

Duolingo XP and Time on Platform

—Modeling the User Experience of a Language Learning Application—

Jeff BRODERICK*

Abstract

Past studies have analyzed the effect of Duolingo on students' TOEIC Bridge Test Scores and found that Duolingo is not an efficient way to improve these scores. These past studies were based on assumptions concerning the relationship between students' earned Experience Points (XP) in the application, and the time spent using the application. This study aimed to explore this relationship in greater detail, drawing on further information made available to educators by updates to the Duolingo application. Results of this study indicate that XP can vary widely from user to user, depending on their starting level. A regression model is provided and past studies are reviewed, with the finding that Duolingo is more effective than previously assessed.

Keywords : Computer-aided language learning (CALL), Duolingo, smartphone applications

1. Introduction

Duolingo is one of the world's most popular language learning applications (or “apps”), with an estimated 500 million users learning 39 languages (Blanco, 2020). The app has several advantages, the main ones being its convenience and ease of use. Users can study using their smartphones at any time, in any place, and this function can lead to greater learner autonomy.

In a 2012 study, Veselinov and Grego found that adult beginner Spanish learners acquired a semester's worth of knowledge in an average of 34 hours. They also determined that student scores on a standardized WebCAPE Spanish test increased by an average of 8.1 points per hour of Duolingo use. Munday (2016) argued for the inclusion of Duolingo in university language classes, while others like Krashen (2014) questioned its utility.

I conducted a study (Broderick, 2021) to

investigate whether Duolingo was similarly effective at increasing Japanese university students' scores on the TOEIC Bridge Test (BT). The results showed that for these lower-intermediate learners, 60 minutes spent studying Duolingo resulted in an average TOEIC BT score increase of 0.121 points. This is clearly not an efficient means of increasing student TOEIC BT Scores, as the score increases as a function of time spent studying with Duolingo were minimal.

These results, however, were based on an assumption that students received 10 Duolingo Experience Points (XP) for the completion of one online lesson, each of which is designed to take approximately 10 minutes to complete. The adoption of this simple model, *10 XP for 10 minutes of study*, was necessitated by the fact that, at the time of those studies, Duolingo did not provide details about student time spent using the app. Since these studies were made,

* 理工学部共通教育群講師 Lecturer, Division of Liberal Arts, Natural, Social and Health Sciences, School of Science and Engineering

however, Duolingo has modified their platform and now provides educators with additional data including both XP and time on platform. With this data, it is possible to examine student performance in detail and create a model of average student performance.

The research questions guiding this study are therefore:

- i.) How much time did students tend to spend on the platform over the course of a 14-week semester?
- ii.) What is the model for average expected XP awarded per minute of Duolingo use?
- iii.) How does this alter the results of the previous studies?

2. Methodology and Participants

Participants for this study were 90 first-year students ($m = 75$, $f = 15$) enrolled in my first-year Communication English classes at Tokyo Denki University. They were enrolled in a variety of majors including mechanical and electrical engineering, biological science, systems engineering and design, and architecture. They were asked to study using Duolingo for 10% of their final grade, with the guideline that completing one 10-minute lesson per day, five days a week, for 12 weeks, (a total of 60 lessons, or approximately 600 minutes study) would provide them with full marks.

Upon enrolling in the free Duolingo online course, students were instructed to select the “casual” study level, which provides lessons of approximately 10-minutes duration. They were also asked to take a short in-app test, which the app designers claim can test a user’s starting ability and provide them with lessons tailored to their level.

Students were reminded several times through the semester to use Duolingo. Specifically, each

class was verbally reminded in Weeks 2, 6, 10, and 12 to keep up their daily Duolingo practice.

At the end of the 14th week, data was collected from each student concerning their allotted XP and the time that Duolingo had recorded them using the app. This data was plotted in Excel and further analyzed using the Excel Analysis ToolPak. Confidence intervals for the linear regression were calculated manually following Field (2009).

During the data analysis, several students were found to have completed none of the assignment, i.e., they received 0 XP and had 0 time on the platform. These students were removed from the study. The regression, however, was still performed with the logical assumption of a y-intercept value of 0.

Furthermore, two students were found who could be considered extreme outliers. One was a student who had obtained over 11,000 XP after using the app for 1860 minutes, or 31 hours. Another student used the app for 780 minutes (13 hours) acquiring over 8000 XP. They were both removed from the study. A third student, who obtained 4926 XP in 508 minutes, was left in the study as this result, while extreme, still lies within two standard deviations of the mean result.

3. Results

The main results of the study are presented below. First, the distribution of the independent variable, *time on platform*, is presented as a frequency histogram in Figure 1, below, with the descriptive statistics in Table 1.

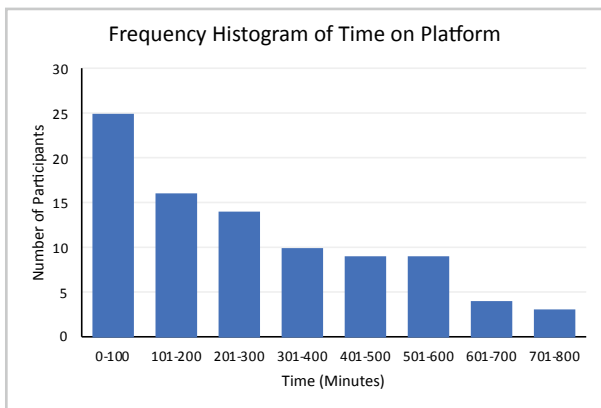


Figure 1

Table 1. Descriptive statistics

Mean	Upper 95% CI	Lower 95% CI
274.3	318.8	229.8
SD	Std. Err.	
211.2	22.38	

Next, the student data of actual XP awarded *vs.* time on platform is presented as a scatter plot in Figure 2. A linear regression was performed to model the average expected XP per minute of Duolingo usage. The lines above and below the regression line are the 95% confidence interval for this average expected value. The statistics for this regression are shown in Table 2.

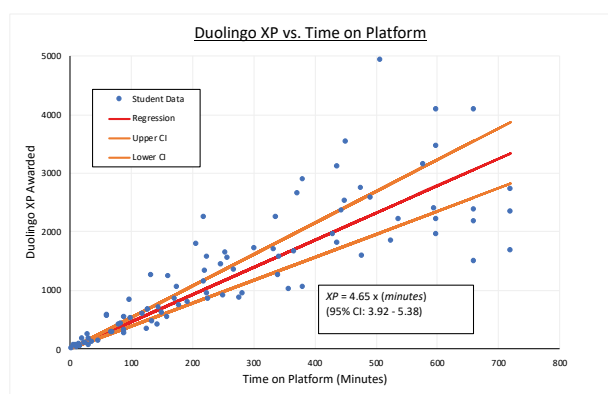


Figure 2

Table 2. Descriptive statistics

R^2	Deg. of Freedom	t -value
87.7%	88	1.987
Predicted Coefficient	Lower 95% Conf. Int.	Upper 95% Conf. Int.
4.65	3.92	5.38

4. Discussion

(i) Time on Platform

In examining the amount of time that participants spent using Duolingo, the largest single group (25 students, or almost 28% of participants) used the app for less than 100 minutes. Overall, 78 participants (almost 87%) did not complete the requirement set by the teacher to study for 600 minutes. These results suggest that, for students who do not have strong motivation to study on their own or who do not possess the self-discipline to schedule their daily study times, the potential benefit of “enhancing learner autonomy” may be missed. The participants in this study are required to take Communication English as a compulsory subject, and past research (Broderick, 2013) profiled students in a similar cohort as having relatively low motivation to study English. Duolingo may not be well matched to such learners.

(ii) Awarded XP and Time

The data for awarded XP and time on platform shows that there is a large amount of variation between users. Users studying at higher levels tend to receive higher XP rewards for completing a given lesson or level test.

A linear regression performed on the student data suggests that the new prediction equation $XP = 4.65 (\text{minutes})$ is a better model for awarded XP than that used in my previous study

(Broderick, 2021) which was $XP = (\text{minutes})$. It is still not a perfect model, considering the large amount of variation between users, but we can conclude that this is a very significant change and means that, on average, students acquire XP at a much faster rate than previously assumed.

(iii) Application to Past Studies

Taking into account these new findings and applying them to the past results means that students potentially improved their TOEIC BT scores at a rate of $0.121 \times 4.65 = 0.563$ points per hour of study. Veselinov and Grego's (2012) study claimed that students acquired a semester's worth of knowledge in an average of 34 hours of Duolingo use. Whether a TOEIC BT score increase of $0.563 \times 34 = 19$ points is equivalent to a semester's worth of knowledge is an open question, but it does warrant further research into this area.

(iv) Directions for Future Research

Future studies should utilize a control group and treatment group comprised of similar students, the former studying with traditional methods, and the latter studying using Duolingo, to more accurately assess the effects of Duolingo on TOEIC BT scores. Now that the app provides educators with actual information regarding user study times, there will be no need to utilize a model such as the one produced in the current study; any future study will be able to correlate TOEIC BT score increases with study time directly and determine the effectiveness of Duolingo.

5. Conclusions

In a past study (Broderick, 2021) I stated that Duolingo does not provide an efficient means of improving standardized test scores such as the TOEIC Bridge Test. In light of the current findings, Duolingo may be up to 5 times more effective than previously measured and further research is very much needed in this area. Future studies can utilize the more comprehensive data reporting tools that are now available to researchers using the platform, and should be able to provide much more accurate results.

Regarding actual time on platform, the majority of participants did not use Duolingo as often as directed, with the suggested level being one 10-minute lesson per day. While Duolingo may be seen as a good way to grant learners autonomy and control over their own studies, these benefits are best realized in students with high levels of motivation and self-discipline.

References

-
- Blanco, C. (2020). 2020 Duolingo Language Report: Global Overview. Retrieved Aug. 2022 from <https://blog.duolingo.com/global-language-report-2020/>
- Broderick, J. (2013). Profiling the motivation of TDU English learners. *Bulletin of Tokyo Denki University, Arts and Sciences 11*, 139-148.
- Broderick, J. (2021). Does Duolingo improve TOEIC scores? Investigating a language learning app in Japanese university classes. *OnCue Journal Special Issue Vol. 2*, 28-34.
- Field, A. (2009). *Discovering Statistics Using SPSS* (3rd edition). London: Sage.
- Krashen, S. (2014). Does Duolingo "trump" university-level language learning? *The International Journal of Foreign Language Teaching 9*(1), 13-15.
- Munday, P. (2016). The case for using Duolingo as part of the language classroom experience. *Revista Iberoamericana de Educacion a Distancia 19*(1), 83-101.
- Veselinov, R., & Grego, J. (2012). *Duolingo Effectiveness Final Report*. Manuscript available through Duolingo.