Supplemental Use of Duolingo in Language Courses: A Behavioral Engagement Perspective

Daniel A. Castañeda Vise[1*], Moon-Heum Cho[2]

[1] Kent State University, USA

[2] Syracuse University, USA

[*]Corresponding author's email: dcastane@kent.edu

[*]ORCID: https://orcid.org/0000-0002-3569-5645

Received: 03/24/2024 Revision: 07/31/2025 Accepted: 12/19/2025 Online: ..../...../2025

ABSTRACT

Few MALL studies have examined learner engagement from the behaviorist dimension using platform-generated data. Addressing this gap, the present study investigated learner engagement with the Duolingo app in both short- and long-term use, comparing face-to-face and online class modalities. Participants were 143 English-speaking college students in Elementary I and II Spanish courses who used Duolingo for 11 of 15 weeks as a supplement to classroom instruction. Engagement was measured through Duolingo for School logs (e.g., time spent, lessons completed) and timestamps. Overall results indicate higher app use in the face-to-face classes, suggesting that teacher presence and peer immediacy support engagement. Findings also show that Elementary I students' usage declined over time, while Elementary II students peaked midsemester. A small number continued using the app after the course ended. This study highlights how platform data can offer valuable and unique insights into learner engagement in MALL contexts.

Keywords: App use, Duolingo, engagement, gamification, MALL

Introduction

One of the most gamified mobile-assisted language learning (MALL) apps (Govender & Arnedo-Moreno, 2020), Duolingo is popular among millions of users around the world because of its gamified features that users can access free of charge (Duolingo, 2024). Previous studies have shown that gamification increases learner engagement (Golesorkhi & Marandi, 2025; Looyestyn et al., 2017; Sukmana et al., 2024); specifically, motivation and increased learning engagement have been cited as the most frequently reported advantages of using gamification in a foreign language class (Dehganzadeh & Dehganzadeh, 2020). These advantages can be attributed to gamification-based learning methods, which allow learners to engage physically and mentally (Ardoiz García, 2017). Research on students' engagement with digital gamification is still at its infancy. That is, peer-reviewed articles related to gamification-based

CITATION | Castañeda Vise, D.A., Cho, M-H. (2025). Supplemental Use of Duolingo in Language Courses: A Behavioral Engagement Perspective. *Computer-Assisted Language Learning Electronic Journal*, *26*(5), pp. 1-21

learning in the field of foreign language learning first appeared in 2015, with most of the studies focusing on the English language, vocabulary content, and the Duolingo app as one of the most used in the reviewed articles (Dehghanzadeh & Dehghanzadeh, 2020). The trend toward the use of Duolingo in the classroom has increased globally not only because of the 2020 pandemic but also practitioners' pursuit of ways to engage students online (Azar & Iskansar Tan, 2020; see also meta-analysis by Shortt et al., 2021).

In addition, contextual factors, including class modality (e.g., in person vs. remote), which could potentially impact students' overall engagement with diverse forms of technologies, require further clarification. For instance, McKellar and Wang (2021) reported that students' academic engagement and sense of connectedness may be heightened when teachers and students share the same physical space. Gómez et al. (2011) also reported that certain students participating in online classes may perceive these courses as less challenging and may expect to dedicate less time on coursework. Given that class modality may impact students' overall engagement, a comparative approach (face-to-face vs. online) was used in the current research to account for the potential differences (e.g., learners' attitudes) that exist in each class modality.

Literature review

Behavioral Engagement in Gamified Applications

Notably, the term "learner or student engagement" has not been clearly operationalized in current research related to student engagement with learning technologies (e.g., social media, online games, and mobile phones), making identification of the degree of similarity with related concepts like motivation difficult (Nkomo et al., 2021). For instance, in technology-mediated learning, "engagement" has been operationalized with indicators that include participation, attendance, assignments completed, time logged, and other on-task behaviors (Henrie et al., 2015). Despite the lack of consensus on what constitutes learner engagement, scholars have conceptualized this construct in terms of three major themes or dimensions: (a) behavioral engagement (e.g., completing assignments), (b) cognitive engagement (e.g., students' evaluation of class content), and (c) emotional engagement (e.g., showing interest or boredom; Bond et al., 2020; Fredricks et al., 2004; Heilporn et al., 2022; McKellar & Wang, 2023; Nkomo et al., 2021).

Also, in technology-mediated learning experiences, no consensus has been reached on how behavioral engagement can be most appropriately measured. Measures can range from self-reported surveys to interviews to assessment scores and behavior counts. In addition, in the case of surveys, the constructs developed to measure engagement may vary, with some containing only one or two items related to engagement while others contain a full battery of questions related to this construct. The differences in measurement variations may lead to conflicting findings between two studies linking student engagement to positive outcomes due to varying definitions or construct conceptualizations (Henrie et al., 2015).

For instance, some researchers like Bacca-Acosta and Garzon (2020) have collected data by combining an automatic monitoring system in the app with self-reported measures. The overall results indicated that participants with higher positive perceptions regarding effort, perceived usefulness, and behavioral intention engaged in longer use of mobile-based assessment systems.

Al-Bogami and Elyas (2020) used a 5-point Likert-type questionnaires and classroom observational data to examine learners' mobile app engagement in an English as a foreign language (EFL) class, which resulted in the students' positive perception of the use of the apps. Similarly, Huang et al. (2024) used a questionnaire analysis and qualitative semistructured interviews to explore how EFL students with lower and upper proficiency levels (LP and UP) behaviorally engaged in their self-initiated mobile-assisted English learning (MAEL). The researchers reported significant differences in both how long and how often students participated in MAEL across proficiency levels, with LP learners investing more time on a more frequent basis than UP learners.

The aforementioned studies on behavioral engagement regarding the use of digital technologies reveal a prevalent reliance on self-reported perception data rather than other data types (See also a systematic review by Nkomo et al., 2021), showing only one end of the spectrum. Additionally, the data collection methods varied across studies in terms of the variables used to measure student engagement. In the field of MALL, at the time of this writing, very few researchers had incorporated and analyzed platform-generated data in terms of the behavioral dimension. For instance, García Botero et al. (2018, 2019) have used some features of the Duolingo for Schools platform tracking system to explore app use with Spanish speakers learning either French or English as a foreign language. (See more details on these studies in the Duolingo and Gamification section).

Given the rising prominence of mobile apps in language acquisition and rapid and continuous technological improvements, more research on app learner engagement from a variety of perspectives and dimensions, using different metrics, and across different populations and varieties of languages is needed to inform teachers and MALL developers how to improve the language-learning process. Our intent in the current study was to contribute to the discussion by investigating the extent to which learners engage in app use (i.e., short and long term) in two class modalities—face-to-face vs. online—when it was used as a supplement to regular classroom instruction.

Definition of Terms

In the digital learning context, students' behavioral engagement is often identified by computer-recorded indicators, such as assignments completed, frequency of logins to websites, and number of postings, responses, and views as well as the time spent in these task-related behaviors (Henrie et al., 2015; Huang et al., 2024). In this study, the term "learner engagement," therefore, refers to the students' app use (e.g., days active on the app, time spent on the app, experience points [XP], and lessons completed) as tracked on the Duolingo for Schools platform (see Research Instrument section for additional information). The expression "short term" refers to one semester of study, whereas "long term" refers to an indefinite period of app use following course completion.

Duolingo and Gamification in MALL: Motivation and Engagement

In the field of MALL, Duolingo is the most widely used digital gamified platform for learning a foreign language (Dehganzadeh & Dehganzadeh, 2020). Research on its gamification features and its potential to engage learners (e.g., in platform use) is limited but evolving. Current

research in this field has indicated that the app features (e.g., gamification) may not be sufficient to foster app use per se; in fact, the need for guidance or support (e.g., scaffolding) from the instructor to foster engagement in using the app has been reported. In particular, García Botero et al. (2019) conducted a study with students of French as a foreign language, inviting two experimental groups to engage in voluntary learning through Duolingo outside class. One of the two experimental groups was trained for self-regulation and received scaffolding for their MALL. The control group followed regular class instruction. Results of their study revealed that students who were trained in self-regulation and received temporary scaffolding participated in Duolingo at a rate significantly higher than the other groups. The researchers also reported a high correlation between extensive use of Duolingo and improvement in French writing skills. In addition, they found that inviting students to engage in voluntary MALL outside class did not necessarily translate into higher test scores.

In another study García Botero et al. (2018) tracked 118 college students to explore their use of Duolingo and to discover whether MALL was used by these students for meaningful, self-directed learning. All students were native Spanish speakers learning either English or French at varying levels of proficiency. Although most students saw value in Duolingo and enjoyed its game-like aspects, the researchers found that three factors inhibited its use: difficulties with motivation, flaws in self-monitoring, and failure of self-management. They also reported that student motivation waned over time, leading to decreased use of the app. In addition, many students either did not know about the self-monitoring features or found them time-consuming or repetitive. Finally, many students prioritized graded coursework over using Duolingo and commented that they would have preferred their teacher to at least be aware of their Duolingo use and incorporate it into their formal instruction.

Similarly, Bende (2017) investigated the effectiveness of Duolingo's gamified elements in practice to examine whether they could support EFL learners' engagement in their own learning as well as learner autonomy. According to Bende (2017) the app's gamified elements did not engage the students effectively in the long term. In other words, game elements in the app—such as leaderboards, restarting the task, and immediate feedback—may not have been sufficient to stimulate competition and increase motivation in the long run. Regarding learning autonomy, results suggested that learners need an instructor's training and support to gradually become more independent in their learning process.

The research summarized above has indicated that gamification can motivate and engage students in their use of mobile apps, which may potentially lead to language learning gains. Researchers examining these gains have considered two major populations: noninstructed learners (i.e., informal learners) and instructed learners (i.e., those using apps as supplements to classroom instruction). For instance, regarding informal learners, Vesselinov and Grego (2012) investigated the effectiveness of Duolingo in language achievement for 8 weeks. The participants were English-speaking learners of Spanish, who used the app mainly for personal reasons, business or work, and travel; most were college graduates or had advanced degrees (MA or PhD). Results from the WebCAPE Spanish Placement Test used as a pre- and posttest indicated a statistically significant improvement. The researchers estimated that a learner with no prior knowledge of Spanish would need an average of 34 hours to complete the equivalent

of one semester of college-level Spanish. Other studies with similar populations have shown that Duolingo users achieved proficiency scores comparable to those of college students (Jiang et al., 2020); however, other researchers have reported that participants found Duolingo helpful but with some mixed effects on language learning. (For instance, Loewen et al. (2019) investigated the learning experiences of participants learning Turkish on Duolingo. Results indicated that participants knew more Turkish at the end of the study; however, after 34 hours of study, only one participant received a score that would be considered a passing grade in the university's first semester Turkish course, calling into question Vesselinov and Gregos' claim regarding the app's efficacy.

Regarding instructed learners, Rachels and Rockinson-Szapkiw (2018) used a pre- and posttest design to investigate the effect of Duolingo on students' Spanish language achievement and academic efficacy for 12 weeks. The students in the control group were instructed in a traditional instructional classroom setting, whereas the quasi-experimental group was taught via Duolingo. Results of the analysis of covariance indicated no statistical difference in students' Spanish achievement or in academic self-efficacy between the control and quasi-experimental group; that is, students who used the Duolingo app achieved results similar to those receiving traditional classroom instruction. Other studies with instructed learner populations, for instance, have shown the effectiveness of Duolingo (in combination with online learning) in developing vocabulary, sentences with simple grammar, and fluency (Al Fadda & Alaudan, 2020).

Research on the efficacy of Duolingo has been called into question by some scholars. For example, Krashen (2014) pointed out that the participants in Vesselinov and Grego's (2012) study were not typical university students: Many were college graduates or had graduate degrees. The reader should also note that the participants in studies by Jiang et al. (2020) and Loewen et al. (2019) had similar educational backgrounds. As to Vesselinov and Grego's (2012) claim that a beginner would need 34 hours to do the equivalent of one semester of college Spanish, Krashen (2014) commented about the difficulty in comparing with typically instructed foreign language learners (e.g., university students), who are often not volunteers, typically less motivated, younger and less experienced, and have few or no opportunities to proceed at their own pace. Krashen's (2014) critique seemed to indicate that—at least for instructed learners—some type of guidance is needed from the instructor.

In summary, research on engagement with Duolingo remains inconclusive but is evolving. Further investigation is needed into different learning contexts (e.g., online vs. face-to-face) and learner characteristics (e.g., instructed vs. non-instructed; young vs adult learners) that may influence engagement outcomes. Additionally, diversifying research methods—such as short-term versus longitudinal studies and varied data collection tools (e.g., surveys, platform logs)—will offer deeper insights. Furthermore, including languages other than English in the research discussion would expand the scope of findings. In the current study we sought to provide instructors and researchers with insight into how instructed learners engage in a gamified platform—from the perspective of behavioral engagement—in two learning contexts (face-to-face vs. online) in the short and long terms, potentially leading to language learning gains. We did not, however, focus on the efficacy of Duolingo in this study.

Research Questions

1. Does a significant difference in app platform use exist between learners who use Duolingo in a face-to-face classroom vs. those who do so in an online synchronous environment?
2. What were the participants' monthly app use patterns during the semester (short term)?
3. To what extent do students engage in app use after course completion (long term) and for how long?

Methods

Participants

An original pool of 167 students enrolled in Elementary Spanish I and Elementary Spanish II participated in this study from fall 2019 to summer 2022. Twenty-four participants were removed from the study because they either had no activity on the platform or were considered outliers because their scores were greater than +3 standard deviations (SD) from the mean. A final pool of 143 participants was used for the current data analysis as detailed in Table 1 below. A total of 57 were men and 86 were women. All participants were enrolled on a regional university campus in northeastern Ohio, USA, where most students (all face-to-face classes) commuted to school.

Table 1

Participants (Final Pool Detailed)

Modality/group	Elementary I Group	Elementary II Group	Total participants
Face-to-face	36	31	67
Online (Synchronous)	69	7	76
Totals	$n = 105$	$n = 38$	$n = 143$ (final pool)

Duolingo Application

Duolingo is a free app available for Android and iOS operating systems; an ad-free premium subscription called Super Duolingo is also available (Duolingo, 2024). The gamification features of its bite-sized lessons are adaptable to the learners' individual needs and help them to remain motivated. Learners can test out of a level if they have some previous knowledge of the target language. Regarding flexibility, the software allows the learner to choose a daily goal; for instance, casual: 5 minutes per day (1 lesson); regular: 10 minutes per day (two lessons). Completion time may differ depending on the learner's engagement with the activity. The types of activities on the app include multiple-choice questions, translation exercises (e.g., writing from Spanish to English and vice versa), dictation activities (e.g., the learner writes the phrase or word the software dictates), and voice recognition activities. The software also provides feedback in the form of a congratulatory message displayed each time the learner enters the correct answer. Conversely, the software displays the correct word or sentence if the learner enters the incorrect response. The software is tolerant of some minor mistakes (e.g., missing

accent marks, misspellings, and lack of punctuation). It also provides grammar review materials that students may read at their convenience.

Notably, Duolingo has met with some criticism because its activities are based on behaviorist learning, translation, competition, and extrinsic rewards, which contrast with more acceptable sociocultural approaches (Lantolf et al., 2015; see also Shortt et al., 2021). In this study, however, Duolingo was used as a supplement to regular classroom instruction (see Procedure section for additional information).

Context for Activities with Duolingo

The learning goals for Duolingo activities were (a) to learn in an engaging manner extra or new content that may or may not have been covered in class and (b) to reinforce long-term retention (e.g., some app activities recycle vocabulary in multiple lessons and skills) of the language learned in the current class or in previous classes (e.g., high school). Before Duolingo activities were implemented, all students were required to download the app to their smartphones or their tablets. Instructions for joining Duolingo for Schools were also provided via the course management system. Duolingo activities were implemented as outside class assignments and accounted for 12% of the final grade. Students were required to accumulate at least 1,000 experience points (XP) by the end of the semester (starting Week 4 and ending Week 15). Additional XP (beyond 1,000) were encouraged but optional and, if completed, were added as bonus points toward the final grade. The latter extra points were given as make-up opportunities for students who missed points during the semester because of unexcused class absences or to compensate for missed homework assignments in ELEteca, the textbook platform used concurrently with Duolingo. If students accumulated more than the required XP, they were given additional bonus points (added to the overall final grade) as follows: 2% = 1,500–1,999 XP; 4% = 2,500–2,999 XP; 6% = 3,500–3,999 XP; 8% = 4,500–4,999 XP; and 10% = 5,500 or more XP. The latter was also implemented to add an extra layer of motivation to engage learners in app use. Importantly, students had the flexibility to work at their own pace and convenience regarding the number of lessons they wanted to complete per day or per week. In other words, no due dates were set, but the required bonus XP were expected to be completed by Week 15. The instructor provided various reminders during the semester and strongly recommended that learners to disperse their work throughout the semester rather than remaining inactive for a while and then pumping up their score close to the due date. These reminders may have influenced students' app use during this time; however, because they were allowed to work at their own pace, some participants did not follow these recommendations as reported in the results and discussion sections. Finally, only Duolingo work completed during the participants' corresponding semester of study counted toward their grade in the current analysis.

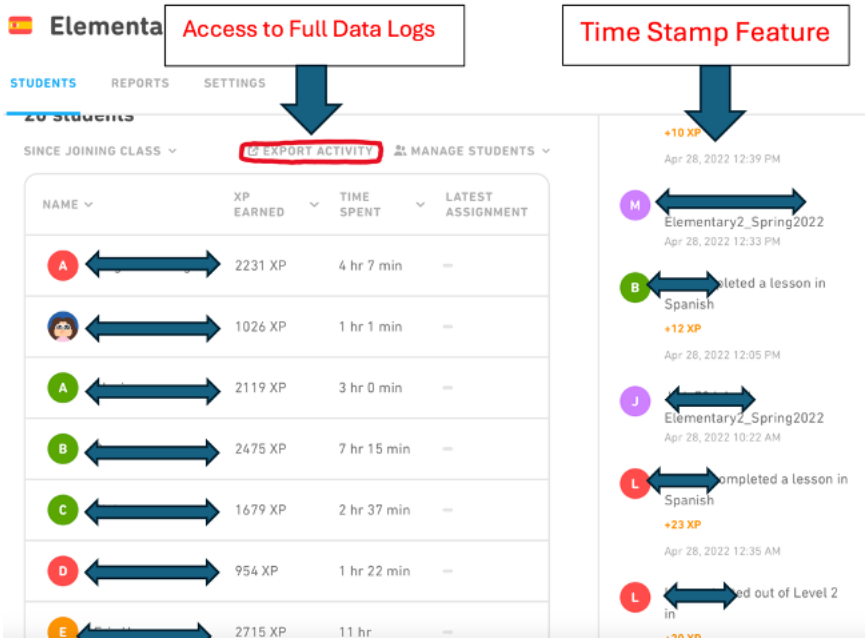
Research Instrument: Duolingo for Schools Tracking System

Duolingo for Schools has a tracking feature that allows teachers and researchers to see students' logs so they can follow their activity (e.g., progress, engagement) on Duolingo even several semesters after course completion. The platform allows the download of raw data, which include days active on the app, time spent on the app, total XP, and lessons completed. Duolingo for Schools also uses a timestamp system to report per student use each time a lesson is

completed: It shows students' progress (e.g., since joining class), unit, level, date and time, type of lesson (e.g., audio, review), and XP, which vary by lesson. Some of these platform tracking features (e.g., lessons completed and timestamp) have been considered appropriate and reliable parameters to examine app use in previous research studies (e.g., García Botero et al., 2019). Duolingo for Schools has other features, such as one that allows instructors to assign specific lessons as homework assignments; however, these features were not used in the current study. The platform was solely used in this study to track students' activity (engagement) on the app.

To analyze app use, we examined students' Duolingo for School logs, considering four main parameters: days active on the app, time spent on the app, total XP, and lessons completed. Full data logs were accessed and exported to an Excel spreadsheet for further analysis via the Export Activity function of the platform We also examined the timestamp feature to analyze average monthly patterns of app use (short term, see sample screenshot in Figure 1). Monthly

Figure 1
Duolingo for Schools Platform: Instructor's dashboard (Sample Screenshot)



active students were those who completed at least one Duolingo lesson in the defined month of measurement (García Botero et al., 2019). In addition, we used the timestamp feature to examine students' lessons completed after the end of the course (long term). Data were analyzed by groups (Elementary I vs. Elementary II) and by instructional modality (face-to-face vs. online). As noted above, our decision to use a comparative approach was based on instructional nuances (e.g., teacher–student or student–student immediacy) that exist in each learning environment despite the primary use of Duolingo for assignments outside the classroom. All data were analyzed after the students completed their corresponding semester.

Procedure

The current research was approved by the Institutional Review Board at the university where one of the authors is employed (Protocol # 19-078) and conducted accordingly. The activities with Duolingo were implemented for 11 of 15 weeks during the regular semester and for all 5 weeks during the summer session (Elementary I only). One of the researchers was the instructor for all sections in both instructional modalities, and the app was used as a supplement to regular classroom instruction. Each class met twice a week for 1 hour and 40 minutes during the regular semester and for 2 hours and 30 minutes daily during the summer session. The class was delivered in two modalities: face-to-face and online synchronous. The latter modality was implemented and delivered virtually via Microsoft TEAMS. Both instructional modalities involved the same curricular materials, including syllabi, classroom activities, textbook, and PowerPoint presentations. Regular classroom activities included those based on sociocultural approaches (e.g., collaborative work), and Duolingo activities were incorporated as part of the curriculum and completed outside classroom instruction. The implementation of this study followed recommendations from previous studies, indicating that Duolingo should be used for elementary-level students (Munday, 2016) as a supplement to regular classroom instruction (Cunningham, 2015; Falk & Götz, 2016; Munday, 2017).

Results/Findings and discussion

Research Question 1

Does a significant difference in app platform use exist between learners who use Duolingo in a face-to-face classroom vs. those who do so in an online synchronous environment?

Independent-Samples t-Tests

We conducted two separate independent-samples *t*-tests for the Elementary I and Elementary II groups to compare the four app use parameters (lessons completed, days active on the app, total XP, and time spent on the app) of the face-to-face group and the online synchronous group as described below.

Elementary I Group

Regarding lessons completed, results indicated a statistically significant difference for the face-to-face group and the online group; t (103) = 2.21, p = 0.29. The magnitude of the differences in the means indicated a medium effect size (Cohen's d = .454).

With days active on the app, no significant difference occurred for the face-to-face group and the online synchronous group; t (103) = .27, p = 0.79. The magnitude of the differences in the means indicated a very small effect size (Cohen's d = .055).

For total XP, no significant difference occurred for the face-to-face group and the online synchronous group; t (103) = .17, p = 0.87. The magnitude of the differences in the means indicated a very small effect size (Cohen's d = .035) as well.

For time spent on the app, no significant difference emerged for the face-to-face group and the online synchronous group; t (103) = .86, p = 0.39. The magnitude of the differences in the means indicated a small effect size (Cohen's d = .18). See Table 2.

Table 2

Elementary I Group

App use parameter	Group	*n*	*M*	*SD*	*Cohen's d*	*Sig.*
Lessons completed	Face-to-face	36	112.31	111.64	.454	.029
	Online (Synchronous)	69	73.91	66.43		
Days active on the app	Face-to-face	36	21.50	29.50	.055	.791
	Online (Synchronous)	69	19.94	27.97		
Total XP	Face-to-face	36	2104.58	1738.54	.035	.866
	Online (Synchronous)	69	2037.29	2034.44		
Time spent on the app	Face-to-face	36	317.89	332.80	.176	.394
	Online (Synchronous)	69	262.49	304.89		

Elementary II Group

Regarding lessons completed, a statistically significant difference emerged for the face-to-face group and the online synchronous group; t (36) = 3.58, p = 0.001. The magnitude of the differences in the means indicated a medium effect size (Cohen's d = .771).

With respect to days active on the app, no significant difference arose for the face-to-face group and the online synchronous group; t (36) = .78, p = 0.442. The magnitude of the differences in the means indicated a small effect size (Cohen's d = .325).

As for total XP, significant differences occurred for the face-to-face group and the online synchronous group; t (33.27) = 3.21, p = 0.003. The magnitude of the differences in the means indicated a medium effect size (Cohen's d = .734).

Regarding the time spent on the app, no significant difference appeared for the face-to-face group and the online synchronous group; t (36) = 1.90, p = 0.65. The magnitude of the differences in the means indicated a medium effect size (Cohen's d = .795). See Table 3.

Table 3

Elementary II Group

App use parameter	Group	*n*	*M*	*SD*	*Cohen's d*	*Sig.*
Lessons completed	Face-to-face	31	72.32	67.48	.771	.001
	Online (Synchronous)	7	24.57	14.84		
Days active on the app	Face-to-face	31	11.48	11.51	.325	.442
	Online (Synchronous)	7	8.00	5.10		
Total XP	Face-to-face	31	1440.58	1064.55	.734	.003
	Online (Synchronous)	7	720.86	310.29		
Time spent on the app	Face-to-face	31	218.97	176.62	.795	.065
	Online (Synchronous)	7	89.43	56.17		

In summary, in terms of lessons completed, statistical difference indicated that students in the face-to-face classroom had higher app use than those in the online (synchronous) classroom in both the Elementary I and Elementary II groups. In terms of total XP, the results were mixed: We found statistical differences for the Elementary II group but none for the Elementary I group. In terms of days active and time spent on the app, no statistically significant difference emerged for either group.

Notably, finding no statistically significant results for some parameters does not necessarily translate into zero differences in terms of overall app use between groups and modalities. A closer look at the mean scores across parameters indicated that participants in the face-to-face group had higher app use than those in the synchronous online classroom.

Research Question 2

What were the participants' monthly app use patterns during the semester (short term)?

To examine the monthly activity patterns during the semester, we analyzed the timestamps of all participants enrolled in the regular semester and summer sessions. The latter included only the Elementary I group because the instructor taught no Elementary II sections during the summer. Because the number of participants in each modality for both groups differed, monthly activity was calculated proportionally to the total population of each modality. This adjustment was made because using raw counts of app use could skew the results, favoring larger groups, and causing unfair comparisons. For example, for Elementary 1, we added all app users for each month and divided the sum by the total number of participants of each modality (e.g., face-to-face = 36); see formula below. By using proportions for each modality, monthly activity can be compared equivalently in face-to-face and online groups regardless of the total number of participants per modality. In other words, this adjustment normalizes the data and ensures that

larger groups do not appear disproportionally active compared to smaller ones across modalities. Those calculations were performed for Elementary I and Elementary II groups separately.

Formula used to calculate the proportions of the participants' monthly app use patterns:

$$\frac{\text{Sum of all active app users for each month (September +October +…+December)}}{\text{Total sample population for each modality (e.g., Elementary I, face-to-face = 36)}}$$

Elementary I Group

Regarding the regular semester users (fall and spring), the overall trend indicated that activity decreased gradually over time for both Elementary I groups. The same pattern was observed in the summer session users. See histograms in Figures 2 and 3.

Figure 2
Monthly App Use (Elementary I Fall and Spring Semesters Combined)

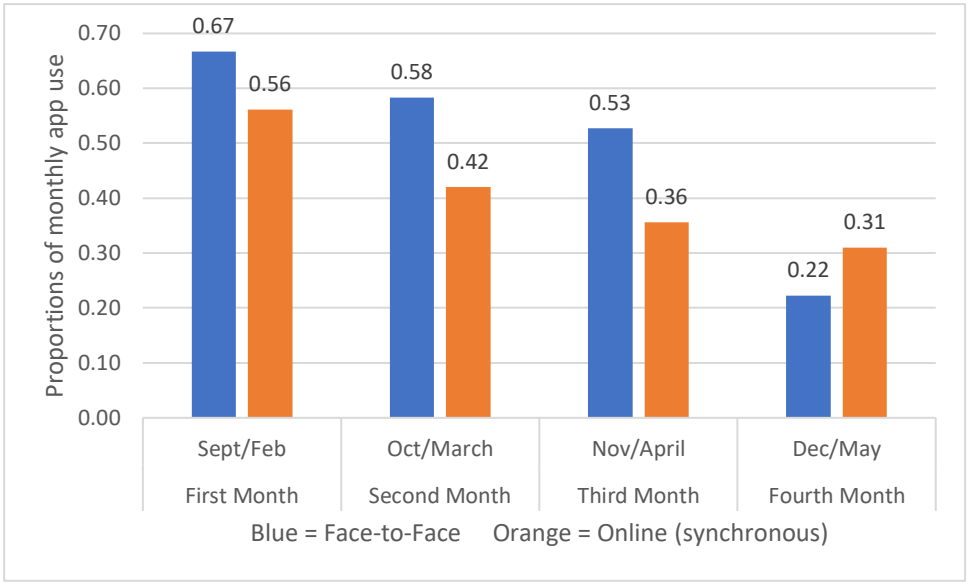
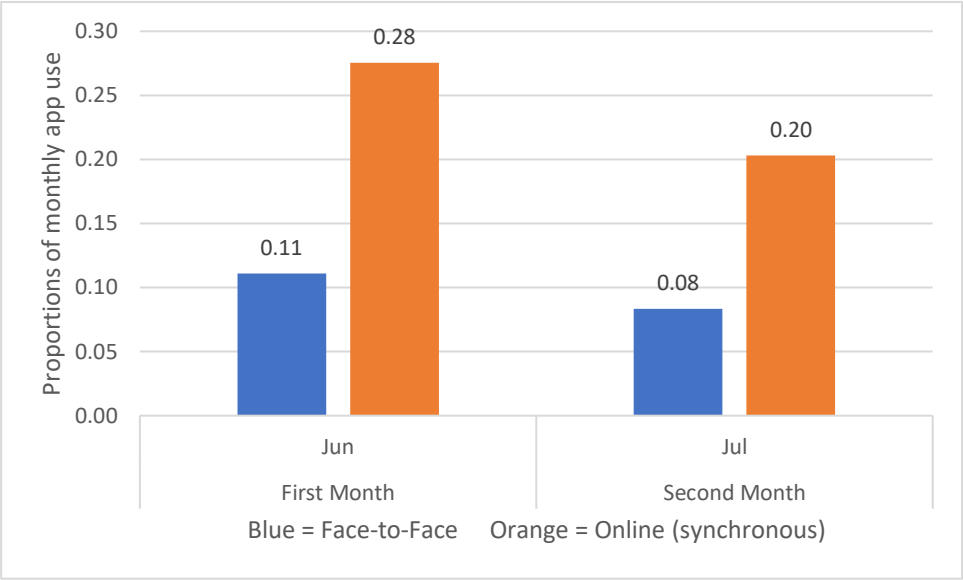


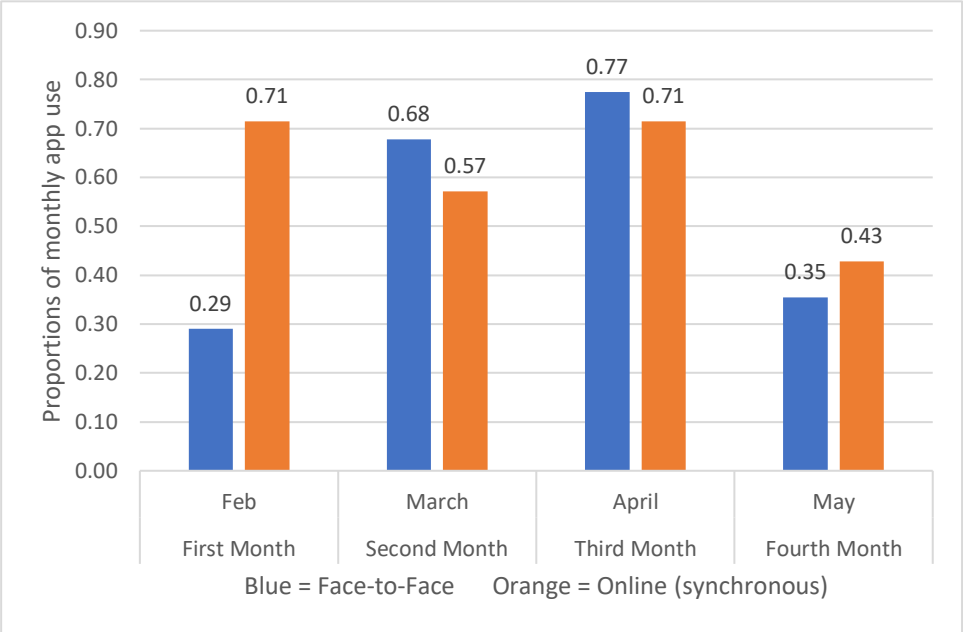
Figure 3
Monthly App Use (Elementary I Summer Session)



Elementary II Group

Overall patterns indicated that Elementary II participants tended to be more active by midsemester than at the beginning or end of the implementation of the study. See Figure 4.

Figure 4
Monthly App Use (Elementary II Spring Semesters)



In sum, Elementary I students' activity gradually declined as time passed, whereas the Elementary II group was on average more active by midsemester than at the beginning or end of the semester.

Research Question 3

To what extent do students engage in app use after course completion (long term) and for how long?

The results from both the Elementary I and Elementary II groups revealed similar trends regarding app use after course completion. In the Elementary I face-to-face group, 91.7% of students stopped using the app after the course ended while a small percentage (8.3%) continued using it sporadically—one for 10 months and two others for 1 and 2 months, respectively. In the online group, 76.8 % did not continue using the app while 23.2% kept using it for varying lengths of time: eight students used it for 1 month, five for 2 months, one for 3 months, another for 5 months, and one for 7 months. A similar pattern was observed in the Elementary II group, where 96.8% of face-to-face students discontinued app usage after the course and only 3.2% used it for 2 months. In the online group, 85.7% of participants stopped using the app, with 14.3% using it for 1 month. A summary of these findings is presented in Table 4 below.

Table 4

Activity After End of Semester

Activity	Elementary I Group		Elementary II Group	
	Face-to-face (n = 36)	Online (n = 69)	Face-to-face (n = 31)	Online (n = 7)
Zero activity	33 (91.7%)	53 (76.8%)	30 (96.8%)	6 (85.7%)
Some activity continued	3 (8.3%)	16 (23.2%)	1 (3.2%)	1 (14.3%)

In sum, for the Elementary I and Elementary II groups most students did not use the app after completing the course, and only a small number continued but mostly for a few months.

Discussion

In this section the results are discussed in terms of the three research questions.

Use of App Platform by Face-to-Face vs. Online Duolingo Learners

One of the purposes of this study was to investigate whether differences existed in the use of the app platform by learners who used Duolingo in a face-to-face environment vs. those who did so in an online classroom when that use was supplemental to regular instruction. The results of the statistical analyses show that the face-to-face group outperformed the online group in app use for both the Elementary I and Elementary II groups in terms of lessons completed. The same statistical differences apply in terms of XP but only for the Elementary II group. An examination of the mean scores across parameters, including the ones with no statistical differences like days active on the app and time spent on the app, further indicate that overall, the participants in the face-to-face group had higher app use than those in the online group.

As per previous research and our interpretation of the current results, learners' attitudes toward class modality could conceivably have had an impact on app use. For instance, both classes were conducted in real time (face-to-face and online), yet the instructor observed that certain online learners faced more distractions during class, such as interruptions from family members or issues with weak internet connections. In addition, some online participants appeared less engaged, simultaneously performing other tasks while being "connected" to the scheduled virtual meetings in contrast to their face-to-face peers. These observations are congruent with previous studies reporting that in-person learners had higher academic engagement and connectedness to teachers and classmates than learners enrolled in hybrid or remote learning modalities (McKellar & Wang, 2021). In addition, these results align with comparable studies reporting the impact of teaching modalities on student engagement (Bagheri & Zenouzagh, 2021; Gómez et al., 2011).

Further, even though learners in both learning modalities had opportunities for pair or group interactions (e.g., practice mini dialogues in the target language) during the regular class, the face-to-face group had more opportunities to engage in informal social conversations related to their personal or academic coursework. For example, before class started, the instructor observed that some students talked or joked about their weekly XP points totals (earned by completing lessons) ranked on the app leaderboard—an app feature that allows users to compare their language learning progress against other users, including classmates enrolled in same level or section. These informal social interactions may have provided a higher sense of informal competition and therefore higher app use when compared to the online learners who had limited opportunities for such interactions outside the virtual classroom setting, impacting their engagement.

Patterns of Participants' Monthly App Use During the Semester (Short Term)

Regarding the monthly Duolingo app use patterns, the results produce a mixed picture. For Elementary I (regular semester and summer session), results indicate that activity decreased over time for the face-to-face and online groups. For the Elementary II group, however, results indicate that participants tended to be more active by midsemester than the beginning or end of semester for both modalities. The findings for the Elementary I group are congruent with those reported by García Botero et al. (2019), who found that app use (as measured by lessons completed) decreased as time elapsed. The decreasing use of the app may be the result of the learners' fading motivation (García Botero et al., 2018, 2019) because using the app is no longer a part of their grades once the semester has ended. The Elementary I group may have used the app more often at the beginning of semester because of its newness (novelty effect) but may have lost interest because of some of its gamification features, such as prioritizing repetition and translation over meaningful feedback and context (Shortt et al., 2021; see also Loewen et al., 2019).

The decrease in use over time exhibited by the Elementary I group does not apply to the Elementary II group, however, because use was lower at the beginning but increased by midsemester and then decreased close to course completion. Similar app use patterns observed in the Elementary II group have not been reported in existing Duolingo literature. The researchers speculate that learners in this group, with more language knowledge and academic

skills, were better equipped to use more efficient learning strategies (e.g., self-regulation) than the Elementary I group who were typically in their first semester and adjusting to new academic routines. This may explain the midsemester spike in app usage, followed by a decline toward the end of the semester. Despite this decrease, the Elementary II group maintained an overall higher app use from mid to end of semester compared to the Elementary I group (see Figures 2 and 4 above).

Overall, these results show a mixed picture of the capability of Duolingo's gamified features in engaging learners with the app.

In the analysis provided above we considered the overall picture (mean) of app use, but a visual inspection of the data shows that not all students consistently used the app for 11 weeks. We noticed that some participants used the app until they completed the required points to obtain full credit for that portion of their assignment or their expected bonus points and then stopped using the app even though the semester was not over. Conversely, not all students completed homework assignments immediately after they joined the class but waited until closer to the end of the semester (e.g., November and December or April and May) to complete their homework.

Students' App Use After Course Completion (Long Term)

Regarding use of the app after course completion, most participants in this study did not use it after the end of the course; only a small number continued but mostly for a few months. These results are congruent with Bende (2017), who reported that the Duolingo app itself—despite its gamified elements—did not engage the students effectively in the long term. Bende argued that students need instructor scaffolding to be autonomous language learners in using Duolingo. In addition, the results in this study may have occurred because of the characteristics of these instructed learners; that is, the participants were students taking this course as a requirement for graduation. Results also suggest that these students completed the Duolingo activities primarily because they were incorporated into the grade or because they needed some extra credit to boost their final score.

In addition, based on the experience of both researchers, who taught on this campus, another possible reason most participants did not use the application in the long term may have been that their personal priorities interfered with their academic work. Most participants in this study were enrolled at a regional university campus, typically nontraditional students who had more responsibilities beyond their academic work than their traditional peers—those who live in campus dorms. Specifically, most participants in this study were commuters (in face-to-face classes) and attended college part-time. Some worked part or full time; others were married with families. Their priorities may have left little time or mental energy for extracurricular learning after the semester was over even if the resources were easily accessible.

However, a visual inspection of some outliers (not included in the main analysis) in the data—those whose scores were greater than +3 standard deviations (SD) from the mean—show that very few participants used the app for a substantial amount of time after their semester ended. For example, one student completed 1,402 lessons, earned 53,794 XP, spent 18,082 minutes, and was active for 872 days. Another student completed 806 lessons, earned 56,642 XP, spent

13,371 minutes, and was active on the app for 425 days. At the time of this writing (January 2024), both students had been active on the app since joining the class in fall 2019. A third student completed 916 lessons, earned 17,237 XP, spent 8,430 minutes, and was 164 days active on the app. The latter student consistently used the app for 15 months after course completion (fall 2019) and stopped using it after March 2021. These students were enrolled in this class as part of the senior guest program at our institution, which allows people aged 60 or older to audit courses on a tuition-free, space-available basis.

Readers should note that the senior guest program in our school does not require these students to do any academic work (e.g., tests or homework are optional); however, according to the instructor observations, they consistently turned in homework and took exams. In addition, the instructor noticed they were typically self-motivated, retired professionals (e.g., nurses, police officers) who took this class for personal development and—as verbally communicated to the instructor—they used L2 when they traveled to Spanish-speaking countries. The characteristics of these senior guest participants resembled those reported in studies by Vesselinov and Grego (2012), Jiang et al. (2020), and Loewen et al. (2019) as opposed to average college students, who were typically instructed foreign language learners, often externally motivated for grades and less experienced with few or no opportunities to proceed at their own pace (Krashen, 2014).

Maintaining motivation and engagement with the app after the semester can be challenging for less experienced learners. Instructors can address this by reminding students that language learning is a long-term process and encouraging them to articulate their plans for continuing their language studies after graduation—with or without the app. One possible solution is to incorporate a reflective component like a learning journal, where students can track and reflect on their progress. Additionally, offering post-semester support though a group chat (e.g. on WhatsApp) could provide a platform for learners to share their metacognitive progress and support one another.

Conclusion

Results indicate that, overall, participants in the face-to-face group had higher app use than those in the synchronous online classroom when the use occurred outside classroom instruction. Results also reveal that app use in the Elementary I group gradually faded as time passed, whereas the Elementary II students were on average more active by midsemester than at the beginning or end of the semester. Except for a few outliers, most students in this study did not use the app after course completion; only a small number continued but mostly for a few months. The results suggest that—at least for instructed learners—the app engages learners during the duration of the course provided the app activities are incorporated as part of the curriculum.

Notably, this study is not without limitations. Specifically, the current study did not include data collection through surveys, interviews, or reflective entries. Future researchers may consider incorporating a mixed-methods approach, offering additional insights into instructed learners' app use. In addition, results cannot be generalized because observed app use patterns pertain to instructed college-level learners rather than noninstructed learners. In a future mixed-methods study app use may be compared in instructed and noninstructed learners.

Paradoxically, the limitations of this study constitute its major strength because it provides unique insights into the discussion of Duolingo app engagement from the behaviorist engagement perspective using platform data logs; that is, most researchers of student engagement with digital technologies have analyzed or measured student engagement through self-reported perception data (Nkomo et al., 2021). Very few researchers studying Duolingo engagement have incorporated and analyzed platform-generated data in terms of the behaviorist dimension as part of their research analysis (e.g., García Botero et al., 2018, 2019). Given that most platform systems can track user activity, further research using computer-generated data is needed to better understand student engagement (Henri et al., 2015). The current study will help practitioners and researchers triangulate information with other studies—related to app engagement—in which other research metrics, such as surveys (Munday, 2016), questionnaires and interviews (Bende, 2017), or journal reflections (e.g., Kessler, 2021) were used. Finally, the current study will inform schools and administrators considering partnerships with online language applications or transitioning to fully online language classes about the advantages and challenges associated with integrating mobile apps into the curriculum.

References

Al-Bogami, B., & Elyas, T. (2020). Promoting middle school students' engagement through incorporating iPad apps in EFL/ESL classes. *Sage Open*, *10*(2). https://doi.org/10.1177/2158244020926570

Al Fadda, H. A., & Alaudan, R. M. (2020). Effectiveness of Duolingo app in developing learner's vocabulary, grammar and pronunciation: A case study of a blended TESOL classroom. *International Journal on Emerging Technologies*, *11*(5), 403–410.

Ardoiz García, L. (2017). *Gamification in English teaching in primary education* [Doctoral dissertation, University of Valladolid]. https://uvadoc.uva.es/bitstream/10324/29552/1/TFG-O-1222.pdf

Azar, A., & Iskandar Tan, N. (2020). The application of ICT Techs (mobile-assisted language learning, gamification, and virtual reality) in teaching English for secondary school students in Malaysia during COVID-19 pandemic. *Universal Journal of Educational Research*, *8*(11C), 55–63. https://doi.org/10.13189/ujer.2020.082307

Bacca-Acosta, J., & Avila-Garzon, C. (2020). Student engagement with mobile-based assessment systems: A survival analysis. *Journal of Computer Assisted Learning*, *37*(1), 158–171. https://publons.com/publon/10.1111/jcal.12475

Bagheri, M., & Zenouzagh, Z. M. (2021). Comparative study of the effect of face-to-face and computer mediated conversation modalities on student engagement: Speaking skill in focus. *Asian-Pacific Journal of Second and Foreign Language Education*, *6*, Article 5. https://doi.org/10.1186/s40862-020-00103-0

Bende, L. (2017). Encouraging high school students to become autonomous EFL learners: Exploring the possibilities of Duolingo. *Edulingua*, *3*(1), 53–65. https://doi.org/10.14232/edulingua.2017.1.4

Bond, M., Buntins, K., Bedenlier, S., Zawacki-Richter, O., & Kerres, M. (2020). Mapping research in student engagement and educational technology in higher education: A

systematic evidence map. *International Journal of Educational Technology in Higher Education*, *17*(2). https://doi.org/10.1186/s41239-019-0176-8

Cunningham, K. J. (2015). Duolingo. *Teaching English as a Second Language Electronic Journal*, *19*(1), 1–9. http://tesl-ej.org/pdf/ej73/m1.pdf

Dehghanzadeh, H., & Dehghanzadeh, H. (2020). Investigating effects of digital gamification-based language learning: A systematic review. *Journal of English Language Teaching and Learning*, *12*(25), 53–93. https://doi.org/10.22034/elt.2020.10676

Duolingo. (2024). *About us*. *Mission*. https://www.duolingo.com/info

Falk, S., & Götz, S. (2016). Interactivity in language learning applications: A case study based on Duolingo. In T. Zeyer, S. Stuhlmann, & R. D. Jones (Eds.), *Interaktivität beim Fremdsprachenlehren und–learnen mit digitalen medien. Hit oder Hype?* [*Interactivity in teaching and learning foreign languages with digital media: Hit or Hype?*] (pp. 237–258). Narr Verlag.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59–109. https://doi.org/10.3102/00346543074001059

García Botero, G., Botero Restrepo, M. A., Zhu, C., & Questier, F. (2019). Complementing in-class language learning with voluntary out-of-class MALL. Does training in self-regulation and scaffolding make a difference? *Computer Assisted Language Learning*, *34*(8), 1013–1039. https://doi.org/10.1080/09588221.2019.1650780

García Botero, G., Questier, F., & Zhu, C. (2018). Self-directed language learning in a mobile-assisted, out-of-class context: Do students walk the talk? *Computer Assisted Language Learning*, *32*(1–2), 71–97. https://doi.org/10.1080/09588221.2018.1485707

Golesorkhi, Z., & Marandi, S. S. (2025). Exploring the impact of gamified web-based collaborative writing on individual L2 writing performance in higher education. *Computer-Assisted Language Learning Electronic Journal, 26*(1), 1–20. https://doi.org/10.54855/callej.252611

Gómez, F. A. A., López, D. C., & Marin, L. F. G. (2011). Reading comprehension in face-to-face and web-based modalities: Graduate students' use of reading and language learning strategies in EFL. *Colombian Applied Linguistics Journal*, *13*(2), 11–28. http://www.scielo.org.co/scielo.php?pid=S0123-46412011000200002&script=sci_arttext&tlng=en

Govender, T., & Arnedo-Moreno, J. (2020, October 21–23). *A survey on gamification elements in mobile language-learning applications*. Eighth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM '20), Salamanca, Spain. ACM. https://doi.org/10.1145/3434780.3436597

Heilporn, G., Lakhal, S., & Bélisle, M. (2022). Examining effects of instructional strategies on student engagement in blended online courses. *Journal of Computer Assisted Learning*, *38*(6), 1657–1673. https://doi.org/10.1111/jcal.12701

Henrie, C., Halverson, L. R., & Graham, C., R. (2015). Measuring student engagement in technology-mediated learning: A review. *Computers & Education*, *90*, 36–53. http://dx.doi.org/10.1016/j.compedu.2015.09.005

Huang, C., Wang, T., & Li, Y. (2024). Mobile-assisted English learning beyond the classroom:

Understanding the effects of language proficiency on Chinese undergraduate students' behavioral engagement. *Education and Information Technologies*, *29*, 737–762. https://doi.org/10.1007/s10639-023-12301-7

Jiang, X., Rollinson, J., Plonsky, L., & Pajak, B. (2020). *Duolingo efficacy study: Beginning-level courses equivalent to four university semesters*. https://duolingo-papers.s3.amazonaws.com/reports/duolingo-efficacy-whitepaper.pdf

Kessler, M. (2021). Supplementing mobile-assisted language learning with reflective journal writing: A case study of Duolingo users' metacognitive awareness. *Computer Assisted Language Learning*, *36*(5–6), 1040–1063. https://doi.org/10.1080/09588221.2021.1968914

Krashen, S. (2014). Does Duolingo "trump" university-level language learning? *The International Journal of Foreign Language Teaching*, *9*(1), 13–15.

Lantolf, J., Thorne, S. L., & Poehner, M. (2015). Sociocultural theory and second language development. In B. Van Patten & J. Williams (Eds.), *Theories in second language acquisition* (pp. 207–226). Routledge.

Loewen, S., Crowther, D., Isbell, D. R., Kim, K. M., Maloney, J., Miller, Z. F., & Rawal, H. (2019). Mobile-assisted language learning: A Duolingo case study. *ReCALL*, *31*(3), 293–311. https://doi.org/10.1017/S0958344019000065

Looyestyn J., Kernot J., Boshoff, K., Ryan, J., Edney, S., Maher, C. (2017). Does gamification increase engagement with online programs? A systematic review. *PLoS ONE*, *12*(3), e0173403. https://doi.org/10.1371/journal.pone.0173403

McKellar, S. R., & Wang, M.-T. (2023). Adolescents' daily sense of school connectedness and academic engagement: Intensive longitudinal mediation study of student differences by remote, hybrid, and in-person learning modality. *Learning and Instruction*, *83*, Article 101659. https://doi.org/10.1016/j.learninstruc.2022.101659

Munday, P. (2016). The case for using Duolingo as part of the language classroom experience. *Revista Iberoamericana de Educación Digital*. http://blogderied.blogspot.com/2015/12/the-case-for-using-duolingo-as-part-of.html

Munday, P. (2017). Duolingo: Gamified learning through translation. *Journal of Spanish Language Learning*, *4*(2), 194–198. https://doi.org/10.1080/23247797.2017.1396071

Nkomo, L. M., Daniel, B. K., & Butson, R. J. (2021). Synthesis of student engagement with digital technologies: A systematic review of the literature. *International Journal of Educational Technology in Higher Education*, *18*(1), Article 34. https://doi.org/10.1186/s41239-021-00270-1

Rachels, J. R., & Rockinson-Szapkiw, A. J. (2018). The effects of a mobile gamification app on elementary students' Spanish achievement and self-efficacy. *Computer Assisted Language Learning*, *31*(1–2), 72–89. https://doi.org/10.1080/09588221.2017.1382536

Shortt, M., Tilak, S., Kuznetcova, I., Martens, B., & Akinkuolie, B. (2021). Gamification in mobile-assisted language learning: A systematic review of Duolingo literature from public release of 2012 to early 2020. *Computer Assisted Language Learning*. *36*(3), 517-554. https://doi.org/10.1080/09588221.2021.1933540

Sukmana, Y., Putri, A., Sulistyaningtyas, T., Suryani, Y., Waskita, D., Sembiring, J., & Rosmansyah, Y. (2024). Gamified mobile virtual laboratory for Indonesian language

learning. *Computer-Assisted Language Learning Electronic Journal, 25*(1), 84–107. https://callej.org/index.php/journal/article/view/90

Vesselinov, R., & Grego, J. (2012). *Duolingo Effectiveness Study.* https://static.duolingo.com/s3/DuolingoReport_Final.pdf

Biodata

Daniel A. Castañeda Vise is a Professor of Spanish at Kent State University. His primary research interests include the effects of instruction involving technology on second-language learning as well as discourse analysis of students' use of tense and aspect and temporal words and expressions in written narratives.

Moon-Heum Cho, Ph.D. is an associate professor in the Instructional Design, Development and Evaluation program at Syracuse University. His research interests include mobile technologies in language learning and the role of motivation and emotion in technology mediated learning environments.