The Interplay of Different English Proficiency Levels and Gamification in Online English Learning Environment

Liping Yang (lxy950239@miami.edu)
Nam Ju Kim (namju.kim@miami.edu)

Abstract: As educational technology becomes increasingly prevalent in a variety of learning environments, gamification has emerged as a frequently discussed method of increasing student interest, engagement, and learning outcomes. This research will examine the relationship between learners' English language competence levels and their gamification involvement using data from one of the most prominent online English learning platforms, controlling for demographic characteristics. The results revealed that students who are at an entrance level of English proficiency were more motivated to participate in the competition of earning more badges and points and participating on leaderboards as their proficiency increases. However, if the students reach a level greater than 6 out of 8, their interest in the activities reflecting gamification decreases. This might be due to the increasing complexity of assignments, which may become too distracting from acquiring the language skill itself. In addition, gamification as it is currently used in research studies is quite an effective method for learning, despite that factors that contribute to high success in gamification are still somewhat unresolved, in particular, for the cognitive learning results.

Introduction
As educational technology becomes increasingly prevalent in a variety of learning environments, gamification has emerged as a frequently discussed method of increasing student interest, engagement, and learning outcomes (Landers, 2014). Gamification has been shown to have a significant influence on students' motivation and engagement, which in turn enhances their overall learning outcomes. While some study has examined the effect of gender, age, and language on gamification participation, the effect of learners' skill level has received less attention. The study by Sun-Lin et al. (2019) showed that the gamification test group performed significantly better than the control group and comparison group. This research will examine the relationship between learners' English language competence levels and their gamification involvement using data from one of the most prominent online English learning platforms having the majority of students from Latin America, while controlling for demographic characteristics. The website includes tens of thousands active members of varying ages and ethnicities, providing us with a significant sample size for analysis. We see that users at a higher level are more engaged in gamification activities. Nonetheless, after a user attains a particular degree of expertise, their gamification activities decrease.
Gamification is quite an effective tool for beginner learners to move up to speed to a certain level of mastery of a subject. Despite the fact that many studies have demonstrated the effects of gamification without the specific theoretical foundation (Hamari et al. 2014; Seaborn and Fels, 2015), many researchers have tried to explain the significant relationship between the
variables gamification and learning via provisions of frameworks such as the theory of gamified learning (Landers, 2014). The theory explicitly puts forth the components: game characteristics, behaviors, and attitudes, and learning outcomes. The theory proposes that the instructional content directly impacts the learning results as well as the behavior of learners. As gamification is usually not used to eradicate instruction, but rather it sets to improve on instruction and thus produce higher quality. Instructional content is a stepping stone to have successful gamification (Landers, 2014). The main motive of gamification is to directly impact the results and attitudes of individuals towards learning (Landers, 2014). These impacts affect the relationship between the instructional content and learning outcomes by both moderation or mediation, regarding the nature of the behaviors and attitudes that targeted by gamification (Landers, 2014).

Research context

**Online English Learning.** Due to the COVID-19 outbreak, online classes have started in earnest in English subjects as well. The characteristics of the English subject, which are relatively high in private tutoring and mainly taught by dedicated teachers, clearly reveal the advantages and disadvantages of online classes compared to other subjects (Famularsih, 2020). First, the advantage of online English classes is that students can take classes at their own pace from anywhere. In addition, since there are many contents available to learners in the EFL environment in online classes, students can use more diverse materials to conduct classes (Fitria, 2020). On the other hand, the disadvantages of online English classes were also varied. Disadvantages of online English classes include lack of teacher feedback, polarization of academic achievement, difficulties in reflecting individual learner characteristics, difficulties in understanding students’ academic achievement, and difficulties in interaction between students and students and between students and teachers (Hsu, 2008). In addition, as English classes, which were conducted through various interactive activities in the classroom, are conducted online, students' interest, satisfaction, and participation in classes decrease, which can have a significant impact on students' English acquisition. Gamification can be regarded as one of the solutions to address the above-mentioned issue in online English learning.

**Gamification.**

It is scientifically described as the use of gaming aspects to non-game circumstances. Levels, points, badges, leaderboards, and avatars are among the most often used game features. Combat, content unlocking, gifting, boss battles, quests, social graphs, certifications, and memes are just a few of the additional methods accessible in gamified platforms. There are a few gamification advantages to consider (design elements, known from video games). The most reported gamification components were points, which were often used as a foundation for other features. Often, points and leaderboards were used to encourage competitiveness among players. For time restrictions, for example, they were utilized as a foundation for determining the level of crowdsources in a level with the ability for comparisons between team members and peers, as well as with badges or missions to visualize particular goals in
different implementations of the system. Nebel (2017) concluded that leaderboard could promote competition, while it had limited positive impact on motivation and effort. In this research, we will look at the gamification features of badges, leaderboard, students’ weeks on leaderboard and their points. In the recent times due to the overall growth in all areas we as a people we have become more technology reliant. Thus, the need to introduce new learning methodologies using the available tools. At first it was not as indulging so research has been in the forefront to bring out new ideas that could increase the rate at which students grasp content and in a fun way in the process. There was the idea about gamification born. Gamification of learning ways has been one of the game changing improvements over the years that has been seen to enhance the learning capabilities of students. However, as there is not yet sufficient evidence on how this enhancement is made, it might be important to see what kind of drastic changes occur in the given data of our interest with the level attained by individuals in the gamified learning platform. This is an upcoming technology whose value is increasing in popularity given what it promises to offer to individuals who are interested in learning something. Whether it’s pursuing a whole new subject or growing upon what they already know. We can say that gamification is quite an innovative a creative solution to many challenges that present themselves when one is dealing with other forms of delivery of content. Students like to compete; they compete in everything that can warrant approval from others given any circumstance.

Platform introduction
This Open English learning platform offered 24/7 learning activities and classes, with trained North American instructors available at all times. Courses and seminars in real time; practice activities. These eight levels are intended for new students who have successfully completed a placement test. Additionally, the course's gamification provides students with "miles", "badges", and “leaderboards”. Each week, fifty students are assigned to fifty groups. Each week, the top 50 learners of the classes are listed. Leagues are formed by the top 10 students. Each week, the leaderboards are updated, and new groups are formed. Leaderboard badges are not a badge type. There are badges for weekly events, performances, and special occasions. If a student completes ten tasks this week, they earn a badge. To earn a performance badge, a student must complete prescribed requirements. The Halloween badge is comparable to the holiday badge.

Literature review
Theoretical framework
As a single topic of interest usually has many implications the theoretical framework acts as the ‘blueprint’ or rather pattern for research (Grant & Osanloo, 2014). Our framework is based on the existing theory in the field of learning that gamification is a likely method to enhance productivity and learning. Gamification has been identified as one of the many technologies that will revolutionize the way schools offer learning materials to students (Johnson et al., 2014) and considering the new approach that has the ability to minimize the gap between students and their teachers (Kapp, 2017; Oblinger, 2004). The studies on
gamification usually stress on the appropriate and strategic use of games so as to have the most positive learning experience and engagement with the learners.

Striking a balance between the game and the quality of learning material is of the utmost importance as games alone are not sufficient to enhance performance neither is crude content. Therefore, one must balance these two components to have optimal outcome of the desired goal. (Apostol et al., op. cit.). Some experts have raised some concerns regarding the limitations of the method. But many have considered the arguments as trivial impediments that can be solved diligently by performing more research in the problem statement.

In this paper we will focus on the learning of the English language at different proficiency levels. We research how students behave at different levels of the learning stages in the gamified structure of study. And generate some insights if at all the level of an individual is significant in their motivation to move to the next level as they earn points and badges in the process.

Methodology
Data
Participants and Online English Learning Platform.
A total of 26,191 learners from 103 countries participated in a commercial online English program (Open English: www.openenglish.com), which includes several features that qualify as CSCL. The age range of the participants was very diverse, from 18 to late 60s. The features and functions in Open English (OE) are described below:

1. OE Provides individualized and structured course arrangements based on learners’ language levels. Before learners start their courses, they are required to take an adaptive placement test, then they are allocated to a level (from level 1 to level 8) according to their English skills. Learners at each level are required to participate in a minimum number of each of the different activities, but are not limited in the number of any activities they prefer.

2. OE provides a host of interactive multimedia content (i.e., interactive videos), which is designed to make the lessons fun and more attractive for learners.

3. Multi-platform modules (unit lesson, and practices), various English tools (dictionary, pronunciation, writing tool, grammar guide, translator), functions and techniques (gamification, standardized test prep, practice lessons, interactive videos) are designed to meet learners varying at baseline language knowledge and proficiency levels and their diverse learning needs. Within each module, a wide range of authentic topics related to business, news, arts, sport, everyday life, etc. in various contexts, are incorporated.

We selected 63,221 learners with gamification activity records as our sample from data provided by Open English. To compare the group differences in gamification performance, a multivariate analysis of covariance was used.

Independent variable (N= 1): Proficiency Level
Dependent variables (N= 4): Practice Points, Leaderboard Points, Badge Earned, Weeks on Leaderboard
Covariates (N=4): age, gender, score, mother language

1. Gender: 0 – female, 1 – male
2. Mother Language: 0 – Spanish, 1 – Portuguese

Variable definitions

Practice: The points the student has accumulated (for the students who currently not on the leaderboard)

Badge Earned: How many times the badge has been won by the student

Leaderboard Points: How many points they've earned from activities. It is the sum of points student has obtained from different leaderboards. Points here only applies to students on the leaderboard (not all students are on leaderboard, there is no overlap with Leaderboard Points)

Weeks on Leaderboard: It reflects how many weeks that student was on a leaderboard

Proficiency Level: It shows the highest English proficiency level this user has reached measured by the test score of the platform.

Research Question:
What was the main effect of users’ current learning level on this learning platform on the gamification activity, controlling the effects of their demographic information such as age, gender, score, and language?

Results

The independent variables mean the user level was divided into 8 levels, level 1 had the biggest user number of 23410, and level 8 had the smallest user number of 6. Practice Points, Badge Earned, Leaderboard Points, and Weeks on Leaderboard was set as dependent variables mean the points the student has accumulated, how many times the badge has been won by the student, how many points they've earned from activities, and how many weeks that student was on a leaderboard, respectively. From the data we have levels 1 to 8 each having a certain value of number of students at each. It is represented in a pie chart in figure 1 below:

![Pie chart of students’ levels](image)

Figure 1. Pie chart of students’ levels
The mean and standard deviation for Practice Points, Badge Earned, Leaderboard Points, Weeks on Leaderboard under 8 levels. The maximum value of Practice Points from 21 users was 5153.00 in level 6, and the minimum value from 23410 users was 1260.22 in level 1. The maximum value of Badge Earned from 21 users was 1096.14 in level 6, and the minimum value from 23410 users was 219.34 in level 1. The maximum value of Leaderboard Points from 21 users was 153305.52 in level 6, and the minimum value from 23410 users was 26470.43 in level 1. The maximum value of Weeks on Leaderboard from 6 users was 3628.83 in level 8, and the minimum value from 23410 users was 329.56 in level 1. It’s understandable that the users in level showed lowest Practice Points, Badge Earned, Leaderboard Points, Weeks on Leaderboard, but the surprising aspect of the data was the level 6 users seems more active than high level users.

We compared the difference in Practice Points, Badge Earned, Leaderboard Points, and Weeks on Leaderboard under different Proficiency Level using pairwise comparison. The Practice Points in Proficiency Level 1 was significantly higher than Proficiency Level 3, 4, and 5 ($p <0.001$), and significantly lower than Proficiency Level 6 ($p <0.001$). The Practice Points in Proficiency Level 2 was significantly higher than Proficiency Level 3, 4, and 5 ($p <0.001$), and significantly lower than Proficiency Level 6 ($p <0.001$). The Practice Points in Proficiency Level 3 was significantly lower than Proficiency Level 1, 2, 6, and 8 ($p <0.001$). The Practice Points in Proficiency Level 4 was significantly lower than Proficiency Level 1, 2, 6, and 8 ($p <0.001$). The Practice Points in Proficiency Level 5 was significantly lower than Proficiency Level 1, 2, and 6 ($p <0.001$). The Practice Points in Proficiency Level 6 was significantly higher than Proficiency Level 1, 2, 3, 4, and 5 ($p <0.001$). There was no significant difference in Practice Points under different Proficiency Level. The Practice Points in Proficiency Level 1 was significantly higher than Proficiency Level 2, 3, 4, and 5 ($p <0.001$). When the significant value was lower than 0.001, and the mean difference (I-J) was positive, which can be considered that the value (I) was significantly higher than value (J). For the detailed and full results of this pairwise comparison, please refer to Appendix A.

To test which is combination of variables performs the best out of all the possible combinations, a multivariate analysis of covariance (MANCOVA) was conducted with Proficiency Level as an independent variable, Practice Points, Badge Earned, Leaderboard Points, Weeks on Leaderboard as dependent variables. After controlling the influence of covariates, significant main effects (Proficiency Level) were found (Wilk’s $\Lambda = 0.99$), $F =11.91, p <0.001, \eta_p^2 =0.002$. These results illustrate that Practice Points, Badge Earned, Leaderboard Points, and Weeks on Leaderboard have significant difference under different Proficiency Level. The $p$ values were significant (<0.001) in all covariates, illustrate that use age, gender, score, and language as covariates were correct.

<table>
<thead>
<tr>
<th>Effect</th>
<th>Value</th>
<th>F</th>
<th>Hypothesis df</th>
<th>Error df</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept, Pillai's Trace</td>
<td>.002</td>
<td>31.174b</td>
<td>4.000</td>
<td>55273.000</td>
<td>&lt;.001</td>
<td>.002</td>
</tr>
<tr>
<td>Wilks' Lambda</td>
<td>.998</td>
<td>31.174b</td>
<td>4.000</td>
<td>55273.000</td>
<td>&lt;.001</td>
<td>.002</td>
</tr>
</tbody>
</table>
We performed tests of between-subjects effects to check whether the Proficiency Level would have any effect on Practice Points, Badge Earned, Leaderboard Points, and Weeks on Leaderboard, the results showed that they were significantly different on Practice Points ($p < 0.001, \eta^2_p = 0.004$), Badge Earned, ($p < 0.001, \eta^2_p = 0.003$), Leaderboard Points ($p < 0.001, \eta^2_p = 0.003$), and Weeks on Leaderboard ($p < 0.001, \eta^2_p = 0.001$) under different Proficiency Levels. The partial eta squared ($\eta^2_p$) statistic reports effect size, the individual variables in our results had a significant difference but may be do not have great effect on the value of all dependent variables.
Data interpretation

Gamification applications can be applied in many different fields, and available research has consistently failed to bring to light that there are many different game design factors that can influence results in different manners for the learners, methods of social interactions, and learning arrangements (Sailer et al. 2017a). Therefore, including different factors to cater for conceptual heterogeneity in gamification. Further, because contextual and the situational influencers may impact the effects of gamification on the learning outcomes (Hamari et al. 2014), and gamification research does lack methodological rigor (Dicheva and Dichev 2015; Dichev and Dicheva 2017), we did include situational, contextual, and methodological factors. The procedure of choosing potential moderating factors for the effects of gamification on learning results was iterative in design. We included moderating factors that were both theoretically interesting and also those that our research could manage to support their inclusion.

The data that we have has 11 columns of which 10 of them are key variables of interest. All the 10 variables are significant in the regression analysis of the data, as the p-values for each is less than 0.05 alpha level of significance thus they are significant. There is almost a total drop of users in the levels 6, 7, 8 which may be attributed to several factors. We have the additive linear model as: Design: Intercept + age + gender + score + language + test level.

Performing a pairwise analysis on the data to test if there are defects using the combination method. Most defects are usually caused by interaction of at most two variables. The data is clean there is no necessity of doing any data cleaning. The dataset is large enough to aid us in our research despite there are lesser values that represent the levels which are higher than the optimum. The data suggests that there is quite a large drop of individuals as soon as the fifth level is reached. We have close to zero individuals that push through the 6th to the 8th level. This can be caused by several factors maybe in the complexity of the tasks in the level or there is a big jump with the difficulty measure.

Discussion

Students portray a very fascinating behavior as we have come to observe in our analysis problem. Humans have a very highly competitive nature; this may be the reason why many students at the beginning of the learning process are highly motivated so that they can win points and awards to prove to others that they are better. This is a quite a fun and engaging way to enhance the intake of knowledge as at the end of the day it is highly encouraged to ingest a high amount of knowledge. That is the optimum as taking in too much is usually pointless as it corrups the learned information. As the individual skill up or level up to a certain point we can call it optimum as per our analysis it’s almost as if they do not want to take in much and their competitive nature starts to lag. Learning professionals have come to find the use of game-based learning methodologies highly effective over the years.

From what we have observed one can say that if students think that something is a challenge, they tend to shy away from it as they do not find it interesting, boring things are not usually one of their strongholds. The points that the individuals get after some hard work
works as a clear goal that they would be inclined to sweat for. By the time they reach an optimum level where most of them are bored and participate less they usually have accrued most of the valuable material as the building blocks of the relevant subject that they indulge in. Given that the complexity of the learning material is set to increase exponentially this can highly challenge students as they have quite a low threshold for challenges, such that they would rather ran as far as they could away from them rather than face the challenges head on. With the continued rise in the use of technological tools, which are most likely to take over traditional styles of content delivery, the gamification learning process is likely to be the best methodology to implement in order to have the best results.

The question we need to pose later on is how we make the individuals that have lost some key interest in gaining it back so as to have a consistent flow and have the productivity set as high as possible. Simplification of challenges in the latter levels may be one way to have a consistent competitiveness of individuals. Also maybe increasing the reward system in the challenging levels so that students can have some motivation not to lag or decrease morale in the learning process. Gamification is yet to be tapped fully to its full capacity of potential. Given that it has only been there for a while and it has proven to work quite effectively it can be researched upon broadly why indeed these characters drop consistency or rather their competitive potential. Knowing these exact reasons can open up a whole new level of productivity not only in learning the English language but also extended to other fields as the notion can be applied across. Social interactions are also implied while undertaking this gamification exercises. When individuals meet other individual on a setting whether physically or virtually, they usually form key bonds. The result for behavioral learning outcomes is evident from the indulging of games, showing that combinations of competition and earning rewards through collaboration in games are promising for learning (Clark et al. 2016).

Limitations
As a result, one drawback of our study is that the sample size for higher level students was quite small, particularly for the behavioral learning outcomes and all subsplit analyses. This limits the generalizability of the results and is also problematic for statistical power because, for random effects models, power is dependent on the total number of participants across all studies as well as the number of primary studies. This limits the generalizability of the results and is also problematic for statistical power (Borenstein et al. 2009). When there is a significant between-study variation, as there was in this meta-analysis, power is likely to be insufficient. This means that statistically insignificant results do not always indicate the absence of an impact; instead, they might be explained by a lack of statistical power, which is especially true when effects are tiny.

Conclusion
We observed that students in the beginning level of English proficiency may be more motivated to compete for badges, points, and leaderboard participation as their proficiency grows. But after level 6, students may lose interest. The rising difficulty of assignments may be distracting from learning the language itself. The study's sample size of higher-level
learners is limited. Also, students that participate in instructional activities may not be competing just for badges or points. Maybe people reach a particular degree of skill and assume that their knowledge is sufficient to use in the field. It's an intriguing outcome that many of the study's participants shared. In general, additional high-quality research is required to investigate other aspects that may influence individual productivity.

The findings indicate that gamification may be an extremely effective strategy for increasing student involvement in learning processes. Individuals are clearly motivated to perform when a reward system is in place. When other variables stay constant, the motivation to compete declines with level. Enhancing the incentive system at more difficult or later levels can assist in retaining a group of people's engagement. This gamification-based learning approach is applicable to a wide variety of context.

Individuals feel more competitive when they are in a group, and learners may share knowledge. Individuals seek to understand things in a variety of ways in order to trade their perspectives on a subject and obtain the most experience and results. This style has been proven to be the most effective, and with future study, it may become the preferred form of instruction. Dropping out after a given level could undoubtedly be improved. These might become the new standard, upending previously accepted procedures.

**Implications**

Our findings mean that we can apply more of this method to the enhancement of the way students learn. More research in the field would generate a lot more information and results that could be highly relied upon. We have found that this method of learning could be of great benefit to both teachers and students. It is quite a practical way of achieving goals. The study could be replicated easily in a setting. The methodology that is used has been explicitly stated and hence the ease. One can also use conceptual replication method to produce the results.
References


