college. *The Journal of Higher Education*, 86, 524–563.
- Hakulinen, L., Auvinen, T., & Korhonen, A. (2013). Empirical Study on the Effect of Achievement Badges in TRAKLA2 Online Learning Environment. In *Proceedings of the 2013 International Conference on Learning and Teaching in Computing and Engineering* (pp. 47–54). IEEE.
- Hamari, J. (2017). Do badges increase user activity? A field experiment on the effects of gamification.
580 *Computers in human behavior*, 71, 469–478.
- Hamari, J., & Eranti, V. (2011). Framework for designing and evaluating game achievements. In *Proceedings of the 2011 DiGRA International Conference*.
- Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does Gamification Work? - A Literature Review of Empirical Studies on Gamification. In *Proceedings of the 47th Hawaii International Conference on System Sciences*
585 (pp. 3025–3034). IEEE.
- Henderikx, M. A., Kreijns, K., & Kalz, M. (2017). Refining success and dropout in massive open online courses based on the intention–behavior gap. *Distance Education*, 38, 353–368.
- Henrie, C. R., Halverson, L. R., & Graham, C. R. (2015). Measuring student engagement in technology-mediated learning: A review. *Computers & Education*, 90, 36–53.
- 590 Hew, K. F. (2016). Promoting engagement in online courses: What strategies can we learn from three highly rated MOOCs. *British Journal of Educational Technology*, 47, 320–341.
- Hone, K. S., & El Said, G. R. (2016). Exploring the factors affecting MOOC retention: A survey study. *Computers & Education*, 98, 157–168.
- Ibáñez, M. B., Di-Serio, A., & Delgado Kloos, C. (2014). Gamification for Engaging Computer Science
595 Students in Learning Activities: A Case Study. *IEEE Transactions on Learning Technologies*, 7, 291–301.
- Jordan, K. (2014). Initial trends in enrolment and completion of massive open online courses. *The International Review of Research in Open and Distributed Learning*, 15, 133–160.
- Khalil, H., & Ebner, M. (2014). MOOCs Completion Rates and Possible Methods to Improve Retention - A
600 Literature Review. In *Proceedings of the 2014 World Conference on Educational Multimedia, Hypermedia and Telecommunications* (pp. 1305–1313).
- Khalil, M., Ebner, M., & Admiraal, W. (2017). How can Gamification Improve MOOC Student Engagement? In *Proceedings of the 11th European Conference on Games Based Learning* (pp. 819–828).
- Khalil, M., Wong, J., de Koning, B., Ebner, M., & Paas, F. (2018). Gamification in MOOCs: A Review of the
605 State of the Art. In *Proceedings of the 2018 Global Engineering Education Conference* (pp. 1629–1638). IEEE.
- (2008). Test of homogeneity, chi-square test of homogeneity, chi-square. In W. Kirch (Ed.), *Encyclopedia of Public Health* (pp. 1386–1386). Springer Netherlands.
- Kizilcec, R. F., Piech, C., & Schneider, E. (2013). Deconstructing Disengagement: Analyzing Learner
610 Subpopulations in Massive Open Online Courses. In *Proceedings of the 3rd International Conference on Learning Analytics and Knowledge* (pp. 170–179). ACM.
- Krause, M., Mogalle, M., Pohl, H., & Williams, J. J. (2015). A Playful Game Changer: Fostering Student Retention in Online Education with Social Gamification. In *Proceedings of the 2nd ACM Conference on Learning @ Scale* (pp. 95–102). ACM.
- 615 Kyewski, E., & Krämer, N. C. (2018). To gamify or not to gamify? an experimental field study of the influence of badges on motivation, activity, and performance in an online learning course. *Computers & Education*, 118, 25–37.
- Margaryan, A., Bianco, M., & Littlejohn, A. (2015). Instructional quality of massive open online courses (MOOCs). *Computers & Education*, 80, 77–83.
- 620 Navidi, W. C. (2008). *Statistics for engineers and scientists*. McGraw-Hill Higher Education New York,

NY, USA.

Nicholson, S. (2015). A recipe for meaningful gamification. In *Gamification in education and business* (pp. 1–20). Springer.

625 O'Donovan, S., Gain, J., & Marais, P. (2013). A case study in the gamification of a university-level games development course. In *Proceedings of the South African Institute for Computer Scientists and Information Technologists Conference* (pp. 242–251). ACM.

Ortega-Arranz, A., Er, E., Martínez-Monés, A., Bote-Lorenzo, M. L., Asensio-Pérez, J. I., & Muñoz-Cristóbal, J. A. (2019). Understanding Students Behavior and Perceptions towards Earning Badges in a Gamified MOOC. *Universal Access in the Information Society*, (p. accepted).

630 Ortega-Arranz, A., Kalz, M., & Martínez-Monés, A. (2018). Creating Engaging Experiences in MOOC through In-Course Redeemable Rewards. In *Proceedings of the 2018 Global Engineering Education Conference* (pp. 1875–1882). IEEE.

Ortega-Arranz, A., Muñoz-Cristóbal, J. A., Martínez-Monés, A., Bote-Lorenzo, M. L., & Asensio-Pérez, J. I. (2017). How Gamification is Being Implemented in MOOCs? A Systematic Literature Review. In *Proceedings of the 12th European Conference on Technology Enhanced Learning* (pp. 441–447). Springer.

635 Penuel, W. R., Roschelle, J., & Shechtman, N. (2007). Designing formative assessment software with teachers: An analysis of the co-design process. *Research and Practice in Technology Enhanced Learning*, 2, 51–74.

Rizzardini, R., Chan, M., & Guetl, C. (2016). An Attrition Model for MOOCs: Evaluating the Learning Strategies of Gamification. *Formative Assessment, Learning Data Analytics and Gamification: In ICT Education*. Elsevier.

Ruipérez-Valiente, J. A., Muñoz-Merino, P. J., & Delgado Kloos, C. (2017). Detecting and Clustering Students by their Gamification Behavior with Badges: A Case Study in Engineering Education. *International Journal of Engineering Education*, 33, 816–830.

645 Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist*, 55, 68.

Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: A survey. *International Journal of Human-Computer Studies*, 74, 14–31.

650 Shah, D. (2017). By The Numbers: MOOCs in 2017. How has the MOOC space grown this year? Get the facts, figures, and pie charts. Retrieved from: <https://www.class-central.com/report/mooc-stats-2017/> last access: June 2018.

Siemens, G. (2013). Massive open online courses: Innovation in education. *Open educational resources: Innovation, research and practice*, 5, 5–15.

655 Zhu, Y., Pei, L., & Shang, J. (2017). Improving video engagement by gamification: A proposed design of MOOC videos. In *International Conference on Blended Learning* (pp. 433–444). Springer.

Zichermann, G., & Cunningham, C. (2011). *Gamification by design: Implementing game mechanics in web and mobile apps*. O'Reilly Media, Inc.