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*(2–B), 816–830.

https://www.ijee.ie/contents/c330217B.html

Detecting and Clustering Students by their Gamification Behavior with Badges: A Case Study in Engineering Education

JOSÉ A. RUIPÉREZ-VALIENTE[a,b], PEDRO J. MUÑOZ-MERINO[a], CARLOS DELGADO KLOOS[a]
[a] Universidad Carlos III de Madrid, Avenida Universidad 30, 28911 Leganés (Madrid) Spain
[b] IMDEA Networks Institute, Av. del Mar Mediterráneo 22, 28918 Leganés (Madrid) Spain
{jruipere, pedmume, cdk}@it.uc3m.es

Abstract: Engineering degrees are often regarded as complex and one usual issue is that students struggle and feel discouraged during the learning process. Gamification is starting to play an important role in education with the objective of providing engagement and improving the motivation of students. One specific example is the use of badges. The analysis of users' interactions and behaviors with the badge system can be used to improve the learning process, e.g. by adapting the learning materials and giving game-based activities to students depending on their interest toward badges. In this work we propose some metrics that provide information regarding the behavior of students with badges, including if they are intentionally earning them, the concentration for achieving them and their time efficiency. We validate these metrics by providing an extensive analysis of 291 different students interacting with a local instance of Khan Academy within our courses for freshmen at Universidad Carlos III de Madrid. This analysis includes relationship mining between badge indicators and others related to the learning process, the analysis of specific archetypal profiles of students that represent a broader population and also by clustering students by their badge indicators with the objective of customizing learning experiences. We finalize by discussing the implications of the results for engineering education, providing guidelines into how instructors can take advantage of the findings of the research and how researchers can replicate experiments similar to this one in other general contexts.

Keywords: badges; gamification; engineering education; distance learning; learning analytics; khan academy; educational data mining; modelling behavior

1. Introduction

Gamification can be defined as the use of game design elements in a non-game context [1]. Gamification techniques in education have been used broadly in the last years. The main reason to implement gamification elements in education is to improve the motivation and engagement of the student towards the learning goals. Engineering education classes are often complex and students might struggle. An issue in engineering education is the lack of motivation and engagement. We can find high rates of attrition and students do not have enough motivation to persist on the task of achieving an engineering degree [2]. Additionally, teachers find that majoring in engineering is harder than in many other subjects [3].

The introduction of gamified environments in engineering education can help ease the path of students and keep them motivated. An example of engineering software with gamification elements is Pex4Fun [4], where students can have duels and earn badges with their interaction with engineering problems. Albeit some students might be very motivated towards gamification, others can feel discouraged, and perceive gamification features as unnecessary and confusing from the real objective [5]. Therefore, here comes the need to develop indicators related to gamification, that can be integrated into the student model and use this information for adaptation purposes. However the inclusion of gamification indicators is not common on educational platforms, e.g. the review performed by Dyckhoff *et al*. [6] collects different indicators related to the learning process, but those that describe gamification features on educational contexts are not considered.

Learning analytics involves the collection and use of educational data, applying modeling techniques to better understand student learning [7]. We propose the application of learning analytics technology to collect, transform and analyze data in order to detect which students are motivated and engaged with gamification features and in which way they use the gamification features. In this study we try to combine both approaches to analyze the behavior of students with respect to the gamification system in an educational case study. Specifically, we focus on one of the most used gamification techniques in education, which is the provision of badges. Badges are awards that students can receive by doing specific actions on an educational platform (e.g. some actions can be participatory like logging into the platform for several days in a row or related to learning skills, such as solving correctly 5 different exercises).

The inference of indicators about students' behaviors and interactions with badges is very important as it would shed some light into the question of which students are really finding benefits on gamification and those who are not. These indicators can be integrated into the student model and use this information for adaptation purposes. For example, if we can infer which students make actions with the intention to acquire badges and which students acquire badges without intention, then we can know which students are motivated by these achievements and we can adapt the learning process by applying specific learning techniques for these students. In addition, if we know the badges that are used more often and compare them with others, then we can make hypotheses about which badges might motivate more students or which of them have requirements too demanding; then we can make the proper actions to improve the learning process.

For the aforementioned reasons we believe that the combination of both gamification and learning techniques can improve the current educational approaches on engineering education where gamification features are implemented with the objective of improving the motivation and engagement of students with the learning flow. Specifically in this study, we focus on analyzing students' interactions and behavior with badges in an experiment with 291 engineering students using the Khan Academy platform at Universidad Carlos III de Madrid. The specific objectives of this research are the following:

- Design and implement technological detectors that provide indicators to measure the intentionality, concentration and efficiency of students for the achievement of badges.
- Apply, calculate and analyze the proposed indicators in a real experiment with 291 different engineering students using the Khan Academy e-learning platform.
    - Analyze the relationship among students' intention, concentration and efficiency for the achievement of badges with other typical indicators of students' performance and behavior.
    - Analyze different profiles of students and be able to group students by their behavior with badges by applying clustering methods with the ultimate goal of being able to find different behavioral profiles towards badges that can be used to adapt the learning experiences of students.
- Discuss the implications of the results for engineering education, giving insights about how to take advantage of these results in order to provide badges or gamification activities to some students, or how to provide the gamification activities to the students.

More specifically, the remainder of this paper is organized as follows. Section 2 presents a review of related works about gamification, particularly its application to education, and learning analytics. Section 3 describes the methods used in the study, where in the first subsection we present the implemented detectors to infer the behavior of students with badges and in the second subsection we describe the case study. Section 4 presents the results regarding relationship and clustering analysis of students in different types of profiles and Section 5 discusses the application of these findings in engineering education providing also conclusions and future work.

2. Related work

The use of games is widely spread among different contexts, which encourages the use of game elements for educational purposes in order to provide a more immersive learning flow and improve engagement [8]. Many studies apply game components in non-game contexts. These techniques have been used in many different contexts; for example the introduction of game achievements in a photo sharing service [9] or the inclusion of gamification elements in eco-driving [10] showing a positive correlation with the use of the proposed eco-driving tips. Gamification has been tested in different e-learning experiments, reporting positive results. Some examples include a gamified AutoCAD tutorial [11] or a gamification system for Blackboard [12]. Hamari, Koivisto and Sarsa [13] presented a literature review of the empirical studies on gamification analyzing 24 research works. The results indicated that gamification yielded positive effects as a general rule but these effects were strongly dependent on the contexts and the users of the experiment.

Additionally, learning analytics has been gaining importance within the educational field over the last years. In this work we use learning analytics techniques in order to model the behavior of the students with badges. In the literature, we can find several learning analytics tools to improve the support for instructors and students. As an example ALAS-KA [14] is a plugin for the Khan Academy platform that provides a set of more than 21 new different indicators related to the learning process; all the indicators have been grouped in 5 different modules of knowledge that can be used by instructors and students. Another example is Moodog [15] which is a plugin for Moodle which uses the data generated by the CMS to provide new statistics which have been categorized into 4 different groups. But indicators about the gamification process are not usually included, and that is one of the reasons why we feel encouraged towards advancing in developing new indicators related to the behavior of students with badges. Learning analytics have already been applied to educational games in different works. The amount of events triggered by users playing games makes harder to select which data can be useful to infer useful information. The work presented by Serrano-Laguna *et al*., [16] uses the source of data from an educational game to feed a learning analytics system to infer knowledge about the effectiveness of the students. Moreover, real-time learning analytics in educational games need to adequately match the dynamics of a game environment [17].

Engineering education is tough for students, and it is often regarded as more difficult than other degrees. The engineering community is trying to find new models that help engineering careers be more engaging [3]. The use of games and also gamification elements in engineering courses and software for learning engineering topics is becoming more and more widespread. As an example we can find successful experiments where mathematics computer games have greatly improved the motivation of students compared to those who did not play them [18]; another example is with the specific problem of teaching computer programming skills to increase the success of novice programmers, where the learning curve might be quite abrupt [19]. We also can find in the literature several studies regarding gamification in education. Some works have reported very positive results using gamification in engineering classes, where students improved their level of understanding, their learning and reduced the stress of complicated classes [20]. Another study applies different elements such as scoring, levels, leaderboards, and badges finding that student satisfaction improved with respect to more traditional classes, as well as lecture attendance and other proactive behaviors [21]. However, gamified experiments can also led to negative side effects that can deteriorate the overall feeling of the learning experience. Specifically, gamified environment needs to be carefully tailored to the specifies of the course and tasks that need to be developed, and also to students that are taking the class [22]. Another possible counterpart is that gamification can be perceived as an extrinsic motivator, that is why a detailed design needs to be done, which should focus on boosting intrinsic motivational factors. We can find a successful experiment which tries to focus on intrinsic motivators within an Object-oriented Programming course, by improving the control of students, enhancing cooperation and the possibility to gain recognition [23]. Moreover, some techniques such as the competition can bring negative emotions but some competition systems have been created such as ISCARE [24], which can motivate students without generating negative emotions [25].

In this work, our context is remedial courses for students accessing engineering degrees, and the environment is Khan Academy, which is considered as one of the pioneer platforms in Massive Open Online Courses (MOOCs). There are many MOOCs on engineering education, however the huge dropout rates on these courses are probably the most important handicap. That is why gamification can play an important role to promote intrinsic motivation and participation of students to help improving completion rates also in MOOCs [26]. Additionally, we focus on a specific gamification element, which is the use of badges. In the literature there are other works which made use of badges in education. Some results [27] concluded that badges can have a positive effect on learner motivations, however as an extrinsic reward, instructors and researchers need to consider the possibility that badge systems can be counter-effective to the real goal, which is learning in this case. In our research, we did find some behaviors that might imply a real interest in badges and increase of participation, albeit that might not imply more learning and in specific students it could be not beneficial. There is a description and evaluation of the use of achievement badges in the TRAKLA2 online learning environment with 281 students [28]. They found that some badges can be used to affect the behavior of students and that differences in students' behavior exist between each type of badge. Our conclusions in that regard are similar, as we also found that some badge types affected more the behavior of students than others. The systems Septris and SICKO [29] implement both learning analytics and gamification techniques to ascertain its viability in medical education obtaining good opinions from the users – although they did not try to address the behavior of users towards badges but only global results of participation. However, most of these works only report quantitative results either by surveys or by summarizing global usage of the platform. In this study we aim at modeling the behavior of students with the badge system, as done in a different context like Stack Overflow [30], [31]. The work of Gant & Betts [31] analyze the behavior of the users with only three types of badges as it is very hard to analyze all of them; that is the same approach that we follow, as we focus mainly on *topic* and *repetitive badges* (as explained in next section). It was demonstrated that the different badges indeed affected the behavior of users, boosting their amount of participation with the system. Nevertheless, after researching on the literature, we have not found other works that try to model the behavior of students with a badge system the same way as we do in this manuscript.

3. Methods

3.1. Description of the case study

The case study is framed within the context of 0-courses that freshmen students take before starting their first year of an engineering degree at Universidad Carlos III de Madrid. These courses have an online period of time in August where students review the concepts by themselves. Next, students start the face-to-face lessons that take place in September. A "flipped classroom" methodology is applied in which students should prepare the lessons at home before receiving the actual lecturing class. This methodology was used for the first time as a pilot initiative in August 2012 for a physics course and repeated again in following years for physics, mathematics and chemistry courses due to the success of the pilot experience. The data analysis that is presented here belongs to the physics, mathematics and chemistry courses of August 2013. The courses are composed by a series of exercises and videos that have been developed by the instructors of each course. A total number of 30 exercises and 25 videos in mathematics, 30 exercises and 30 videos in physics, and 49 exercises and 22 videos in chemistry are available.

Students who participate in these courses are freshmen who have enrolled to an engineering degree. Most of them have an age comprised from 17 to 19 years. The number of students is different for every course, taking into account that students might have different requisites depending on the engineering degree that they have enrolled. The number of students is 167 for the physics course, 73 in chemistry and 243 in mathematics. Note out that we have removed students who did not log into the Khan Academy platform. In addition, as some students had to enroll in more than one course depending on their engineering degree, the number of different students taken into account for this study is 291.

3.2. Gamification system in Khan Academy

Khan Academy incorporates innovative features such as learning analytics visualizations and gamification elements. These features are oriented to improve the learning experience as well as the engagement of students with the platform. Among the different gamification features of the Khan Academy platform, we can find a scoring system, setting up goals, and badges. We considered which one of the gamification features should be used for the study; we decided to remove the scoring system since it is not clear for students how to get more points and goals because only a few students made use of this specific feature. Khan Academy incorporates a wide badge system, which facilitates this study where we focus on the interaction with badges. Additionally, it is very easy for students to access the overview badge page[1] with the requisites to earn each badge, thus we think since the badge system is very transparent for them, it is the best choice to try to infer motivation and intentionality in the behavior of students. Some of the badges are easy to achieve, and others are really challenging. In addition, there are different types of badge requirements, e.g. those related to problem solving, video watching or social activity, among many other things. Overall, there are way too many badges to try to infer the behavior of students with all of them, that is why in this study we focus mainly on two categories of badges, which is an approach that has been followed previously in other studies [29]. The two categories have specific characteristics as we describe next:

- ***Topic badges***: These badges are awarded to students when they accomplish to earn proficiency in a set of exercises (skills). In our experiment, the required exercises to earn one badge are always different from the others. This means that each problem will belong only to the requirements of one *topic badge*. Each one of these badges can be earned only once by student. As part of the experiment the badge system was customized and new badges were added in the case of *topic badges*, to match

[1] https://www.khanacademy.org/badges

the exercises that were developed for each one of the courses. A total number of 7, 12 and 16 *topic badges* were designed for the mathematics, physics and chemistry courses respectively. The amount of topic badges was in relationship with the amount of exercises in each course and the relationship between exercises, as related exercises were united to provide a *topic badge* about a specific area of knowledge.

- ***Repetitive badges***: We classify within this category those badges that can be earned repetitively by the same student as many times as students want (as long as they keep fulfilling the required conditions). Specifically in our experiment, we have two types of badges that fall within this category, which are called as *Timed Problem* and *Streak badges*. The first ones are delivered when solving problems rapidly, and the second ones when solving several exercises correctly in a row. Each one of these two types of badges have 5 different levels; for example the *Streak* type have 5 different badges which are called *Nice*, *Great*, *Awesome*, *Ridiculous* and *Ludicrous Streak badge* which are earned whenever the student correctly solves 20, 40, 60, 80 and 100 exercises in a row respectively. The different levels of the *Timed Problem* type are quite similar as the former one. So there are a total number of 10 badges that can be earned repetitively and is the same amount for the three courses.

As mentioned before, despite we focus on the two aforementioned categories of badges, there are many more in the system. For metrics that take into account the total number of badges, we consider a total amount of 53, 61 and 65 badges for mathematics, physics and chemistry courses respectively.

3.3. Modelling students' behavior with badges

Some indicators which model students' behavior towards gamification based on badges are proposed in this section. The objective is to propose some measures which can provide a deep insight about the interaction of users with badges. The defined metrics are four: intentionality on topic badges (ITB), intentionality on repetitive badges (IRB), concentration on achieving badges (CAB) and time efficiency in badges (TEB). The specific Python scripts implemented to calculate these badge metrics mining the App Engine Datastore of our Khan Academy instance have been open sourced in a public GitHub repository[2]. For each indicator, we will first introduce the base idea about the behavior we are trying to detect and next how to operationalize it in our specific case study with Khan Academy.

3.3.1. Intentionality for earning badges

The requisites needed to receive each one of the badges are available in the achievements interface of the Khan Academy platform. Therefore, students have a full access to this information and they can know what actions must be fulfilled to obtain each award. In this direction, it is interesting to get to know if students are trying to maximize the amount of badges that they earn with their actions or if they are receiving these badges just as part of their interaction with educational activities but without any intention to get them. This knowledge would allow us to determine which students are motivated about earning badges and to adapt some decisions to this knowledge. For example, a student might be solving exercises in a way to get badges according to the badge requisites, but another one might be solving exercises in a way that he/she might not get any badges. As the badge system in the Khan Academy is very wide, we cannot infer this information from all the different types of badges. For example, how would it be possible to know if a student has commented on a video in order to obtain a social badge or just because he wanted to make a question? However, there are other cases in which we can try to infer the students' intention for earning badges. This is the case of the *topic* and *repetitive badges*. We previously presented preliminary results regarding these two intentionality metrics but in this work we provide a much more thorough analysis [32].

3.3.1.1. Intentionality on *topic badges* (ITB)

The subset of exercises that are required to earn a *topic badge* are always different; therefore, we can try to infer, if a student is trying to maximize the number of *topic badges* that he/she is acquiring, or if they are earning them as just part of the learning process. As an example, think out a student who has achieved proficiency in 21 exercises earning only two badges; now imagine a second student that has achieved

[2] https://github.com/jruiperezv/clustering-badges

proficiency in the same amount of exercises earning seven badges. Most probably we assume infer that the first student was not very interested in maximizing the amount of badges whilst the second one is showing more interest and intentionally fulfilling badge requirements.

Figure 1 shows the flow diagram that we implemented in the Khan Academy platform in order to infer if a student is maximizing the number of badges acquired, and thus showing an intentional behavior. The implemented algorithm gets the number of problems that the student has achieved as well as the number of *topic badges* earned by the student; then it calculates the maximum number of *topic badges* that the student could have earned with that number of proficient problems in case he/she intended to do that. Finally the metric provides a percentage comparison that can range from 0, meaning that the student does not maximize at all, to 100, meaning that the student maximized the number of *topic badges* that he/she could earn.

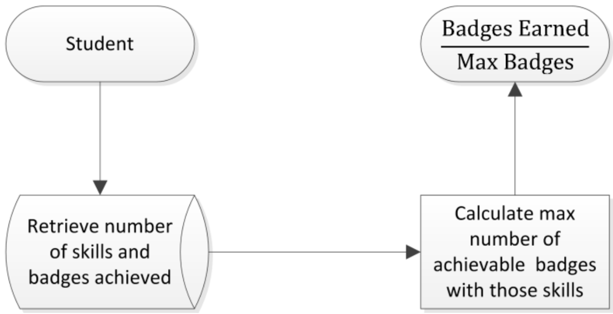

Figure 1. Flow diagram to calculate the intentionality for *topic badges*.

3.3.1.2. Intentionality on *repetitive badges* (IRB)

In our experiment students should solve exercises until they achieve a proficiency level on the exercise (skill). We focus on measuring intentionality on these badges because we can know if students are earning the badge as part of the normal learning process or not. Therefore, if we see a student that keeps earning *repetitive badges*, then we would be able to detect if the student is intentionally earning more and more of the same type.

The specific rationale that follows this implementation is based on the idea that students should solve exercises of the same type, until they get a proficient level and master the skill. Upon achievement of proficiency, students receive a notification from the system, and they should stop doing exercises of this same skill and move on to the next one. In the case that students keep doing exercises of a skill in which they are already proficient, and they keep earning *repetitive badges* this way, we hypothesize that are earning those badges on purpose. Finally, we can compute a percentage on the amount of *repetitive badges* that were earned intentionally, and that would give us a clue about if the student is trying to earn badges on purpose instead of as part of the learning process.

3.3.2. Concentration on achieving badges (CAB)

Students can devote all their consecutive actions into fulfilling the requirements of one badge. On the other hand, students can carry out different actions in the middle, which are not related to the requirements of that badge before actually receiving the badge. The definition of concept of concentration is independent from the one of intentionality, because one aspect is the intention to achieve a badge while another aspect is if the student is focused on earning a specific badge and all his/her consecutive actions are to this end or if there are intermediary actions in between which are not related to earning that specific badge. However, the two concepts might have a correlation since if students are concentrating all their consecutive interaction into achieving a badge, then they are intentionally earning them.

As stated before, the wide variety of requisites in the badge system makes very difficult to propose a model that could be applicable for all of them. In addition, some requisites of some badges do not apply well into this indicator. For example, it does not make sense to say that a student is concentrated into voting a video. This concentration indicator would apply only to those badges which are awarded for a specific chain of students' actions. A similar approach [31] is found on a work on Stack Overflow, where they focus on badges in which the progress of the user towards achieving that certain badge can be determined and not on those badges that are earned as a result of just one action.

We have designed and implemented an algorithm to detect students' concentration on *topic badges*. Since *topic badges* have as requisites a fixed set of exercises, we can track if students have done the required exercises in a consecutive way or others in the middle before earning a *topic badge*. Following this criterion we can infer the proportion of the previous exercises that a student attempted that actually belong to the requisites. It is important to remark that students have freedom to select the next exercise they are going to do in the platform and they can select any of them at any moment.
Figure 2 shows the flow diagram to infer the concentration of students when earning topic badges. We will retrieve all the badges earned by the user and analyze the previous actions for only those which are *topic badges*. For each *topic badge* that the student has received, the algorithm retrieves all the attempted exercises between the last *topic badge* (in case is the first one, it retrieves all the exercises attempted until the current time) and the current badge. Next, the algorithm analyzes which of the exercises belong to the requisites of the current badge. The following step is to calculate the concentration of the student when acquiring the current badge. Finally repeating this process for all *topic badges* that the student has earned, we can calculate a global level of concentration for the student. This level of concentration gives an idea about if the student is focused in fulfilling the requisites to receive badges in a continuous and concentrated way or just attempting exercises and earning the badges by chance or they like to earn badges but in a disperse way and not a concentrated one.

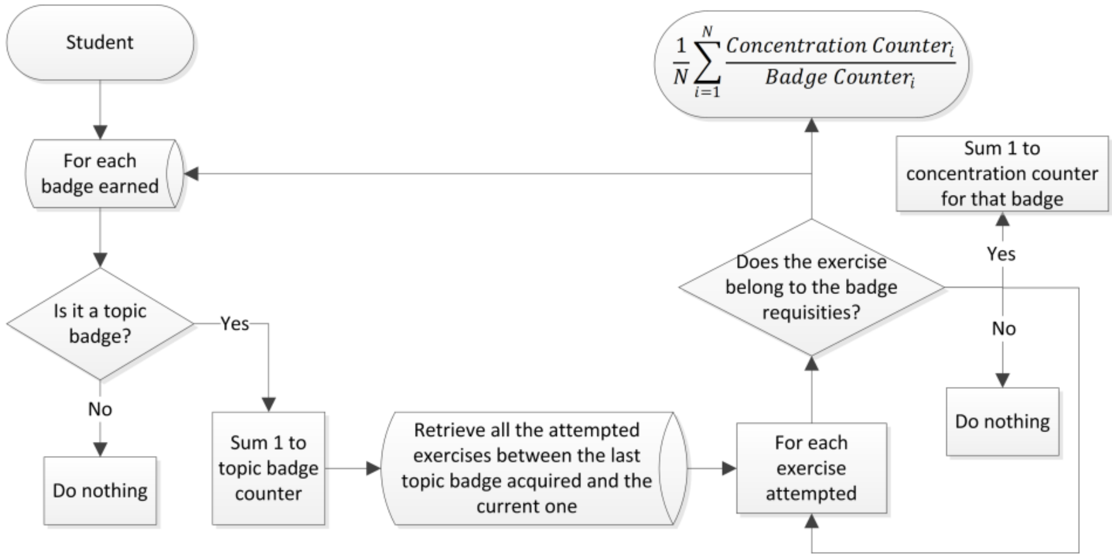

Figure 2. Flow diagram of the algorithm that calculates CAB indicator.

3.3.3. Time efficiency in badges (TEB)

Students devote a certain amount of time interacting with the platform. In addition, this time can be divided into different activities such as watching videos or solving exercises. The time invested by students can be used to obtain measures which give insight about the number of badges they earn per unit of time. This can allow a comparison between the different students to see which of them are more efficient in terms of time towards achieving badges. In this direction we propose an indicator for the time efficiency in badges (TEB) defined as the total number of badges divided by total time in the platform. Additionally, this indicator could be particularized for different badge types, i.e. only badges related to exercises and time spent in exercises. This indicator can provide supplementary information regarding the efficiency of students towards earning badges.

Although the definition of the time efficiency in badges indicator is independent from the intentionality definition, a correlation might be found with the intentionality because students who have intention to get the badges might have a better efficiency and earn more badges in less time.

4. Results

In this section we present the results after implementing and applying all the indicators of section 3.3 to the data generated by students in the case study described in section 3.1. We divide the results in three subsections, the first one performs relationship mining analysis, the second one reports some archetypal profiles of students and the last one clusters students in regards of their badge indicators. All the results of this section have been performed using R software, except for the Two Step Cluster analysis of Subsection 4.3 which was performed using SPSS.

4.1. Relationship between badge indicators and others

This subsection presents the correlation of the badge metrics with other metrics related to students' behavior during the learning process. In addition, we calculate the relationship of the gamification indicators among them. The objective is to present the existing and underlying relationships between the badge indicators and others, in order to learn more about the relationships of gamification behavior towards badges of the students in a Khan Academy course and make reasonable hypothesis. Table 1 describes the Bivariate Pearson correlation (N=291, two-tailed) of the badge indicators with several indicators. We apply a Bonferroni correction in order to reduce the probabilities of finding false significant correlations. The asterisk marks those correlations which are significant at the 99% level after applying the Bonferroni correction. We have divided this analysis into three sections which are separated by lines within the table. In addition to the badge metrics we include here other metrics related to the learning process which are exercises accessed (EAC), videos accessed (VAC), exercise abandonment (EAB), video abandonment (VAB), total time (TT), use of optional activities (OA), proficient exercises (PE), and completed videos (CV). These indicators are part of the two first sections of the table and have been previously described in our work [33][34]. The third group shows the correlation between the four badge metrics with themselves.

The first section of indicators is composed by EAC, VAC, EAB, VAB, TT and OA which are related to the total use of the platform. As we can see many of these indicators have been found statistically significant with badge metrics, which makes sense and we can hypothesize that the more use of the platform students do the more badges they earn. First thing to notice is that all these indicators have been found statistically significant with the indicator TEB. In addition, EAB and VAB have been found negatively and significantly correlated with several badge metrics. Although this correlation is pretty low, we believe that it might indicative that engagement plays an important role in the motivation towards earning badges. The indicators which were less correlated to badge metrics are VAC and VAB, which are coherent results taking into account that most of badge metrics does not take into account video activity. Finally EAC, EAB, TT and OA are strongly correlated with most of the badge metrics.

The second section of relationships is related to the correct progress in the platform. The correlation with CV is significant but low, as there are less video badges than those who are earned by solving exercises. The correlation with PE is the highest of all indicators which is probably related to the big amount of exercise badges that can be earned repeatedly and also due to *topic badges*. Two of these correlations are especially significant, first with ITB (0.737, $p < 0.000$), which makes sense since students who have more proficient exercises are more likely to earn more *topic badges* as well. Secondly with TEB (0.625, $p < 0.000$), which also makes sense due to the more proficient exercises, more badges the student will earn, hence TEB will also be higher.

The last section of the table which is separated by a double line presents the correlations among the badge metrics with themselves. All the correlations have resulted to be statistically significant probably due to the fact that when one student shows motivation towards earning badges it will be reflected in several of these indicators. We should point out the correlation between ITB and CAB (0.859, $p < 0.000$), which is the highest of all the correlation analysis. One hypothesis is that this correlation is very high as the students who are concentrated earning *topic badges* are also probably maximizing and earning as many *topic badges* as possible.

Table 1. Bivariate Pearson correlation of badge indicators with others indicators

Bivariate Pearson Correlation two-tailed (N=291)	*Intentionality on Topic Badges*	*Intentionality on Repetitive Badges*	*Concentration on Achieving Badges*	*Time Efficiency in Badges*
Exercises Accessed	0.456*	0.464*	0.361*	0.438*
Videos Accessed	0.305	0.322*	0.228*	0.225*
Exercise Abandonment	-0.456*	-0.327	-0.399*	-0.352*
Video Abandonment	-0.177*	-0.125*	-0.168	-0.049
Total Time	0.510*	0.409*	0.372*	0.338*
Optional Activities	0.489*	0.358*	0.345*	0.446*
Proficient Exercises	0.737*	0.511*	0.563*	0.625*
Completed Videos	0.326*	0.302*	0.251*	0.279*
Intentionality on Topic Badges	1	0.445*	0.859*	0.552*
Intentionality on Repetitive Badges	0.445*	1	0.417*	0.573*
Concentration on Achieving Badges	0.859*	0.417*	1	0.459*
Time Efficiency in Badges	0.552*	0.573*	0.459*	1

4.2. Analyzing the behavior of specific students

This subsection focuses on specific users and tries to select some archetypal students that can derive to represent a broader subset of the population. We provide a radar chart in figure 3 in order to visualize different types of profiles of students and be able to do a comparison regarding the different indicators. We will analyze their behavior as well as the interest for badges of the selected students based on the indicators. Figure 3 shows the radar chart of Students A, B, C, D and E. The values of the indicators represented in the plot have been normalized in order to be in the interval [0~1], where 1 would represent the highest possible value of the indicator among all students and 0 would represent the lowest one. In the radar chart we represent starting from the bottom and going counterclockwise the four badge indicators: TEB, CAB, IRB, and ITB in that order. Additionally, we present the indicators CV, PE and TT to help us have an idea about the interaction with the platform done by those students.

Student A has devoted a great amount of time in the platform, actually we can see that TT indicator goes straight to 1 which is the maximum normalized value, meaning that is the student who spent most time in the platform (2458 min). Just an additional note, that amount of time doubles the one of the second student in terms of TT in the experiment. Additionally we can also see also that Student A completed all videos and achieved proficiency in almost all exercises. He/she was able to acquire a total amount of 934 badges, and 43 different types of badges. Student A showed an impressive interest on the badge system, as we can see in both ITB and IRB indicators. But he did not do the actions necessarily in a consecutive way to achieve the badges, i.e. he did several activities in parallel to achieve different types of badges. In addition, he was not super-efficient in achieving badges, this can also explain the big amount of time that he needed to spend on the platform. Our hypothesis is that Student A has probably invested that impressive amount of time because he was very interested in earning badges and he just kept going and going becoming the top earner among all students. However he was not very efficient in achieving the badges and did several actions in parallel to get different types of badges instead of concentrating in doing all the actions consecutively to achieve a specific type of badge.

Student B devoted less time (1168 min) but still managed to complete a high percentage of the course (100% of videos completed and 60% of exercises), and showed a high IRB and CAB. However Student B did not have a high ITB. This might be caused due to the fact that achieving these types of badges is more difficult and requires more effort. Student C spent a similar amount of time (762 min) and made a similar progress (80% of the videos completed and 40% of the exercises) as Student B, however his/her badge indicators are very low which indicates that Student C did not have a high interest in badges.
Student D devoted a good amount of time (927 min) completing almost all videos and exercises in the course. The results also show that Student D has average badge indicators which pointed out that he/she

was not exceptionally motivated by badges but made use of them. Finally Student E invested a low amount of time (249 min), showing low progress where he only completed 16% of the exercises and 20% of the videos; however the results pointed out that he was very interested in badges as both intention and concentration indicators are almost in the maximum.

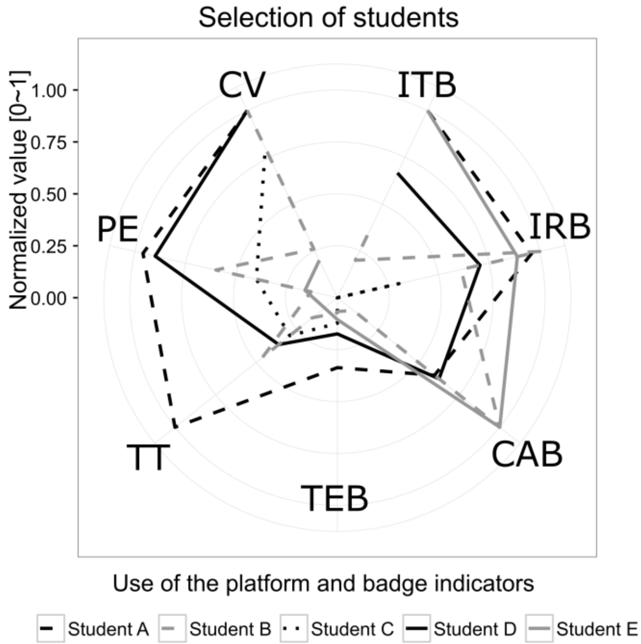

Figure 3. Radar chart of the indicators of five selected students with different profiles.

4.3. Clustering students by their badge indicators

In this subsection we apply an unsupervised learning algorithm in order to be able to cluster students in groups with similar badge indicators. We would like to know which students have been motivated by the badge activity and group them together according to their preferences. This information can be used to divide all the students into different groups in order to customize the teaching methodology of each group.

We have decided to use a Two-Step Cluster analysis since we do not know how many groups we should find. This algorithm gives its best performance when all variables are independent and have a normal distribution, however it is very robust and it performs well when these conditions are not met. Both Kolmogorov-Smirnov and Shapiro-Wilk test confirm that the distribution of the four badge metrics can be considered as normal and the variables are independent, thus the conditions are met.
We apply the Two-Step Cluster algorithm leaving the number of groups to be determined by the execution of the algorithm automatically, with a Log-likelihood measure for distance and a Schwarz's Bayesian Criterion for clustering. Finally, we choose the badge indicators ITB, IRB, CAB and TEB as continuous variables for the clustering algorithm. We will also use PE, CV and TT indicators as evaluation fields to support the interpretation of the results from the algorithm output.
The Two-Step algorithm has selected three clusters providing a good cluster quality in terms of cohesion and separation (0.75 in a quality ranging from -1 to 1). The smallest cluster has 70 students (24.1%) whilst the largest has 149 students (51.2%) providing a ratio of sizes of 2.13; the middle-sized cluster has 72 students (24.7%). The predictors' importance for the four continuous variables has been 1.0, 0.94, 0.79 and 0.31 for IRB, CAB, ITB and TEB respectively.

Figure 4 shows a boxplot visualization with the results about the distribution of the clusters. The horizontal axis represents the different indicators, and the cluster is encoded by greyscale, and always follows cluster 1, 2, 3 order from left to right. The vertical axis represents the scaled values (0 minimum to 100 maximum) of the distribution among all students. The upper plot represents the four badge metrics that were used to cluster students, and the bottom plot represents the evaluation fields.

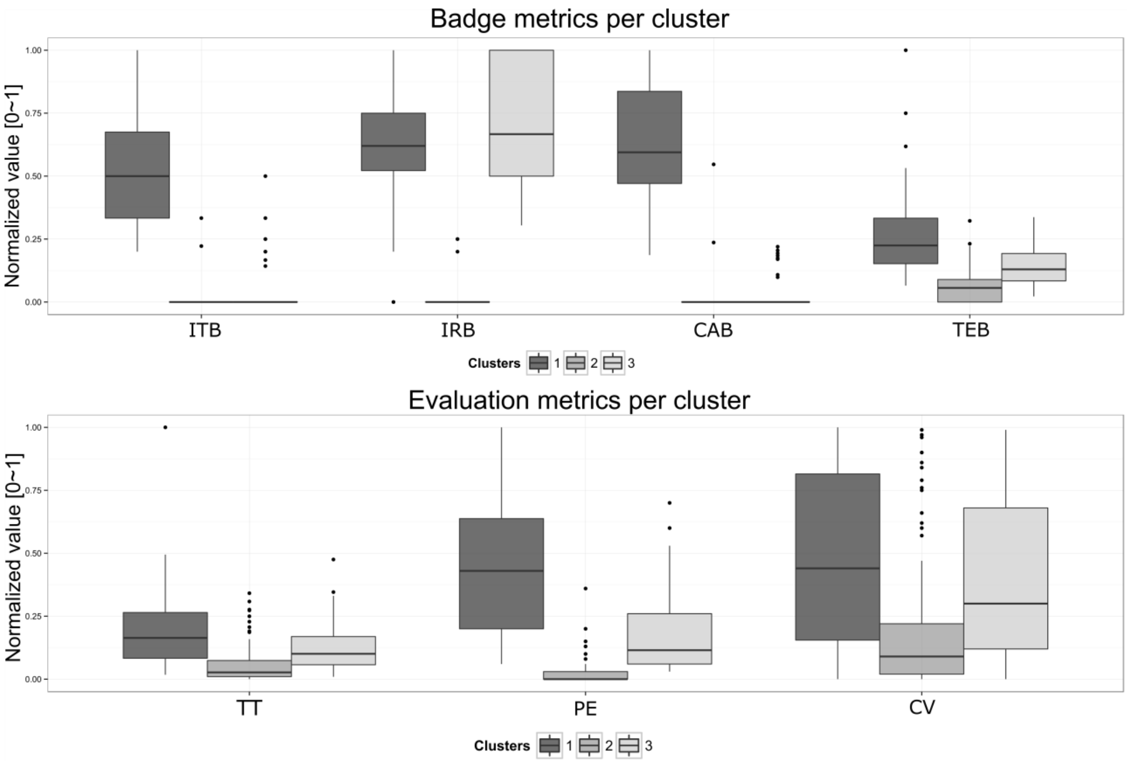

Figure 4. Overview of the results of a Two-Step Cluster analysis using badge indicators as inputs.

The information provided by figure 4 showing the boxplot visualization of the badge indicators and evaluation fields per cluster, can be used to learn what type of students compose each group:

- Cluster 1: The first cluster is composed by the 24.7% of students. We can rapidly perceive that the students who belong to this cluster are those who have put the greatest effort in the platform in terms of amount of PE, TT and CV. The mean value of PE is 46.88%, for CV 46.76%, and for TT is 489.2 min per user on average, which are all high values. In addition, they have high values in all badge metrics when compared to the rest of the clusters. The average user of this cluster made an important investment on time, as well as progress in exercises and videos, showing also interest in the badge system.
- Cluster 2: The second cluster is composed by the 51.2% of the students and it is quite the contrary of the first one. These students have made a low effort using the platform, we can see that on average they have invested 125.8 min per user obtaining only 1.72% in PE and 18.93% in CV. In addition, they have not shown interest in any of the badge indicators.
- Cluster 3: The third cluster is less clear than the others two being composed by the 24.1% of the population. We can see that students within this cluster have invested a decent amount of time with 314.8 minutes per student, however their progress has not been so good with only 17.91% of PE which is much lower than cluster 1 and 39.07% in the case of CV, which is lower than cluster 1 but not that low. The badge metrics show that ITB and CAB have very low average values (1.94 and 3.20) but IRB is even higher than in the in cluster 1. Finally the TEB indicator shows a moderate value. Therefore, this cluster concentrates students who have very low ITB and CAB and therefore they are not doing an organized effort towards achieving badges, but the very high IRB value demonstrates that they are very eager to earn those repetitive badges, thus they are interested in the badge system. These students have invested a moderate amount of time but they have not achieved a great progress, additionally they have shown low ITB and CAB but the highest average IRB indicator of all clusters.

The interpretation of these results will be further discussed in next section. Additionally, we use a high dimensional data visualization denominated as parallel coordinates to represent the tendency of each one of the students in the different clusters. Parallel coordinates represents consecutive parallel axes in which an *n*-dimensional point will be represented as a polyline with vertices on the parallel axes. This type of visualization can be used to detect 2D patterns and it is often use for clustering purposes. Figure 5 shows the visualization where each polyline represents a student characterized by his/her badge indicators, and the vertical facet represents clusters 1, 2 and 3 in that order starting from above. The representation shows students that belong to clusters 2 and 3 are very alike as there are almost no outliers. Cluster 1 is composed of a more diverse source of students but all of them have shown an interest in the gamification indicators. As we can see the data representation is very similar as the discussion we offered previously, and the students belonging to each cluster are very similar, thus we think that the results and groups from the clustering analysis are good.

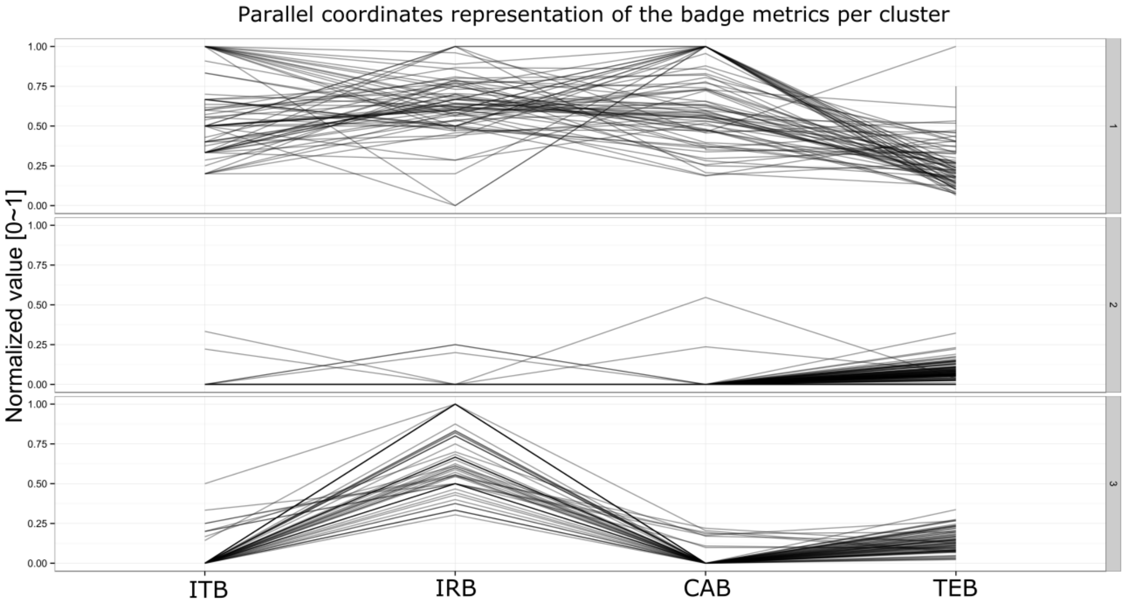

Figure 5. Parallel coordinates visualization of the badge metrics per student and separated by cluster.

5. Discussion of implications for engineering education

Engineering education presents several issues regarding engagement and motivation. Engineering degrees are often regarded as difficult and students in their youth do not really feel excited towards following an engineering career. A report by ASQ informed that the overwhelming amount of 85% of kids said that they were not interested in following an engineering career [35]. Moreover, it was found that both the expectancy- and value-related motivational beliefs of students decreased over the first year of an engineering degree [2]. The introduction of game-thinking and game elements within educational experience can increase the motivation and engagement of students to participate in a course, as shown in an artificial intelligence course for computer science engineers [36]. There are numerous approaches that seek to improve motivation and awareness of engineering classes both at precollege and college level, however, creating a good technology-enhanced learning experience is not trivial and requires a very good design and to have the correct alignment with the course [37]. As an example we did find in our study students that their behavior indicated that they were motivated by gamification features and others who were not. Thus it is important to be able to discern students who enjoy using these features and those who do not, so that the learning experience can be adapted to achieve the final goal, which is improving the motivation and interest in learning of all students within an engineering course.

The correlational analysis shows that the different gamification indicators are highly related although their definitions are different. For example, it provides the level of relation of a student who has intention to obtain *topic badges* has also of obtaining *repetitive badges*, the concentration or the efficiency towards them. We also found that the badge indicators are strongly correlated with those indicators related to the use of the platform and also to those about the progress in the platform. On the one hand, the most natural line of thinking is that making activities in the platform leads to achieve badges if they are done according to the badge requirements, therefore this relationship is quite obvious; on the other hand these results can also suggest that students who are motivated towards badges improve their engagement and participation with the platform, and that would eventually lead to more time and interaction with the learning resources, which could imply that students would also learn better the contents. This hypothesis would suggest that the use of badges can imply an increment of participation and interaction with the platform, for those engineering students that feel motivated towards the use of badges. Additionally we saw how the VAC and CV indicators were correlated also with badge metrics, although the specific definitions of these badge metrics do not take into account at all interaction with videos. Despite these correlations are low, we speculate they exist due to the fact that those users who completed a lot of videos probably show an engagement with the platform and this engagement can also be in relationship with their interest on gamification. The very high correlation between ITB and CAB can indicate that students who are really attempting to learn topics are concentrating in this task, as this would be indeed consider as good behavior which would be worth encouraging, e.g. by providing in-time feedback on how they are advancing towards a learning goal and achieve a badge in order to increment their motivation.

We have argued why it is important to increase motivation for engineering students and how gamification can help do that. However some students might feel very motivated and others will not, that is why we have characterized students depending on their badge indicators. We also reported on the findings of other works where not all the students felt motivated, but quite the opposite, where some students would reject this new methodology and have a negative effect on learning [27][21][22]. Here comes the need of knowing which students can get benefits from gamification and who will not instead of forcing all students to go by the same path. Moreover, in subsection 4.2 we showed how students with a similar interaction with the platform, can clearly have different badge indicators, which helps us know if the student is interested in badges, which type of badges or if he/she is concentrated in earning them. This information can be useful to customize learning experiences to the preferences of each student. This would be particularly interesting in engineering education to improve the motivation of students, where virtual learning environments could use this type of profile and metrics to boost gamification aspects of those who are really interested, or recommend it to students who are not really using it.

In addition to customizable learning environments for students, we can also be interested in clustering students with similar preferences for group formation purposes, i.e. to make working groups or classes setting up together students that might feel motivated by the same features. Usually in engineering courses there are plenty of practice sessions where groups are often assigned either by decision of students or by random assignment, but not on more objective criteria that can help students working together benefit from similar preferences. We presented results after performing a Two-Step Cluster analysis by clustering students according to their badge indicators. We were able to see three clusters of students who had a similar behavior while interacting with the platform. We can use these three clusters to divide them into different classes or groups and adapt the learning pedagogy of each one. For example, students within cluster 2 did not show a high motivation using Khan Academy platform and its learning concept by exercises and videos, and they were not very interested in earning badges. Therefore, these students could perform better if a more traditional learning approach is applied. Students within cluster 1, have found interest in the Khan Academy format with online videos and exercises, and their badge indicators are also high. Therefore, this class will probably perform well using innovative methodologies like ‘flipping the classroom’ and if part of the learning process involves gamification elements. Finally, cluster 3 is composed of students who accomplished a moderate progress, but they have shown a very important motivation towards earning repetitive badges. Therefore instructors teaching this class could add some gamification elements in order to motivate students in their learning process. Additionally, we also performed a cross tabulation between the clusters that have been obtained and the gender and course of students. This categorical variable analysis reveals if there exists any association between these variables. The Pearson Chi-Square Test reports that there are not significant associations between the cluster of the student and his/her gender or the course they were taking. Therefore, we assume that the gender and the course did not have an important influence on how students were elected for each cluster, and that there is not significant impact of gender and the course on the badge metrics results.

One issue is the possibility that students start earning badges for the mere joy of earning them instead of learning, which is the actual goal of the course. In this case, badges become extrinsic rewards that might actually significantly and substantially undermine intrinsic motivation as found in the literature [38]. For example, previous work found that participatory badges that are eventually earned by most of the learners, had minimal to none relationship with measure of skill [27]. In our experience we described in subsection 4.2 the example of Student A, who managed to earn more than 900 badges in the experiment by earning many *repetitive badges*, which clearly shows an intentional behavior and interest in earning badges. Furthermore we were able to confirm that this type of student profile is not an exceptional singular case, but quite the opposite as the clustering algorithm detected a cluster of students who had a very high interest in *repetitive badges*, but not that much in the *topic badges* (that are more related to actual learning achievement), and this cluster of students did not accomplish to make a very good advancement in the platform. We wonder if *repetitive badges*, that can be earned repetitively by solving the same exercise correctly many times, are not having a negative impact on the intrinsic motivation of some students, since they keep solving the same exercises to earn more badges, although they already mastered that skill and they are not learning anything new. In this direction badges related to learning skills, as the ones we denominated in our research as *topic badges*, can be more connected to internal motivation [27]. That is why although we do believe that the use of gamification in engineering education can help improve participation and engagement, when implementing external awards the design should be focused on promoting intrinsic motivation where students want to do something and find meaning in their work [39], to accomplish this the key is to use the correct approach and suitable motivation techniques [23]. We should also mention that as main limitation to our conclusions is that we are attributing to the online behavior of students in our experiment, a cognitive and pedagogical interpretation that cannot be confirmed. That means that we are making an interpretation of the students' actions based on a model based on reasoning (e.g. if all the student actions on the platform are to achieve badges then we infer that the students intentionality is to get badges). But it is very difficult to know the students' real intentions since this is something internal to the students so there is some uncertainty about it. A possibility for future work is to perform a survey to students about their interest and motivation in badges to know their self-report of intentionality and compare with the proposed indicators. Regarding the external validity and generalization of ours results we think that if an experiment is replicated under similar circumstances and badges (*topic* and *repetitive badges* in the case of the intentionality behavior), similar clusters of students should be found and similar conclusions might be applied. We note out two important characteristics of the experiment; the first one is that most students were around 17-18 years old and accessing an engineering degree, the second is that the interaction with Khan Academy was non-mandatory. If we change these characteristics, e.g. a more diverse range of demographics and/or a mandatory learning activity, we might find that different results are reached (e.g. older and busy students might not be that interested in spending extra time in the platform only to earn more badges).

With the objective of enhancing the reproducibility of this research study, we have prepared an additional report using R Markdown which can be found in the GitHub repository[3], and it includes the R code used for all the analysis and visualizations of Section 4 as well as aggregate statistics for all the students and each cluster of students separately. This way the entire process, code, data, libraries and algorithms are much more transparent to the readers, and can be replicated if they are interested. Additionally, even when these indicators have been initially defined for the specifics and context of Khan Academy platform, they can be easily extended to other learning environments that use badges. For ITB indicator, practitioners should take into account *topic badges* that are delivered when a subset of predefined exercises are completed. It is important to note out that each exercise can be part of only one of the requisites among all the *topic badges*. IRB indicator should consider those badges that can be earned repetitively by students without much effort. CAB indicator should consider badges that are earned by a chain of requisites, while TEB is quite general without any restrictions. If the requirements for these four metrics are not fulfilled in other e-learning platforms, they should be redefined although the general idea from them can be taken.

6. Conclusions

The work conducted in this paper has been focused on analyzing the interaction of students with the badge system of the Khan Academy platform in an educational experiment with three different courses. We have proposed and implemented four high level indicators about the students' interactions with the gamification system. These indicators measure the intentionality of students to earn badges (both *topic* and *repetitive* ones), the concentration or disparity to achieve them, and the time efficiency on this task. We provided correlation results between these badge indicators and other metrics related to the use of the platform and progress with exercises and videos. Finally, we used the badge metrics as inputs to apply cluster analysis finding three different clusters of students and discussing how these clusters could be used for adaptation purposes in order to enhance the learning experience of students.

As future work several approaches can be followed. For once we would like to extend some of our models to detect intentionality in different types of badges, but also test these models in different educational contexts with virtual learning environments that might have different badges. An interesting direction would be to survey students about their interest in badges during the experiment, and check for correlations with our badge metrics; this would allow us to learn if the online behavior of students is

[3] https://github.com/jruiperezv/clustering-badges/blob/master/ReportClustering.pdf

actually related to how students feel with respect the badge system. Finally, we would like to know if in the case of those students who felt motivated by badges, that motivation actually led to a better learning, and we could do that by inferring learning gains by doing a pre- and post-test before and after interacting with the learning environment.

Acknowledgements: This work has been supported by the “eMadrid” project (Regional Government of Madrid) under grant S2013/ICE-2715, the “RESET” project (Ministry of Economy and Competiveness) under grant RESET TIN2014-53199-C3-1-R, the SimLap project (Ministry of Economy and Competiveness) under grant RTC-2014-2811-1 and the European Erasmus+ SHEILA project under grant 562080-EPP-1-2015-BE-EPPKA3-PI-FORWARD.

References:

[1] S. Deterding, D. Dixon, R. Khaled, and L. E. Nacke, “Gamification : Toward a Definition,” in *ACM CHI Conference on Human Factors in Computing Systems*, 2011, pp. 12–15.

[2] B. D. Jones, M. C. Paretti, S. F. Hein, and T. W. Knott, “An Analysis of Motivation Constructs with First-Year Engineering Students: Relationships Among Expectancies, Values, Achievement, and Career Plans,” *J. Eng. Educ.*, vol. 99, no. 4, pp. 319–336, 2010.

[3] J. Douglas, E. Iversen, and C. Kalyandurg, “Engineering in the K-12 classroom: An analysis of current practices and guidelines for the future,” *A Prod. ASEE Eng. K-12*, vol. 12, pp. 1–22, 2004.

[4] T. Xie, N. Tillmann, and J. De Halleux, “Educational software engineering: Where software engineering, education, and gaming meet,” in *3rd International Workshop on Games and Software Engineering: Engineering Computer Games to Enable Positive, Progressive Change, GAS 2013*, 2013, no. Section V, pp. 36–39.

[5] K. Berkling and C. Thomas, “Gamification of a software engineering course and a detailed analysis of the factors that lead to it’s failure,” in *2013 International Conference on Interactive Collaborative Learning, ICL 2013*, 2013, pp. 525–530.

[6] A. L. Dyckhoff, V. Lukarov, A. Muslim, M. A. Chatti, and U. Schroeder, “Supporting action research with learning analytics,” in *Proceedings of the Third International Conference on Learning Analytics and Knowledge - LAK ’13*, 2013, p. 220.

[7] C. Bach, “Learning Analytics : Targeting Instruction , Curricula and Student Support,” 2010. [Online]. Available: http://www.iiis.org/CDs2010/CD2010SCI/EISTA_2010/PapersPdf/EA655ES.pdf. [Accessed: 20-May-2014].

[8] S. Arnab, R. Berta, J. Earp, S. De Freitas, M. Popescu, M. Romero, I. Stanescu, and M. Usart, “Framing the Adoption of Serious Games in Formal Education,” *Electron. J. e-Learning*, vol. 10, no. 2, pp. 159–171, 2012.

[9] M. Montola, T. Nummenmaa, A. Lucero, M. Boberg, and H. Korhonen, “Applying Game Achievement Systems to Enhance User Experience in a Photo Sharing Service,” in *Proceedings of the 13th International MindTrek Conference: Everyday Life in the Ubiquitous Era*, 2009, vol. 4, no. 46, pp. 94–97.

[10] V. Magaña and M. Organero, “The Impact of Using Gamification on the Eco-driving Learning,” in *Ambient Intelligence - Software and Applications SE - 5*, vol. 291, C. Ramos, P. Novais, C. E. Nihan, and J. M. Corchado Rodríguez, Eds. Springer International Publishing, 2014, pp. 45–52.

[11] W. Li, T. Grossman, and G. Fitzmaurice, “GamiCAD : A Gamified Tutorial System For First Time AutoCAD Users,” in *25th annual ACM symposium on User interface software and technology*, 2012, pp. 103–112.

[12] A. Domínguez, J. Saenz-de-Navarrete, L. de-Marcos, L. Fernández-Sanz, C. Pagés, and J.-J. Martínez-Herráiz, “Gamifying learning experiences: Practical implications and outcomes,” *Comput. Educ.*, vol. 63, pp. 380–392, Apr. 2013.

[13] J. Hamari, J. Koivisto, and H. Sarsa, “Does Gamification Work? -- A Literature Review of Empirical Studies on Gamification,” in *2014 47th Hawaii International Conference on System Sciences*, 2014, pp. 3025–3034.

[14] J. A. Ruipérez-Valiente, P. J. Muñoz-Merino, D. Leony, and C. Delgado Kloos, “ALAS-KA: A

learning analytics extension for better understanding the learning process in the Khan Academy platform," *Comput. Human Behav.*, vol. 47, no. Learning Analytics, Educational Data Mining and data-driven Educational Decision Making, pp. 139–148, 2015.

[15] H. Zhang, K. Almeroth, A. Knight, M. Bulger, and R. Mayer, "Moodog: Tracking students' online learning activities," in *World Conference on Educational Multimedia, Hypermedia and Telecommunications*, 2007, pp. 4415–4422.

[16] Á. Serrano-Laguna, J. Torrente, P. Moreno-Ger, and B. Fernández-Manjón, "Tracing a Little for Big Improvements: Application of Learning Analytics and Videogames for Student Assessment," *Procedia Comput. Sci.*, vol. 15, pp. 203–209, Jan. 2012.

[17] M. Minović and M. Milovanović, "Real-time learning analytics in educational games," in *First International Conference on Technological Ecosystem for Enhancing Multiculturality*, 2013, pp. 245–251.

[18] M. Kebritchi, A. Hirumi, and H. Bai, "The effects of modern mathematics computer games on mathematics achievement and class motivation," *Comput. Educ.*, vol. 55, no. 2, pp. 427–443, 2010.

[19] S. Mladenović, D. Krpan, and M. Mladenović, "Using Games to Help Novices Embrace Programming: From Elementary to Higher Education," *Int. J. Eng. Educ.*, vol. 32, no. 1, pp. 521–531, 2016.

[20] S. Kim, "Effects of the Gamified Class in Engineering Education Environments," *Converg. Inf. Technol.*, vol. 8, no. 13, pp. 253–260, 2013.

[21] G. Barata, S. Gama, J. Jorge, and D. Goncalves, "Engaging Engineering Students with Gamification," in *5th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES)*, 2013, pp. 1–8.

[22] G. Barata, S. Gama, J. Jorge, and D. Gonçalves, "So Fun It Hurts–Gamifying an Engineering Course," in *Foundations of Augmented Cognition*, 2013, pp. 639–648.

[23] M. Zirk, "Gamification for Software Engineering Education," 2014.

[24] P. J. Muñoz-Merino, M. Fernández Molina, M. Muñoz-Organero, and C. Delgado Kloos, "An adaptive and innovative question-driven competition-based intelligent tutoring system for learning," *Expert Syst. Appl.*, vol. 39, no. 8, pp. 6932–6948, 2012.

[25] P. J. Muñoz-Merino, M. Fernández Molina, M. Muñoz-Organero, and C. Delgado Kloos, "Motivation and emotions in competition systems for education: An empirical study," *IEEE Trans. Educ.*, vol. 57, no. 3, pp. 182–187, 2014.

[26] O. Borras-Gene, M. Martinez-Nuñez, and Á. Fidalgo-Blanco, "New Challenges for the Motivation and Learning in Engineering Education Using Gamification in MOOC," *Int. J. Eng. Educ.*, vol. 32, no. 1, pp. 501–512, 2016.

[27] S. Abramovich, C. Schunn, and R. M. Higashi, "Are badges useful in education?: it depends upon the type of badge and expertise of learner," *Educ. Technol. Res. Dev.*, vol. 61, no. 2, pp. 217–232, Mar. 2013.

[28] L. Hakulinen, T. Auvinen, and A. Korhonen, "Empirical Study on the Effect of Achievement Badges in TRAKLA2 Online Learning Environment," in *Learning and Teaching in Computing and Engineering*, 2013, pp. 47–54.

[29] J. Tsui and S. Edtech, "Septris and SICKO: Implementing and Using Learning Analytics and Gamification in Medical Education," *Educ. Rev.*, no. March, pp. 1–7, 2014.

[30] A. Anderson, D. Huttenlocher, and J. Kleinberg, "Steering User Behavior with Badges," in *22nd international conference on World Wide Web*, 2013, pp. 95–106.

[31] S. Grant and B. Betts, "Encouraging User Behaviour with Achievements : An Empirical Study," in *10th Working Conference on Mining Software Repositories*, 2013, pp. 65–68.

[32] J. A. Ruipérez-Valiente, P. J. Muñoz-Merino, and C. Delgado Kloos, "Analyzing Students ' Intentionality towards Badges within a Case Study using Khan Academy," in *The 6th*

Internacional Learning Analytics & Knowledge Conference, 2016.

[33] J. A. Ruipérez-Valiente, P. J. Muñoz-Merino, C. Delgado Kloos, K. Niemann, and M. Scheffel, "Do Optional Activities Matter in Virtual Learning Environments?," in *Ninth European Conference on Technology Enhanced Learning*, 2014, pp. 331–344.

[34] P. J. Muñoz-Merino, J. A. Ruipérez Valiente, and C. D. Kloos, "Inferring higher level learning information from low level data for the Khan Academy platform," in *Proceedings of the Third International Conference on Learning Analytics and Knowledge - LAK '13*, 2013, pp. 112–116.

[35] American Society for Quality, "Engineering Image Problem Could Fuel Shortage, ASQ Survey: Career Not on Radar for Kids or Parents," 2009.

[36] R. Mas-Sansó and C. Manresa-Yee, "Gamifying an Artificial Intelligence Course in Engineering Education," *Int. J. Eng. Educ.*, vol. 32, no. 1, pp. 513–520, 2016.

[37] M. Riojas, S. Lysecky, and J. Rozenblit, "Educational Technologies for Precollege Engineering Education," *IEEE Trans. Learn. Technol.*, vol. 5, no. 1, pp. 20–37, 2012.

[38] E. L. Deci, R. Koestner, and R. M. Ryan, "Extrinsic Rewards and Intrinsic Motivation in Education: Reconsidered Once Again," *Rev. Educ. Res. Spring*, vol. 71, no. 1, pp. 1–27, 2001.

[39] T. Seifert, "Understanding student motivation," *Educ. Res.*, vol. 46, no. 2, pp. 137–149, 2004.

José A. Ruipérez-Valiente completed his BSc and MSc in telecommunications engineering at Universidad Católica San Antonio de Murcia and Universidad Carlos III de Madrid (UC3M) respectively. He has worked both in the private (Accenture and Group Multimedia Vocento) and public sector (UC3M). Right now he is a PhD candidate at UC3M and Research Assistant at Institute IMDEA Networks. He has completed two research visits of three months each, the first one at Massachusetts Institute of Technology and the second one at the University of Edinburgh. As a student and researcher he has received several awards. He has currently published more than 20 works in important journals and international conferences related to his main fields of research. His research is focused nowadays in learning analytics and educational data mining.

Pedro J. Muñoz-Merino received his Telecommunication Engineering degree in 2003 from the Polytechnic University of Valencia, and his PhD in Telematics Engineering in 2009 from the Universidad Carlos III de Madrid. He is a Visitant Associate Professor at the Universidad Carlos III de Madrid. He has done two long research stays: one in Ireland for more than 3 months at the Intel company in 2005, and another in Germany for more than 6 months at the Fraunhofer Institute of Technology in 2009-2010. He is author of more than 70 scientific publications and has participated in more than 20 research projects. He has been PC member of different conferences and invited as a speaker in different events in topics related to learning analytics and educational data mining. He is also an IEEE Senior Member from 2015.

Carlos Delgado Kloos received the Ph.D. degree in Computer Science from the Technical University of Munich and in Telecommunications Engineering from the Technical University of Madrid. Since 1996, he is Full Professor of Telematics Engineering at the Universidad Carlos III de Madrid, where he is the Director of the online Master's program on "Management and Prodution of e-Learning", Holder of the UNESCO Chair on "Scalable Digital Education for All" and of the GAST research group. He is also Vice President for Strategy and Digital Education. He coordinates the eMadrid network on Educational Technology in the Region of Madrid. He is an IEEE Senior Member. His main interests are centered on educational technologies.

List of tables and figures: