Studies in Second Language Acquisition (2024), 1–27
doi:10.1017/S0272263124000354

RESEARCH ARTICLE

Acceptance and engagement patterns of mobile-assisted language learning among non-conventional adult L2 learners: A survival analysis

Hyun-Bin Hwang[1], Matthew D. Coss[1], Shawn Loewen[1] and Kaitlyn M. Tagarelli[2]

[1]Second Language Studies, Michigan State University, East Lansing, MI and [2]Mango Languages, Farmington Hills, MI
Corresponding author: Hyun-Bin Hwang; Email: hwangh16@msu.edu

(Received 23 October 2023; Revised 18 January 2024; Accepted 19 March 2024)

Abstract

Research on mobile-assisted language learning (MALL) has revealed that high rates of attrition among users can undermine the potential benefits of this learning method. To explore this issue, we surveyed 3,670 adult MALL users based on the Unified Theory of Acceptance and Use of Technology (UTAUT) and also conducted an in-depth analysis of their historical app usage data. The results of hierarchical k-means cluster analysis and recurrent event survival analysis revealed three major findings. First, three distinct profiles of learners were characterized by different MALL acceptance and engagement experiences. Second, those with greater MALL acceptance displayed more intense, frequent, and durable app usage (behavioral engagement). Lastly, high levels of MALL acceptance were associated with more frequent pauses in app usage but also (a) longer active usage, (b) shorter breaks before returning to the app, and, ultimately, (c) fewer dropouts. We argue that persistence is a multidimensional *process* involving cyclical phases of engagement, disengagement, dormancy, and reengagement, with each aspect, like intensity, frequency, and duration, building up cumulatively over time. Implications for promoting persistent MALL engagement are discussed.

Keywords

Mobile-assisted language learning (MALL); engagement; Unified Theory of Acceptance and Use of Technology (UTAUT); persistence; survival analysis; recurrent event analysis

Introduction

Persistence is key to second language (L2) learning because of the amount of time it takes learners to develop L2 proficiency (e.g., Winke & Gass, 2019). In instructed classroom contexts, external accountability like course requirements and the pressure to perform in class can encourage learners to study for extended periods of time; however, in self-study contexts, learners often need to find other reasons to persist (Reinders & Benson, 2017). Particularly in the realm of mobile-assisted language learning (MALL), in which increased MALL engagement has been found to predict proficiency gains (e.g., Burston & Giannakou, 2021; Loewen et al., 2020; Sudina & Plonsky, 2024), learner attrition is recognized as a considerable problem (Kessler et al., 2023; Nielson, 2011).

The current research project investigated the factors that contribute to learners' persistent autonomous L2 study when using one commercially available language learning app, Mango Languages (hereafter Mango). Mango provides structured, conversation-based courses that cover reading, listening, speaking, vocabulary, pronunciation, grammar, and culture in over 70 languages. As of March 2024, Mango had approximately 100,000 monthly active global users. Stepping outside the typical bounds of second language acquisition (SLA) research, the present study included adults in nonuniversity settings across a wide age range, predominantly learning languages other than English. Applying the Unified Theory of Acceptance and Use of Technology (UTAUT) as a theoretical framework, we used survey data to systematically identify learner profiles in terms of their MALL dispositions and then characterized each group with respect to their MALL acceptance and self-reported app experiences. Additionally, we analyzed their historical app usage data to explore MALL engagement patterns across various time scales using survival analysis, which, to our knowledge, has not yet been used in SLA research. Findings are discussed in relation to the nature of persistence and traits of persistent L2 learners in MALL.

Background literature

Unified Theory of Acceptance and Use of Technology for L2 learning

To benefit from technology-enhanced learning (e.g., MALL), learners must willingly adopt particular technologies and use them for extended periods of time. Several theoretical models such as the Technology Acceptance Model (Davis, 1989), which was originally developed to increase work productivity in noneducational settings, have recently been adapted to explore the components that can impact technology adoption and engagement in various educational settings (see Granić & Marangunić, 2019, for a systematic review). Particularly in MALL contexts, previous studies have shown that perceptual factors (e.g., perceived usefulness and ease of use for L2 learning, positive attitude toward mobile technology), supportive elements (e.g., social pressure, teacher/peer supports, instructional design), and individual differences (e.g., motivation, self-regulation skills, self-efficacy) can play significant roles in affecting L2 learners' decisions to use mobile technology for language learning (e.g., Hoi & Mu, 2021; Hsu & Lin, 2021; Lai, 2013; Lai et al., 2022; Zhang et al., 2023). Previous studies, therefore, have provided valuable insights on the learning conditions that can promote successful technology adoption and thereby enhance persistence in MALL engagement.

The UTAUT has been proposed as a meta-theoretical model (Venkatesh et al., 2003) to synthesize 32 elements from eight previously proposed technology acceptance models into one integrated framework. This unified framework brings together cognitive, emotional, and contextual elements, all of which have been shown to influence individuals' acceptance of technology (Straub, 2009). The UTAUT includes four core exogenous predictor variables (performance expectancy, effort expectancy, social influence, and facilitating conditions) and two endogenous outcome variables (behavior intention and use behavior). *Performance expectancy* refers to the degree to which individuals perceive a specific technology as being likely to result in performance gains. *Effort expectancy* is the degree to which individuals perceive technology as easy to use. *Social influence* is the degree to which individuals feel social pressure from others to use a specific technology. Finally, *facilitating conditions* represent the degree to which individuals evaluate the supportiveness of their surrounding environment toward using a specific technology. The fundamental premise of the UTAUT model is that these four determinants can either directly or indirectly influence individuals' intentions to use a particular technology as well as their actual engagement with the technology. Researchers have used the UTAUT to explore the acceptance and use of a wide variety of educational technologies, including e-learning platforms (e.g., Khechine et al., 2020), social robots (e.g., Guggemos et al., 2020), and artificial intelligence (e.g., An et al., 2023). However, there is still limited research on technology-enhanced L2 learning within the UTAUT framework (e.g., García Botero et al., 2018; Hoi, 2020; Hsu, 2023; Zou et al., 2022). Furthermore, the few existing studies have yielded inconsistent results depending on specific L2 learning contexts. For example, in an out-of-class context, Hsu (2023) reported that effort expectancy, social influence, and facilitating conditions, as well as three additional motivational components (autonomy, competency, relatedness), had positive influences on Taiwanese English as a Foreign Language (EFL) learners' intention to use and actual use of massive open online courses (MOOCs) for language learning. In a classroom setting, however, Zou et al. (2022) reported no effects of performance expectancy and effort expectancy on the use of smartphones for L2 learning. Instead, social influence and facilitating conditions along with perceived enjoyment and learner control were significant predictors of Chinese EFL learners' satisfaction, intention to continue language learning with smartphones, and perceived learning performance. Specifically looking at MALL, García Botero et al. (2018) demonstrated that performance expectancy, social influence, facilitating conditions, and attitude toward MALL positively influenced L2 learners' MALL usage. Alternatively, Hoi (2020) found that effort expectancy, when mediated by attitude toward MALL, was also a significant predictor of Vietnamese L2 learners' MALL usage. Overall, these findings indicate that the extent to which the UTAUT constructs explain learners' acceptance of and engagement in MALL may vary depending on specific conditions such as educational contexts and technological tools.

Multidimensional aspects of MALL engagement

Engagement can be manifested in various facets (e.g., behavioral, cognitive, and affective dimensions; see Hiver et al., 2021). In this study, we focus on the behavioral component of MALL engagement, using the term "usage" interchangeably with "(behavioral) engagement." When it comes to engagement in MALL settings, how *much* and how *often* learners use apps for L2 learning is of particular importance, given the positive correlations found between time spent on apps and L2 development. Measurements of engagement in previous MALL studies largely relied on learners' self-reported data, either using frequency-based scales (García Botero et al., 2018; Hoi, 2020; Hsu, 2023) or asking learners to retrospectively provide the amount of time they spent on an app on a weekly basis (e.g., Kessler et al., 2023; Loewen et al., 2019). However, such delayed self-reports might be vulnerable to "rosy retrospection" (van Berkel et al., 2018), where individuals recall events more positively than they actually experienced them. Real-time data collection, on the other hand, can provide a more reliable approach to investigating MALL engagement patterns (Reinders et al., 2023).

Relatedly, there have been attempts to identify distinct groups of learners who share similar MALL experiences in previous research. Such categorizations have been based on either levels of consistent MALL engagement (e.g., García Botero et al., 2019; Jeon, 2022) or types of MALL activities (e.g., Lamb & Arisandy, 2020; Peng et al., 2022). From a methodological perspective, cluster analysis holds an advantage in profiling diverse learner groups because it allows for patterns to naturally emerge from the data, rather than relying on a priori categories. This approach is more ecologically valid and can offer valuable insights into understanding distinct MALL learner profiles that share similar MALL experiences (see Peng et al., 2022, for a "person-centered" approach to MALL research).

Another meaningful way to explore MALL engagement patterns is to study MALL engagement as a multidimensional construct. To our knowledge, Sudina and Plonsky's (2024) study is the first MALL study that used fine-grained, multidimensional constructs of in-app engagement patterns based on behavioral data. In their study, they operationalized app engagement in terms of duration (total minutes per participant across the 6-month period of study), frequency (the number of times the learner opened the app in a given week and the number of days the learner completed at least one lesson), and intensity (the number of content-related/curriculum-oriented activities completed). They found that intensity measures were more dependable predictors of L2 proficiency gains than total minutes of app exposure. This suggests that more frequent and purposeful in-app engagement with an aim to complete activities can be more impactful than periods of extended engagement alone. Overall, previous studies provide useful reference points for understanding patterns of MALL engagement, but none have directly explored the issue of persistence in MALL.

Indeed, while learner engagement is strongly related to persistence (e.g., Rumberger & Rotermund, 2012), there is no universally accepted approach to measuring persistence. Previous research has mainly operationalized persistence as learners' intention to complete tasks or stay enrolled in courses (e.g., Jung & Lee, 2018) or their intention to continue identical or related tasks or courses in the future (e.g., Feng & Papi, 2020; Lakhal et al., 2021). However, several complications can arise when using learners' intention to use technology as a proxy for sustained engagement with it. For example, Feng and Papi (2020) asked learners to respond to the statement "I want to keep on learning Chinese *as long as possible*" as a way of measuring persistence. However, each learner might have their own interpretation of "as long as possible." Given that actual engagement manifests as observable actions following an initial intention (Hiver et al., 2021), a mere *desire* to use apps may not necessarily result in persistent MALL engagement.

Although previous studies (e.g., García Botero et al., 2021) have identified high attrition rates as a potential barrier to reaping the benefits of MALL, it is noteworthy that very little research has directly explored the interface of MALL engagement and persistence. Studies that have attempted to measure persistence have used various approaches, such as the total amount of time spent on apps (e.g., Kessler et al., 2023), the number of in-app activities completed (e.g., Sudina & Plonsky, 2024), or students' attendance logs (e.g., Jeon, 2022). It should be noted, however, that these approaches generally treat persistence as a one-dimensional construct and thus might overlook the complex interplay of diverse engagement patterns (i.e., intensity, frequency, duration, pauses, dormancy, dropouts; see Table 1) that culminate in the dynamic nature of

https://doi.org/10.1017/S0272263124000354 Published online by Cambridge University Press

Table 1. MALL engagement pattern indices

Engagement indices	Descriptions
Intensity	The mean time spent on the app in a single session
Frequency	The mean amount of active app usage per week
Duration	The maximum continuous number of weeks that the app is actively used
Pause	An event in which a user stops active app usage, but resumes active app usage later
Dormancy	A pause period when a user is not actively using the app but could become active again
Dropout	An event in which a user fails to resume active app usage after the most recent active app usage ends

Note: 'Active' is defined as usage that exceeds a specified threshold for each index (See Analysis section for our chosen thresholds). As a result, app usage below a specified threshold level was considered inactive.

persistence. In light of these challenges, there is a need to consider engagement trajectories using real usage data to understand persistence in MALL in relation to L2 learner characteristics.

The present study

We explored acceptance of and engagement in MALL by addressing two research gaps in previous studies. First, we expanded the scope of the population to encompass more diverse profiles of L2 learners, including less-often-researched older adults (i.e., nonuniversity students) from various backgrounds. It is essential to consider a wide range of individuals with distinct needs and motivations within the broader MALL context (Puebla & García, 2022). This is particularly important because technology has not only expanded the landscape of language learning beyond traditional classrooms (Reinders & Benson, 2017) but also because the "lifelong mobility of MALL" (Stockwell, 2022) has transformed language learning into a continuous lifelong endeavor, rather than a pursuit confined to a specific point in one's lifetime (e.g., schooling). The second research gap is that, although the UTAUT provides a model for explaining L2 learners' intention to use technology and actual engagement, the latter component has been overlooked in previous technology acceptance studies. Indeed, there are varying conceptualizations and measurements of the construct of engagement across studies, but few studies directly explored learners' behavior patterns (cf. Sudina & Plonsky, 2024). This lack of direct investigation into engagement makes it challenging to evaluate the extent to which learners' intention to engage in MALL is associated with their actual behaviors and usage patterns.

The present study consists of two related components. First, we used cluster analysis to identify distinct learner profiles based on UTAUT variables and then characterized each group with respect to their MALL acceptance and self-reported app use experience. Subsequently, we analyzed usage data obtained from the Mango database to explore MALL engagement patterns with six key indices along increasing time scales, as summarized in Table 1 (see Analysis and Result sections for detailed rationales in defining an active app usage threshold for each index).

Additionally, we employed survival analysis to explore the relationship between MALL acceptance and persistence-related engagement variables (i.e., pause, dormancy, dropout). Broadly, survival analysis is a regression-based analysis commonly used to predict the time leading up to an event of interest (e.g., dropping out of school; see Plank et al., 2008) and to estimate the likelihood of groups or individuals experiencing the

https://doi.org/10.1017/S0272263124000354 Published online by Cambridge University Press

event of interest while considering other influencing factors. We specifically employed recurrent event analysis (Chiou et al., 2023), a subcategory of survival analysis that is applicable when an event of interest occurs more than once. In the present study, the two target events are the recurrence of pauses in active MALL engagement and the failure to resume it (i.e., pausing without eventually restarting), in other words, a complete dropout from app usage. In survival analysis, the latter is called a *terminal* event (see Analysis 2 for more detailed information).

We explored the extent to which L2 learners' MALL acceptance is associated with their MALL experiences and engagement patterns and, ultimately, their *survival* in MALL (i.e., persistence) along increasing time scales (i.e., days, weeks, months, years). Three guiding research questions (RQs) were formulated as follows:

RQ1. How do MALL experiences differ depending on different UTAUT profiles?

RQ2. How do the patterns of (a) intensity, (b) frequency, and (c) duration in MALL engagement differ depending on different UTAUT profiles?

RQ3. How do the patterns of (a) pause, (b) dormancy, and (c) dropout in MALL engagement differ depending on different UTAUT profiles?

Method

Participants

All individuals who used Mango at least once from October 1, 2022, to March 14, 2023 (N = 52,432) were invited to participate in this study via email. Additionally, recruitment advertisements were shared on Mango's social media pages. Between March 14 and June 1, 2023, we collected responses from 5,056 participants (9.64% response rate). Among these respondents, 3,670 individuals completed the entire survey and were therefore eligible for inclusion in the analysis. All participants voluntarily joined and gave informed consent prior to completing the survey. The vast majority of participants accessed Mango for free through their public libraries—less than 2% had a paid account. Table 2 summarizes the participants' characteristics.

Participants reported learning 65 different languages or language varieties on Mango, with the following occurring most commonly: Spanish (n = 953), French (n = 451), Italian (n = 311), Japanese (n = 269), German (n = 236), Chinese (n = 169, all varieties), Arabic (n = 119, all varieties), Korean (n = 110), Russian (n = 103), and English (n = 90). The location of survey respondents was also geographically diverse (see Figure 1). Approximately 76% (n = 2,784) were located in the United States, followed by Canada (n = 376), Australia (n = 98), Germany (n = 32), the United Kingdom (n = 31), and Mexico (n = 20). The remaining 9% of the participants came from 55 additional countries. Compared to many studies in SLA (see Andringa & Godfroid, 2020; Godfroid & Andringa, 2023; Plonsky, 2023), this sample is much more diverse in terms of educational level, target language, and geographic location. Full participant information is provided in Supplementary Material A.

Instruments

Survey

The survey, which was administered using Qualtrics, was iteratively piloted by the research team and then by three Mango users of similar backgrounds to the target

https://doi.org/10.1017/S0272263124000354 Published online by Cambridge University Press

Table 2. Characteristics of the app users

Characteristics	Response frequency				
Age (*M* = 43.95, *SD* = 17.02, Range = 18–93)	20s and below	30s	40s	50s	60s and above
	897 (24.4%)	885 (24.1%)	575 (15.6%)	453 (12.3%)	860 (23.4%)
Highest level of education	High school or below	Some college	2-year degree	4-year degree	Graduate degree
	230 (6.2%)	357 (9.7%)	203 (6.2%)	1,358 (37%)	1,522 (41.4 %)
Number of L1s	One	Two		Three	Four
	3,046 (82.9%)	491 (13.3%)		93 (2.5%)	40 (1%)
Number of L2s learned	One	Two	Three	Four	More than four
	511 (13.9%)	1,137 (30.9%)	868 (23.6%)	386 (10.5%)	768 (20.9%)
Number of L2s learned using mobile apps	One	Two	Three	Four	More than four
	1,502 (40.9%)	1,049 (28.5%)	514 (14%)	187 (5%)	418 (11.3%)
Current target L2 proficiency	No or limited communication	Basic communication	Extensive communication	Limited professional ability	Extensive professional ability
	1,778 (32%)	1,479 (40.2%)	299 (8.1%)	87 (2.3%)	27 (0.7%)
L2 proficiency goal	Limited communication	Basic communication	Extensive communication	Limited professional ability	Extensive professional ability
	71 (1.9%)	645 (17.5%)	1,997 (54.5%)	426 (11.6%)	531 (14.4%)

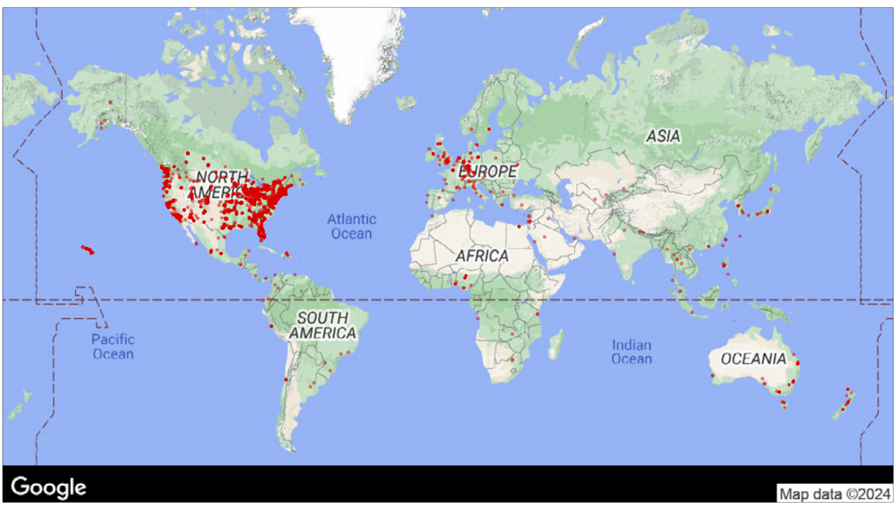

Figure 1. Geographical distribution of the survey respondents.

participants. Survey items and survey flow were revised to eliminate bugs, clarify wording, and reduce the total number of items as much as possible to minimize fatigue. The final survey consisted of two major sections: (a) self-reported MALL experiences and (b) MALL dispositions (i.e., the various components of the UTAUT framework). Survey items utilized either yes/no binary choices or 5-point Likert scales (1 = "never" or "strongly disagree"; 5 = "always" or "strongly agree"). The survey was administered in English. The full original survey is available in Hwang et al. (2024a) through the

https://doi.org/10.1017/S0272263124000354 Published online by Cambridge University Press

Instruments and Data for Research in Language Studies (IRIS) database (https://www.iris-database.org).

Part 1: MALL experience questionnaire

The purpose of this section was to collect data on respondents' MALL experiences. If respondents were learning multiple languages on Mango, they were first asked to choose one primary target language on which to focus their responses to survey items in this section. 32 total items elicited data about (a) reason(s) for learning the target L2 (6 items), (b) L2 learning environment(s) (four items), (c) use of non-app resources for L2 learning (four items), (d) perceived importance of developing different L2 skills (six items), (e) frequency of individual L2 skill practice (six items), and (f) satisfaction with L2 skill practice opportunities (six items).

Part 2: MALL disposition questionnaire

The purpose of the MALL disposition survey was to measure respondents' tendencies on variables related to MALL acceptance. Survey items were developed based on instruments published in previous studies using the UTAUT model (e.g., García Botero et al., 2018; Hoi, 2020). This section included 22 items tied to the five UTAUT determinants that can affect L2 learners' MALL acceptance and their actual engagement: (a) performance expectancy (three items, α = .86), (b) effort expectancy (two items, α = .83), (c) facilitating conditions (six items, α = .82), (d) social influence (four items, α = .78), and (e) attitude toward MALL (three items, α = .91). Additionally, level of MALL acceptance was measured as the intention to continue to engage in MALL-related activities in the future (four items, α = .74). Cronbach's alpha levels indicate very good reliability of all survey items.

In-app usage data

We extracted historical app usage data covering a 14-year period from May 23, 2009, to June 29, 2023, from all participants who completed the survey and provided a valid email address and/or app user ID (n = 3,319 of the 3,670 individuals; 90.4%). The in-app engagement data included information about the time spent during each app session (in seconds), with a total of 1,854,253 data points. We transformed this raw data into six key engagement indices across different temporal scales: intensity, frequency, duration, pause, dormancy, and dropout (see Table 1).

Overview of analyses and results

In our data analysis process, we employed hierarchical k-means cluster analysis, principal component analysis (PCA), and linear regression for RQs 1 and 2, and recurrent event (survival) analysis for RQ3. Given that the statistical results of the first two RQs informed methodological decisions for the subsequent recurrent event analysis, we present two independent sections for the analysis and results for RQs 1 and 2 separately from those for RQ3. For effect size interpretation, we followed Funder and Ozer's (2019) guidelines specifically suggested for human psychology constructs (e.g., cognition, emotion, behavior). An effect size of r = .05 suggests a very small impact on single events that could matter in the future. An effect size of r = .10 has a small impact at the level of single events but is ultimately more consequential. An effect size of r = .20 indicates a medium effect that provides some explanatory and practical use even in the short run, and an effect size of $r = .30$ indicates a large effect that has potentially powerful influences in both the short and the long term. In the context of psychological research, very large effect sizes ($r = .40$ or higher) are likely to be overestimations and will rarely be observed in a large sample or in a replication. Full descriptive statistics and statistical analysis results can be found in Supplementary Materials B to G. The R code used for data analysis is available in Hwang et al. (2024b) through the IRIS database.

Analysis 1: Cluster analysis, PCA, and regression analysis

For each individual, a mean score for each UTAUT variable was calculated. Subsequently, a hybrid (hierarchical k-means) approach to cluster analysis (Crowther et al., 2021) was employed in order to discover distinct sub-groups of learners based on their scores on the five UTAUT determinants. Using the *NbClust* package in R, we first identified the number of clusters using an agglomerative hierarchical clustering algorithm with Ward's linkage and squared Euclidean distance methods. Following the majority rule, the optimal number of clusters was determined to be three, as proposed by nine of 24 clustering validity indices. After examining initial cluster solution plots, six influential cases were identified as outliers because they noticeably intruded into the other cluster area, making the cluster distinction unclear. As a result, those outliers were excluded from the analysis, resulting in a total of 3,664 participants. Subsequently, we performed a new cluster analysis and a PCA using the *factoextra* and *FactoMineR* packages, respectively. We used the resulting clusters as a primary variable for the subsequent descriptive statistics (RQ1) and as a predictor in the regression analyses of intensity, frequency, and duration (RQ2).

As for the regression analyses, we calculated the individual-wise mean values and their confidence intervals (CIs) for each engagement index and performed winsorization, a robust statistical method to address possibly inflated standard error from skewed data distribution in a linear regression model (Mair & Wilcox, 2020; Wilcox, 2005; see Hui, 2020, for winsorization in SLA). Instead of simple data trimming, a winsorization procedure pulls in the extreme tails of the distributions by replacing the smallest and largest 5% of individuals' values with the values of individuals at 5% and 95%, respectively. Consequently, this winsorization impacted 330 of the 3,319 data points (9.94%) for intensity, 328 of the 3,290 data points (9.96%) for frequency, and 132 of the 2,862 data points (4.61%) for duration (see Results 1 RQ2 for a data inclusion criterion for each index). Full descriptive statistics and separate sensitivity analyses with the nonwinsorized data are reported in Supplementary Material F.

Results 1

RQ1. Different learner profiles showed different levels of MALL acceptance and experiences

Figure 2 displays a PCA plot with the result of the cluster analysis. The resulting three clusters contained 1,620, 1,276, and 768 individuals, respectively. The PCA results confirmed that performance expectancy ($r = .87$), effort expectancy ($r = .86$), and MALL attitude ($r = .90$) exhibited strong correlations with Component 1. These three variables closely aligned and most strongly explained the differences across the three clusters (eigenvalue = 2.47 with 49.4% variance explained). On the other hand, social influence explained the cluster differences in its unique way and was highly correlated with Component 2 ($r = .93$), being marked as the second most influential factor (eigenvalue = 0.99 with 19.8% variance explained). In Figure 2, the distances between cluster centers

https://doi.org/10.1017/S0272263124000354 Published online by Cambridge University Press

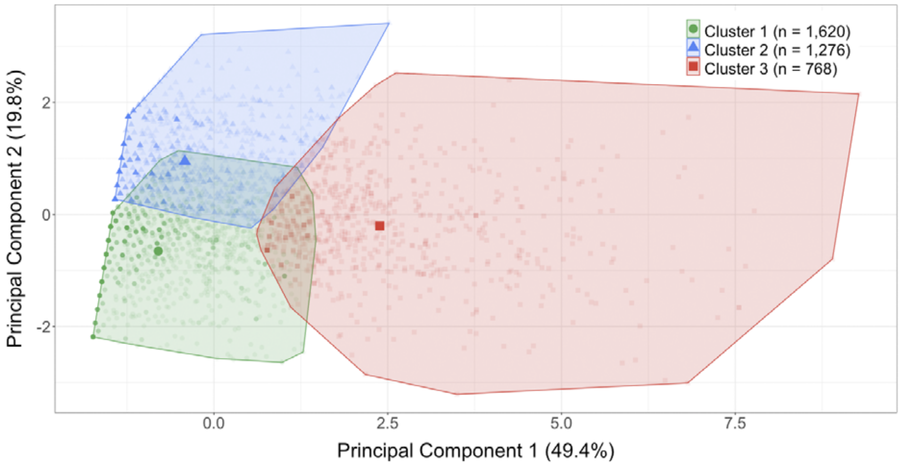

Figure 2. Three-cluster solution with PCA (the axes represent PCA scores).

along an axis indicate the degree of similarity between clusters in the corresponding principal component. Along the *x*-axis, the centers of Clusters 1 and 2 are relatively close to one another, suggesting their similarity in Component 1 (as opposed to Cluster 3). Along the *y*-axis, Clusters 1 and 3 are relatively close to one another, suggesting their similarity in Component 2 (as opposed to Cluster 2).

Figure 3 summarizes the centroid of each cluster, and the five individual points on each plot represent the mean of each UTAUT variable in that cluster. On the five-point Likert scale used in our survey, a score of 3 corresponds to a neutral response (i.e., “neither agree nor disagree”), mean scores above 3 are positive, whereas mean scores below 3 are negative. Clusters 1 and 2 displayed very high levels of performance expectancy, effort expectancy, and MALL attitude (i.e., Positive), whereas Cluster 3 had more neutral responses to these UTAUT variables. On the other hand, Cluster 1 showed neutral-positive responses (i.e., +Social), but Cluster 2 showed negative responses in social influence (i.e., –Social). Cluster 3 was in-between, with a neutral-negative response. This indicates that Cluster 1 has experienced some level of social influence (e.g., encouragement, recommendation) to engage in MALL activities, while Clusters 2 and 3 might be less affected by such influences. Regarding facilitating conditions (e.g., easy and stable access to technology devices, stable internet, and access to places to study), it was similarly high across all three clusters and therefore did not contribute to differentiating the clusters. Taken together, Clusters 1 and 2 both had overall positive responses to MALL expectancy and attitude, differing from each other only in social influence. Cluster 3, on the other hand, was overall more neutral across most categories. We therefore renamed the clusters as follows: Cluster 1 as Positive/+Social, Cluster 2 as Positive/–Social, and Cluster 3 as Neutral.

Figure 4 demonstrates responses to the four questions about MALL acceptance, each of which respectively indicates their willingness to (a) continue their current MALL engagement (i.e., *I will continue to use the app to learn the target language*), (b) recommend MALL to others for learning their target languages (i.e., *I recommend that people use the app for learning additional language(s)*), (c) allocate more time for future MALL engagement (i.e., *I would like to spend more time learning the target language on the app than I currently spend*), and (d) expand their efforts to new

https://doi.org/10.1017/S0272263124000354 Published online by Cambridge University Press

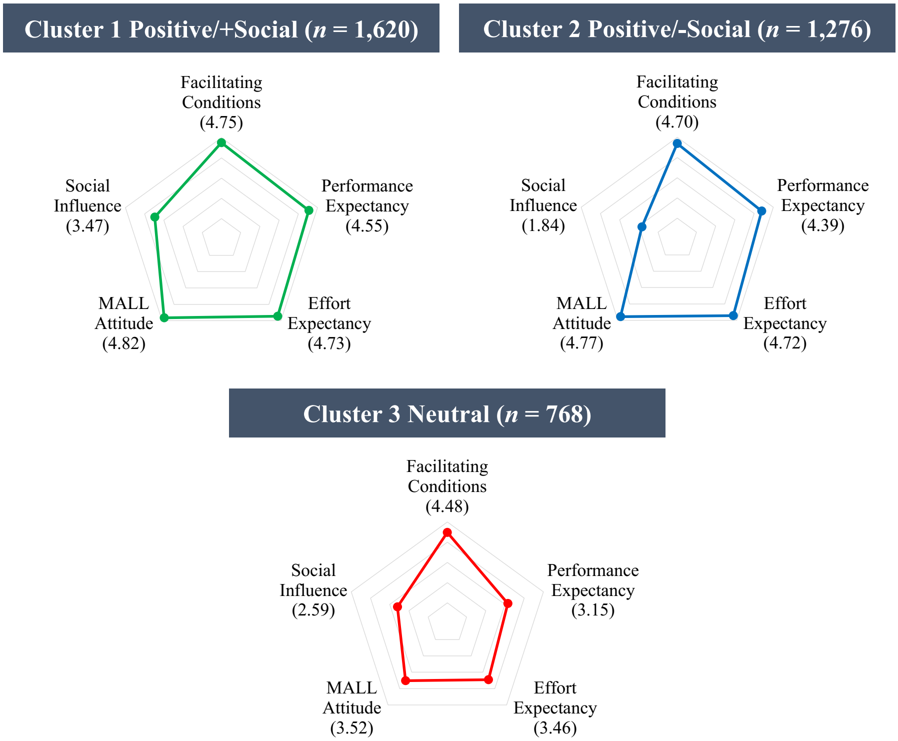

Figure 3. Cluster centroids with the means for each UTAUT determinant.

language learning (i.e., *I will use the app for learning one or more additional language(s) in the future*). Overall, Clusters 1 and 2 demonstrated high levels of MALL acceptance across all four questions, as evidenced by more frequent strongly positive responses. On the other hand, Cluster 3 showed neutral-to-positive MALL acceptance levels. The means across all four questions, indicating overall MALL acceptance levels, were 4.59 for Cluster 1 (95% CI [4.56, 4.61], and 4.43 for Cluster 2 (95%CI [4.40, 4.45]) and 3.52 for Cluster 3 (95% CI [3.46, 3.58]). Nonoverlapping 95% CIs indicate that the levels of MALL acceptance of Clusters 1 and 2 were significantly higher than Cluster 3.

We also explored distinct traits and MALL experiences of each learner group identified by the cluster analysis. Figure 5 illustrates the age distribution of users by cluster. Clusters 1 and 3 skewed younger, with about 50% of learners in the 20s–30s age groups, and only about 18% of learners in the oldest age group (60+). In contrast, Cluster 2 skewed older, with about 32% of learners in the 60+ age group, nearly double the percentage of learners in their 20s and 30s. Across all three clusters, learners in their 40s and 50s represented the smallest age groups.

Figure 6 overviews learners' reasons for learning their primary target languages. Similar trends were found across all three clusters, in which extrinsic factors (i.e., academic and job requirements) were not important reasons, whereas intrinsic factors were the predominant drivers for target language learning. Among these, self-satisfaction was most important, followed by learning about cultures and then communication with people in daily life. However, the proportion of individuals

https://doi.org/10.1017/S0272263124000354 Published online by Cambridge University Press

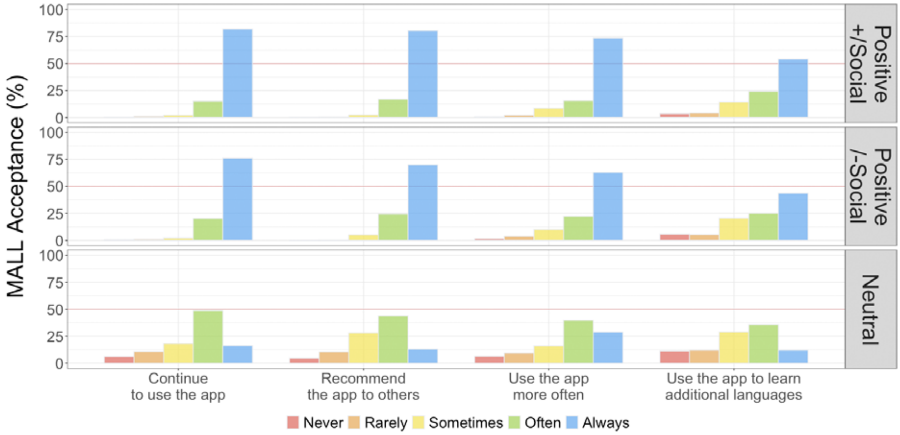

Figure 4. MALL acceptance levels.

Figure 5. Age distribution by cluster.

who strongly agreed on the importance of the cultural and communicative aspects of the target language was the highest for Cluster 1 (about 45%). Additionally, Cluster 2 exhibited the highest "strongly disagree" responses (about 20%) and lowest "strongly agree" responses (about 29%) regarding communication with others as a reason for learning the target language. These findings highlight that the learners in Cluster 1 are motivated by multiple reasons for target language learning; they are more communicatively and culturally oriented compared to the other two groups. Learners in Cluster 2, on the other hand, are less driven by the need to communicate with people in their daily lives. Regarding the environments where their target language is regularly used, Cluster 1 showed a higher proportion (about 39%) of individuals who have social situations (e.g., family, friends) that necessitate regular use of the target language compared to Cluster 2 (about 21%). It is noteworthy that the differences between Clusters 1 and 2 in terms of (a) the value placed on communication for L2 learning and (b) the richness of the L2 learning environment are consistent with the earlier finding that Cluster 1 experienced a neutral-positive social influence on MALL engagement, whereas Cluster 2 did not have such influence.

https://doi.org/10.1017/S0272263124000354 Published online by Cambridge University Press

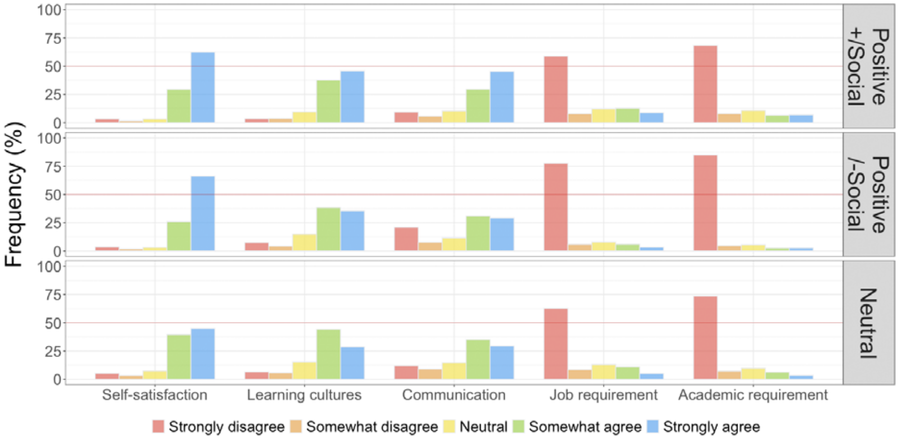

Figure 6. Reasons for learning the primary target language.

Regarding the opportunity to practice L2 skills, Clusters 1 and 2 reported more frequent practices in all skills on the app, as well as higher satisfaction with the opportunities to practice these skills, compared with Cluster 3 (see Figures 7 and 8). Additionally, 60.99% (988 of 1,620) of Cluster 1 and 64.34% (821 of 1,276) of Cluster 2 described Mango as an "important and primary source of learning." Conversely, for Cluster 3, only 24.87% (191 out of 768) viewed Mango as their primary learning source, with the small majority considering it to be a "useful supplement for learning" at 35.02% (269 out of 768). These findings indicate that higher levels of MALL acceptance are associated with greater in-app engagement and a higher level of satisfaction from such app engagement.

Table 3 summarizes the characteristics and MALL experiences of each cluster. Overall, Cluster 3 remained relatively neutral across all measured variables and, therefore, serves as the reference group for the following regression analyses.

RQ2. Differences in MALL acceptance had cumulative impacts on MALL engagement

Figure 9 displays the descriptive statistics of usage *intensity* (i.e., the average time spent on the app per session) with a horizontal line indicating the overall mean value (M = 339.61 seconds, SD = 128.67 seconds, $Range$ = 148.92–614.34 seconds). Cluster 2 spent 357.43 seconds per session (95% CI [349.80, 365.07]) and Cluster 1 spent 338.20 seconds per session (95% CI [331.84, 344.57]). Cluster 3 spent 311.61 seconds per session (95% CI [302.19, 321.02]). Despite the marginal time difference in average time spent on the app per usage, the nonoverlapping 95% CIs indicate that the differences in usage intensity among the three clusters were significant. A linear regression confirmed that Cluster was a significant predictor of usage intensity ($F(2, 3316) = 27.57$, $p < .001$, r = .13). Although the effect size was small-to-medium at the level of single events, it still may potentially be impactful on cumulative language learning outcomes.

Based on the intensity findings, we established a threshold of 300 seconds and above per session (i.e., at least 5 min) to define "meaningfully active MALL engagement" at an individual session level for subsequent frequency analysis. In our study, *frequency* represents the average number of active app usage per week on which a learner used the

https://doi.org/10.1017/S0272263124000354 Published online by Cambridge University Press

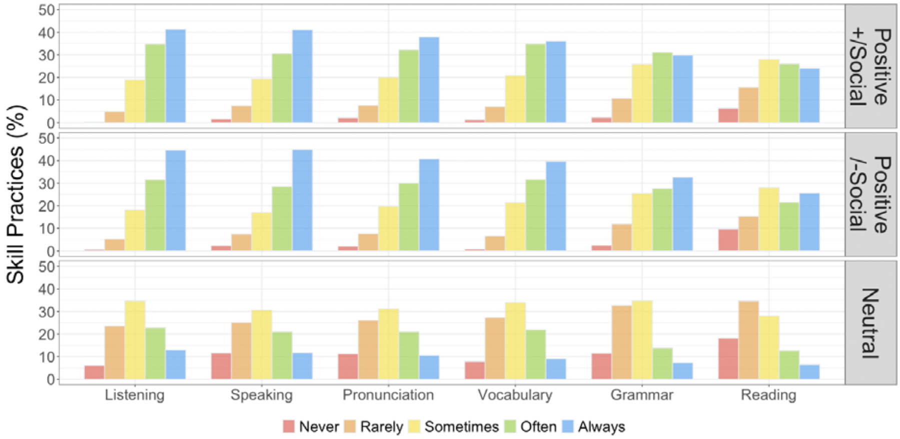

Figure 7. Frequency of L2 skill practice on the app.

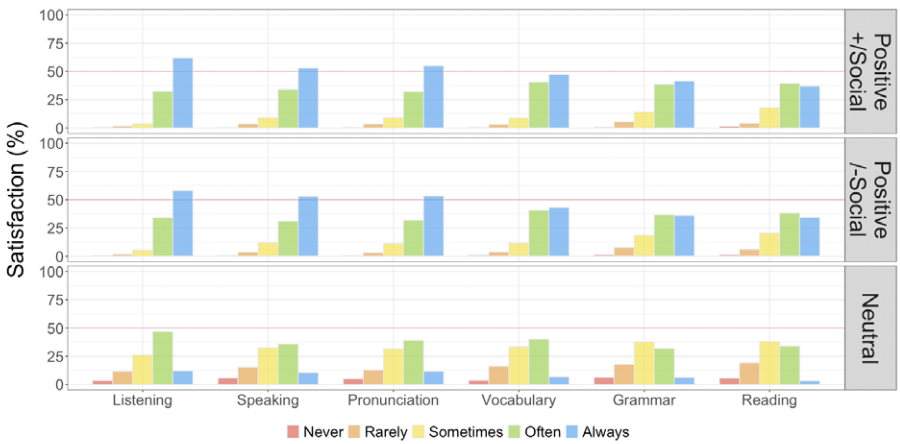

Figure 8. Satisfaction with L2 skill practice on the app.

app for at least 5 min in each single session. This threshold excluded 29 of the 3,319 participants (0.9%) from the frequency analysis. Figure 10 presents the descriptive statistics of usage frequency, with a horizontal line indicating the overall mean value (M = 4.80 times, SD = 3.05 times, *Range* = 1.50–12.83 times). On average, participants in Cluster 2 used the app 5.19 times per week (95% CI [5.00, 5.37]), those in Cluster 1 used the app 4.81 times per week (95% CI [4.65, 4.96]), and those in Cluster 3 used the app 4.11 times per week (95% CI [3.90, 4.32]). The nonoverlapping 95% CIs indicate that the differences in usage frequency among the clusters were significant. A linear regression confirmed that Cluster was a significant predictor of usage frequency (F(2, 3287) = 26.77, p < .001, r = .13). The effect size was small-to-medium at the level of single events but again potentially still consequential cumulatively speaking.

Based on the intensity and frequency findings, we established a threshold to determine a meaningfully active weekly level of MALL engagement for subsequent analysis of *duration*. An "active week" was defined as one in which a learner had at least

https://doi.org/10.1017/S0272263124000354 Published online by Cambridge University Press

Table 3. Summary of the MALL user clusters

Characteristics	MALL user cluster		
	Cluster 1 (*n* = 1,620)	Cluster 2 (*n* = 1,276)	Cluster 3 (*n* = 768)
Cluster label	Positive/+Social	Positive/–Social	Neutral
Preperformance expectancy	High	High	Medium
Effort expectancy	High	High	Medium
Social influence*	Medium–high	Low	Medium–low
Facilitating condition	Favorable	Favorable	Favorable
MALL attitude	Positive	Positive	Medium–positive
MALL acceptance	High	High	Medium
Main age*	20s, 30s	60s and above	20s, 30s
Reasons for L2 learning*	For self–satisfaction For communication For learning culture	For self–satisfaction	For self–satisfaction
Practice on the app	Frequent	Frequent	Sometimes
Satisfaction of the app	High	High	Medium

Note: * indicates variables on which differences between Clusters 1 and 2 were observed.

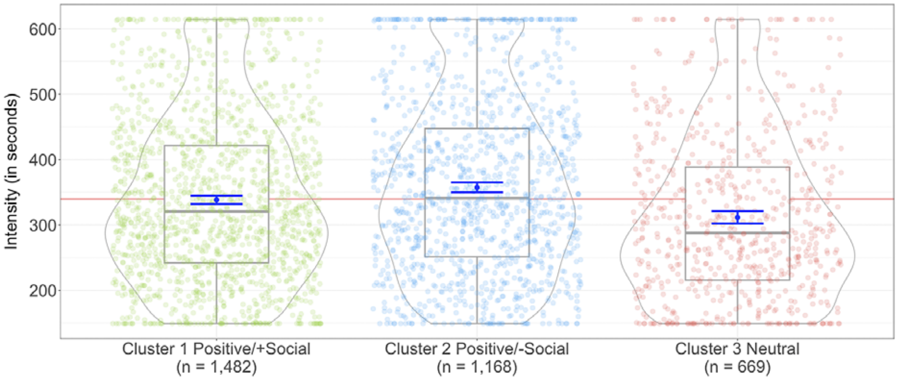

Figure 9. Descriptive statistics of usage intensity: Winsorized means and 95% CIs.

four individual usage sessions that lasted at least 5 min (i.e., 300 seconds) each. Duration is defined here as the maximum consecutive active weeks of MALL engagement, and this threshold excluded 428 of the 3,290 participants (13%) for the duration analysis. Figure 11 demonstrates the descriptive statistics of usage duration, with a horizontal line indicating the overall mean value (M = 7.69 weeks, SD = 9.42 weeks, *Range* = 1–36 weeks). Participants in Cluster 2 maintained an average continuous streak of usage spanning 9.32 weeks (95% CI [8.69, 9.95]), and those in Cluster 1 exhibited continuous app usage for an average of 7.28 weeks (95% CI [6.79, 7.77]). Participants in Cluster 3 had an average usage streak of 5.46 weeks (95% CI [4.80, 6.12]). The nonoverlapping 95% CIs indicate that the differences in usage duration among the three clusters were significant. A linear regression confirmed that Cluster was a significant predictor of usage duration ($F(2, 2859) = 32.12$, $p < .001$, $r = .15$). The effect size was small-to-medium at the level of single events but, again, potentially impactful in terms of cumulative language learning gains.

https://doi.org/10.1017/S0272263124000354 Published online by Cambridge University Press

Figure 10. Descriptive statistics of usage frequency: Winsorized means and 95% CIs.

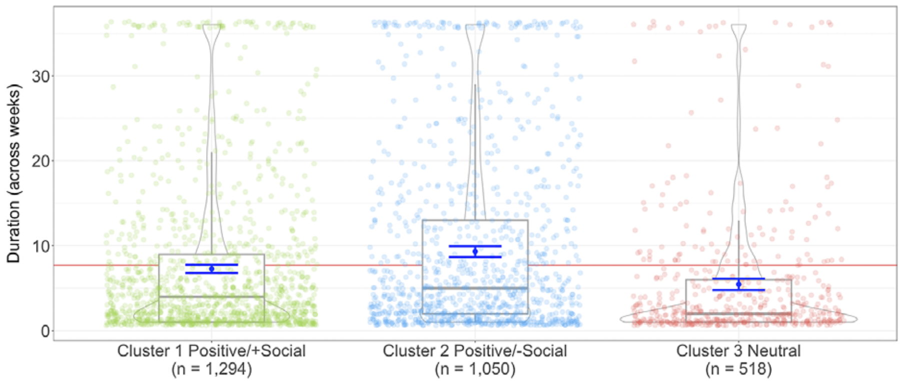

Figure 11. Descriptive statistics of usage duration: Winsorized means and 95% CIs.

To sum up, the learner groups with high levels of MALL acceptance showed greater MALL engagement—defined according to usage-related indices along different time scales—and this effect was cumulative over an extended period.

Analysis 2: Recurrent event analysis (survival analysis)

In order to investigate patterns of persistent MALL engagement, we built a regression model for recurrent event processes and a terminal event. Conceptually, recurrent events in this study refer to instances where learners paused their app use in a particular month but later resumed and continued using the app. The terminal event was a complete dropout, indicating that the learners discontinued using the app and did not return to it. For instance, a learner might use the app for 3 months, pause for 1 month, and resume using the app (i.e., the first recurrent event). The same learner can pause their app use again and resume once more (i.e., the second recurrent event). However, if the learner eventually stops using the app and does not return to it, the terminal event (i.e., a complete dropout) occurs. In this regard, we defined recurrent events as pauses in *active* monthly MALL activities and a terminal event as a dropout when individuals did

https://doi.org/10.1017/S0272263124000354 Published online by Cambridge University Press

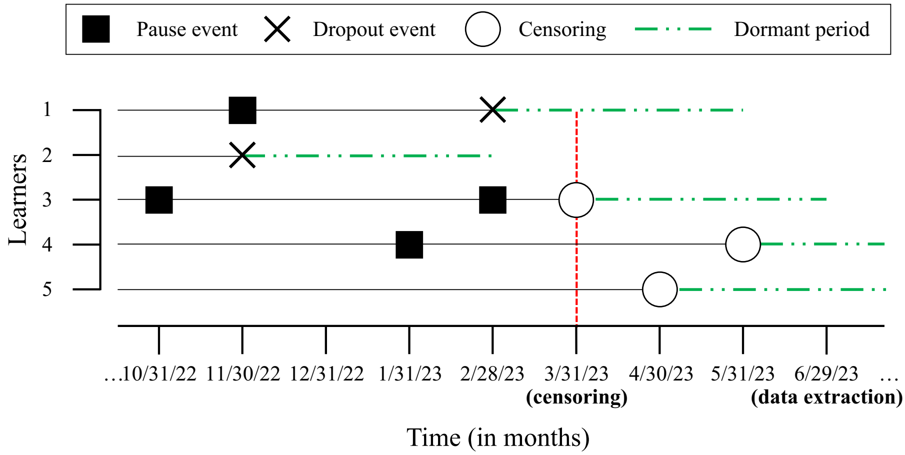

Figure 12. Schematic plot for censoring.

not resume *active* monthly MALL engagement within a dormancy period after the last active app usage.

Based on the findings summarized previously relating to intensity and frequency, active monthly MALL engagement in our study was operationalized as a month with a total amount of usage exceeding 80 min (equivalent to 5 min × 4 times per week × 4 weeks). Cluster 3 showed the highest rate of exclusion, with 204 of the 669 participants (30.5%), compared to the exclusion rate of Cluster 1 with 248 of the 1,482 participants (16.7%) and Cluster 2 with 140 of the 1,168 participants (12%). As a result, 2,727 of the 3,319 participants (82.1%) were included in survival analysis. Furthermore, we chose a 3-month duration for a *dormant* period in which a user is not actively using the app at the present time (i.e., app usage below the 80-min threshold) but could become active again later. This decision was based on the finding that the median time for individuals to resume their active monthly MALL engagement after a pause was 3 months.

In survival analysis, censoring is used for subjects whose target event (i.e., in this case, dropout) has not occurred by the conclusion of the study or period of interest. In the present study, learners were censored as of March 31, 2023, to indicate that their MALL engagement had not ended with a dropout by the time of the usage data extraction. As illustrated in Figure 12, learners 1 and 2 had a dropout experience (as indicated by ×) because they did not return to another active MALL engagement within 3 months (i.e., the dormant period) after their last active MALL engagement. On the other hand, learners 3, 4, and 5 were assumed to be still active on the app (as indicated by ◯) because their most recent active MALL engagement occurred by or beyond the censoring time, and their dormant period had not ended at the time of data extraction. As a result, 797 participants were identified to have dropped out, while 1,930 participants were still active as of June 29, 2023.

To assess the effects of different learner profiles on the likelihood of pause and dropout events, we built a joint frailty Cox-type regression model with Cluster as a covariate. This approach is especially beneficial for examining both recurrent and terminal events simultaneously for two reasons. First, the Cox-type model addresses the shortcomings of other common methods (e.g., the Andersen–Gill model), which require censoring to be independent of the recurrent event process. However, this independent censoring assumption can be violated when the recurrent event process (e.g., pauses) is correlated with the failure time (e.g., dropouts). Although frailty variables (i.e., unobserved random effects on time-to-event data) were introduced to allow for an association between the recurrent and terminal events, some conventional frailty models require a parametric assumption on the frailty distribution. In comparison, the (frailty) Cox-type model has the benefit of avoiding the independent censoring condition and parametric assumption, while inducing proportionate effects of covariates on the recurrent event process and the failure event over time (Chiou et al., 2023). Second, given that the recurrence(s) of the event and the failure time are both of interest, we chose the joint frailty Cox-type, proposed by Huang and Wang (2004) because it allows us to model the association between intensity of the recurrent event process and hazard of the failure time (e.g., higher rate of recurrent events are potentially more or less likely to experience a failure event). This joint modeling works via a two-step fashion by first estimating the value from the recurrent event data and then using the recurrent event information in the terminal event model (see Amorim & Cai, 2015; Chiou et al., 2023; Huang & Wang, 2004, for more detailed information on recurrent event analysis and model selections).

We built the joint frailty Cox-type model using the R package *reReg* with the argument *model* = *"cox|cox."* This approach allows for the simultaneous specifications of the rate (intensity) function and hazard function in the model, with each side of the '|' representing a Cox-type proportional model. The variance estimate was obtained by using a nonparametric bootstrap approach with 200 replications. We determined the hazard ratios (i.e., the relative risk of the event of interest occurring in one group compared to another) by exponentiating the estimated regression coefficients. Typically, a hazard ratio greater than 1 indicates an increased risk of the event occurring in one group compared to a reference group, while a hazard ratio less than 1 indicates a decreased risk. Cluster 3 was used as the reference group because of its relatively neutral stance toward MALL.

Results 2

RQ3. Persistent learners showed more frequent pauses in their app usage

Figure 13 displays the distribution of the first year of app adoption by cluster. A modest majority of learners, about 55.40% (1,511 of 2,727), started using the app in or after 2020. The patterns of frequency of app adoption per year were similar across the three clusters.

Table 4 shows the descriptive statistics of the recurrent events and terminal events by cluster. Although Clusters 1 and 2 had more recurrent pauses in active monthly MALL engagement than Cluster 3 during the given follow-up period (i.e., the time between the first and last active usage months), these two groups experienced lower dropout rates than Cluster 3. It is noteworthy that Clusters 1 and 2 maintained longer continuous active months and had shorter dormancy periods before resuming active app usage than Cluster 3. Additionally, the period between the first and last active months was also longer for Clusters 1 and 2, indicating more consistent efforts of these two groups to return to the app (i.e., resilience against dropout), further evidenced by their longer span before a dropout than Cluster 3. However, as demonstrated in Figure 14, it should be highlighted that about 31% of Cluster 1 (99 of 317), about 29% of Cluster 2 (79 of 274), and about 37% of Cluster 3 (77 of 206), dropped out after their first month of active MALL engagement. Given that about 43.16% of the learners (344 of 797)

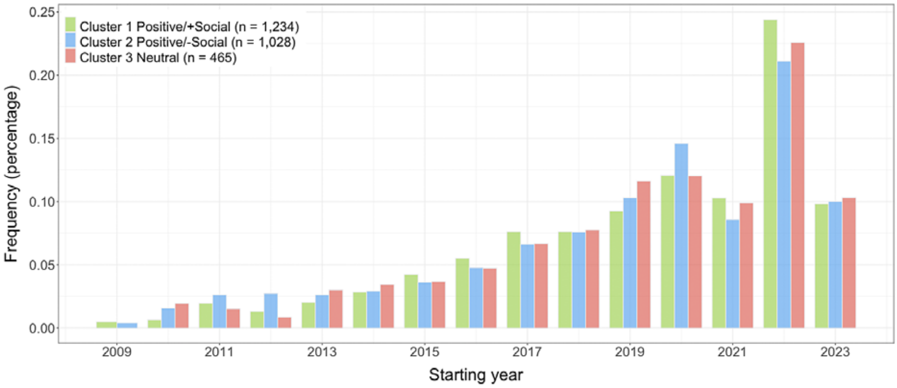

Figure 13. Frequency of app adoption by year.

Table 4. Descriptive statistics of recurrent events and terminal events

Variables	Cluster 1 (*n* = 1,234)	Cluster 2 (*n* = 1,028)	Cluster 3 (*n* = 465)
Median time between the first and last active engagement (months)	14	18	9
Mean number of pauses observed[a]	3.28	3.53	2.74
Number of individuals who dropped out	317 (25.68%)	274 (26.65%)	206 (44.30%)
Median active engagement streak (months)	2	2.5	1.75
Median length of dormancy period[a] (months)	5	5.33	6.86
Median time leading up to a dropout[a] (months)	7	9	5

Note: [a] indicates that only individuals who experienced the target event at least once were included in the corresponding analysis.

dropped out of the app within the first 3 months, the early stage of MALL was particularly susceptible to massive attrition regardless of the learner types.

Table 5 presents the results of the joint Cox-type regression model. The top panel of the summary table indicates statistically significant positive effects of Cluster on the rate function of the recurrent event process, 0.26 for Cluster 1 and 0.23 for Cluster 2. This suggests that Clusters 1 and 2 experience more frequent pauses throughout the follow-up period than Cluster 3. The hazard ratios confirm that Clusters 1 and 2 have 30% (95% CI [15%, 47%]) and 26% (95% CI [12%, 43%]) increased risk of experiencing pauses, respectively, compared to Cluster 3. In addition, the bottom panel of the summary table shows statistically significant negative effects of Cluster on the dropout function, −0.60 for Cluster 1 and −0.63 for Cluster 2. This suggests that Clusters 1 and 2 are at a lower risk of dropping out than Cluster 3. The hazard ratios confirm that Clusters 1 and 2 have 45% (95% CI [35%, 54%]) and 47% (95% CI [37%, 55%]) reduced risk of experiencing a dropout compared to Cluster 3. In sum, these findings are consistent with the observed pattern in the earlier descriptive statistics that the more frequent pauses in active app engagement were associated with fewer dropouts, ultimately leading to longer periods of MALL engagement overall (i.e., persistence).

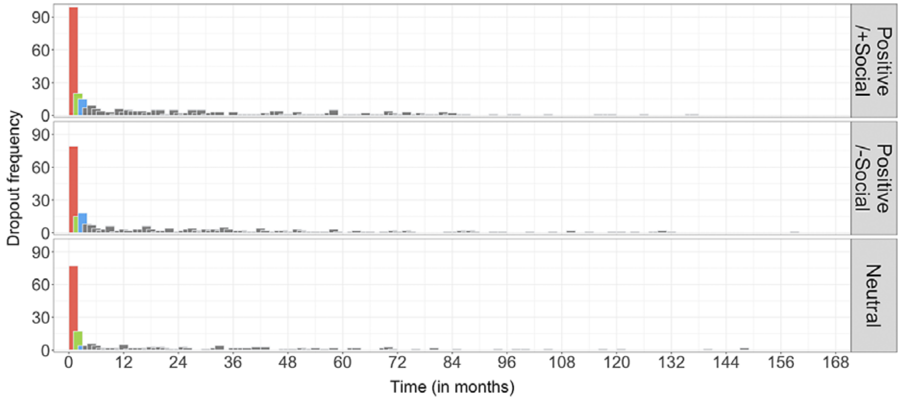

Figure 14. Frequency of dropout over time with the first 3 months highlighted.

Table 5. Results of the joint Cox-type regression model.

(a) Recurrent event process							
			95% CIs				
	Estimate (coefficient)	Hazard ratio	LCI	UCI	*SE*	*z*	*p*
Cluster 1	0.26	1.30	1.15	1.47	0.06	4.17	<.001
Cluster 2	0.23	1.26	1.12	1.43	0.06	3.77	<.001
(b) Terminal event							
			95% CIs				
	Estimate (coefficient)	Hazard ratio	LCI	UCI	*SE*	*z*	*p*
Cluster 1	−0.60	0.55	0.46	0.65	0.09	−7.11	<.001
Cluster 2	−0.63	0.53	0.45	0.63	0.09	−7.30	<.001

Note: Cluster 3 as the reference group.

Discussion

In the study, we identified three distinct UTAUT-based MALL learner profiles. Clusters 1 and 2 showed considerably higher levels of performance expectancy, effort expectancy, and positive MALL attitude, while Cluster 3 displayed relatively neutral responses to these variables. Such group differences were triangulated by additional measures of MALL acceptance. A strong intention to use the app for L2 learning in Clusters 1 and 2 was evidenced by their higher frequency of L2 skill practice and satisfaction in contrast to the more cautious and reserved stance of Cluster 3. More importantly, different levels of MALL acceptance were further linked to engagement patterns. Despite a large dropout rate in the first 3 months for all three groups, Clusters 1 and 2 displayed more intense (daily), frequent (weekly), and durable (monthly) MALL engagement compared to Cluster 3. It is also important to highlight that the higher instances of pausing and restarting observed in Clusters 1 and 2 were associated with their greater persistence in MALL engagement overall.

https://doi.org/10.1017/S0272263124000354 Published online by Cambridge University Press

Context-dependent UTAUT variables for MALL engagement

Although prior studies clustered L2 learners based on similarities in their MALL experiences (e.g., Lamb & Arisandy, 2020; Peng et al., 2022), we adopted the UTAUT as a theoretical framework to identify distinct learner profiles in relation to their MALL acceptance. The results of a PCA confirmed that the collective influence of the four UTAUT variables (i.e., performance expectancy, effort expectancy, MALL attitude, social influence) accounted for about 69.9% of the variance, suggesting the applicability of the UTAUT for future studies into profiling learners in technology-enhanced L2 learning contexts.

Furthermore, we explored MALL with a focus on a diverse group of adult L2 learners in informal settings, a shift from previous studies that largely centered on university students (Godfroid & Andringa, 2023; Plonsky, 2023). Such contextual differences might influence the extent to which each UTAUT variable could affect L2 learners' intention to use apps. Indeed, previous studies did not find clear influences of performance expectancy and effort expectancy on technology acceptance for L2 learning (García Botero et al., 2018; Hoi, 2020; Hsu, 2023; Zou et al., 2022). However, our results suggest that both performance expectancy and effort expectancy coupled with MALL attitude can play an important role in MALL acceptance, while the impacts of social influence and facilitating conditions were less pronounced. Given that 78.4% of our participants already held at least a 4-year degree, their L2 learning settings might differ from those of currently enrolled university students (cf. Puebla & García, 2022). L2 learners in educational institutions might have access to more diverse L2 learning resources (e.g., regular language courses provided by the university) or be driven by external regulations (e.g., grades, course requirements). In contrast, our participants primarily learned languages for intrinsic reasons (e.g., self-satisfaction, communication, cultural awareness). That is, it is possible that the participants in our study were already intrinsically motivated, thus neutralizing the influence of external regulatory (e.g., rewards) or introjected factors (e.g., approval from others) from social contexts. Instead, this might heighten the importance of performance and effort expectancy when deciding to engage in MALL, as these learners are likely to seek tools that align with their higher-order motivations and provide efficient and enjoyable learning experiences. These different MALL contexts can influence how UTAUT variables can impact MALL engagement.

Interestingly, the fifth UTAUT variable, facilitating conditions, did not stand out as a pivotal variable in discriminating learner groups because all three clusters scored similarly high in this area. This might be because a wider proliferation of smart devices and easier internet access in general has reduced the significance of facilitating conditions compared to the early 2000s when the UTAUT model was first introduced. However, the importance of facilitating conditions should not be overlooked because possessing smart devices and stable internet connectivity is still essential for MALL activities to be possible. Consequently, unfavorable learning conditions could potentially impede MALL, particularly in contexts where technologies (i.e., smartphones, tablets) are not universally or reliably accessible, with downstream effects on MALL acceptance and engagement. These findings may therefore highlight the context-dependent nature of MALL engagement (cf. Straub, 2009) and the need to examine these questions in contexts that are less conducive to MALL.

MALL acceptance and multidimensional engagement patterns

We found an association between the acceptance of MALL and actual engagement. In essence, individuals who perceive the usefulness of apps, find them easy to use, and have a positive attitude toward MALL are more likely to engage in MALL activities. Individuals with high levels of MALL acceptance (i.e., Clusters 1 and 2) showed a higher frequency of L2 skill practice and greater and more sustained engagement than those who showed relatively less certainty about MALL (i.e., Cluster 3). It is noteworthy that at daily and weekly levels, on average, Clusters 1 and 2 spent only 30 to 50 seconds more time on the app per session and showed one more instance of active app usage per week than Cluster 3. Nonetheless, such seemingly minor differences had a cumulative impact over an extended period. Indeed, Cluster 3 maintained active app usage for a continuous span of 5 weeks, but Clusters 1 and 2 extended their active engagement beyond 7 weeks. This significant difference at the monthly level led to lower dropout rates on an annual scale. This observation serves to provide additional confirmation of the fact that small effect sizes for single events can ultimately have meaningful consequences (Funder & Ozer, 2019). In this regard, it becomes crucial to discuss realistic benchmarks for the evaluation of effect sizes in longitudinal and informal L2 learning contexts.

When comparing Clusters 1 and 2, it is important to note that Cluster 1 had somewhat lower engagement levels than Cluster 2. However, this should not overshadow the role of important others (e.g., teachers, family, friends) in MALL engagement (Noels et al., 2019). It is possible that Cluster 1's higher perception of social influence led them to be more communication-focused than Cluster 2. In fact, Cluster 1 used more extra apps (1.17 additional apps) for target language learning compared to Cluster 2 (0.93 additional apps), indicating Cluster 2 relied on a single primary app. Given that Cluster 1 might use a combination of multiple apps to build L2 proficiency, their MALL engagement might be distributed across multiple apps, potentially resulting in lower intensity, frequency, and duration on Mango in this study. When considering data across all apps, Cluster 1's engagement might very well surpass that of Cluster 2. Further studies considering engagement across multiple apps might show the importance of social influence in facilitating MALL engagement in informal contexts.

Recurrent pauses and persistence in MALL

Although all three clusters experienced large attrition in the first 3 months of their active MALL engagement, one important finding is that Clusters 1 and 2 showed more frequent pauses in their active monthly MALL engagement than Cluster 3. This observation might initially appear counterintuitive, as increased instances of disengagement could be indicative of lower persistence. However, together with their high levels of MALL acceptance and belief in the importance of Mango as a primary source of L2 learning, such recurrent pauses may signify *resilience* against dropout and consistent efforts to resume and sustain MALL engagement (cf. Kim & Kim, 2017). Indeed, the fewer pauses observed in Cluster 3 suggest that once the flow of MALL engagement was disrupted, they were more prone to complete dropout, rather than attempting to resume active app usage. This could be due to their lower level of MALL acceptance. In other words, more frequent pauses showed greater reengagement effects, enhanced by high MALL acceptance, leading to more persistent MALL engagement. This interpretation is further corroborated by the fact that Clusters 1 and 2 exhibited longer periods of active and continuous app usage but shorter dormancy periods than Cluster 3. Our findings highlight that persistence in MALL engagement is not just about the sum of time spent on apps (Kessler et al., 2023; Loewen et al., 2019) or future intention alone (Feng & Papi, 2020; Jung & Lee, 2018; Lakhal et al., 2021). Instead,

https://doi.org/10.1017/S0272263124000354 Published online by Cambridge University Press

persistence is a multidimensional *process* involving cyclical phases of engagement, disengagement, dormancy, and reengagement, with each aspect, like intensity, frequency, and duration, building up cumulatively over time. Importantly for future research and theory construction, persistence, when understood this way, is relevant to instructed and self-directed L2 learning in general and not limited to MALL only. Therefore, we call for further investigations into persistence operationalized as a multidimensional process and how it impacts actual L2 development, both in L2 learning broadly and for technology-enhanced L2 learning, specifically.

Implications for language learning app development

Recognizing the tendency of persistent learners to experience greater recurrent pauses in MALL engagement, it is important to create conditions for L2 learners to conveniently resume app use. In this regard, the present study provides implications for app companies aiming to enhance learners' MALL engagement. First of all, it could be more effective to provide learners with multiple short learning modules instead of a single extended lesson, as Sudina and Plonsky (2024) suggested. For instance, in informal L2 learning contexts, learners who intend to complete a 15-min lesson might feel obliged to finish the entire session once initiated, perceiving the task as onerous because they need to stay focused for 15 min uninterrupted. This sense of pressure could potentially prevent learners from initiating the learning at all. Or, if they start but do not finish, the unfinished session might undermine their sense of accomplishment. In contrast, if the same 15-min lesson were presented as three distinct 5-min modules (see intensity data in Results 1), learners would gain increased flexibility in determining when to engage (e.g., using "found time," in transit or between other obligations throughout the day). This would also allow learners more clear stopping points so that they can resume learning on a new module, rather than in the middle of a lesson. Consequently, distributing the learning across multiple manageable modules will lower (physical and psychological) hurdles to learning and provide more opportunities to feel a sense of accomplishment, encouraging learners to maintain their engagement momentum.

Additionally, apps can send notifications that provide learners with brief, encouraging prompts to resume their learning, particularly during the first 3 months, which is the critical period for avoiding complete dropout. This notification could take the form of quick, fun activities such as daily vocabulary challenges or quizzes that feature intriguing questions, which could entice learners back into the app without overwhelming them with time-consuming tasks or negative pressure. Rewards or incentives could be provided for learners who consistently return to the app after a pause, such as unlocking bonus content or accessing exclusive features. By incorporating such strategies, app developers can create an environment that encourages learners to seamlessly reinitiate app use after pauses, promoting persistent engagement in the long run.

Ultimately, the core finding of this study is that app users' dispositions toward technology are influenced by a multitude of internal and external variables and that these dispositions have significant cumulative impacts on whether or not individuals persist in MALL engagement. As such, understanding, at least at a basic level, individuals' dispositional profiles may allow both MALL providers and instructed contexts that integrate with MALL the ability to maximize engagement and avoid user dropout.

Limitations and future directions

The present study suggests multiple potential avenues for future research. First, there is presumably a sampling bias in the present study. Learners in Cluster 3, which represents relatively low levels of engagement in our sample, still used the app for roughly 20 min per week (i.e., 5 min per usage and four times a week). In this regard, we acknowledge that these learners do not represent those with *low* levels of engagement. Given this potential sampling bias, it is important to directly explore MALL acceptance levels and engagement patterns of low-usage individuals because they might be vulnerable to much lower engagement levels and higher dropout rates in informal (MALL) L2 learning settings. The second suggestion is that (Instructed) SLA and MALL researchers in different contexts may benefit from the establishment of specific operationalizations of active engagement and persistence. In the study, we applied various active app usage thresholds (e.g., 300 seconds per usage, four active usage sessions per week) based on the mean/median values of our sampled data. However, different populations in various contexts (e.g., young adolescents in school settings) might need different activeness thresholds that can determine overall persistence. In future research, it is important to both provide context-specific operationalizations of persistence (e.g., on daily, weekly, monthly scales) and explicitly justify thresholds applied to relevant engagement indices (e.g., intensity, frequency, duration, pauses, dormancy, dropouts in our study). Timely, transparent reporting practice can enhance comparability and generalizability across different persistence studies. Third, the present study relied on quantitative survey and usage data. In order to enhance our understanding of the nature of MALL engagement in informal settings, it would be beneficial to incorporate more qualitative approaches. To this end, our team has recruited individuals from each cluster for follow-up interviews to further probe MALL persistence qualitatively. A final and important limitation of the present study is that no language learning (i.e., outcome, achievement) data were collected or analyzed. Future studies could deepen our collective understanding of the relationship between longitudinal persistence in MALL, the factors that contribute to differing levels of persistence, and the achievement levels reached by individuals and groups with varying levels of persistence, by collecting L2 learning evidence as well.

Conclusion

We investigated 3,670 adult L2 learners' dispositions toward MALL using the UTAUT framework. We additionally triangulated their app usage experience self-report with in-app behavioral data, providing specific operationalizations of engagement across multiple indices and time scales. Bottom-up cluster analysis revealed multiple learner-type groups based on the UTAUT variables, and these groups were found to engage differently in MALL across time. This study offers important implications for conceptualizing language learning engagement and persistence, and it offers a novel analytical method (i.e., survival analysis) to understand language learner behavior over time. Given the diverse nature of language learners in instructed and self-directed contexts as well as their differential access to resources and reasons for learning additional languages, this study constitutes an important foundation for future explorations of learner persistence, triangulating multiple data sources for richer and more informative conclusions to inform theory and (technology-enhanced) language learning offerings. Furthermore, understanding learner persistence is an important step toward keeping learners engaged in L2 study, which can lead to increased L2 ability—the primary goal of many MALL learners.

https://doi.org/10.1017/S0272263124000354 Published online by Cambridge University Press

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/S0272263124000354.

Acknowledgments. This research was presented at the American Association for Applied Linguistics (AAAL) 2024 in Houston (USA). We are grateful to Anthony Lizan, Madina Dokhova, and Olúwasèyí Fásunhàn for their input as the project was developed. We thank *SSLA* editor Dr. Luke Plonsky and the three anonymous reviewers for their constructive feedback. Any remaining shortcomings in the manuscript are ours.

Competing interest. This project was conducted as an academic-industry partnership. Mango Languages (the app of focus of the study) contributed access to user data for the project. While Kaitlyn M. Tagarelli (the fourth author) is an employee of Mango Languages, neither the research team nor participants were compensated in any way for the study, nor did Mango impose any restrictions or otherwise influence the reporting of the study. Therefore, we have no conflicts of interest to report beyond Kaitlyn's employment at Mango Languages.

References

Amorim, L. DAF., & Cai, J. (2015). Modelling recurrent events: A tutorial for analysis in epidemiology. *International Journal of Epidemiology*, *44*(1), 324–333. https://doi.org/10.1093/ije/dyu222

An, X., Chai, C. S., Li, Y., Zhou, Y., & Yang, B. (2023). Modeling students' perceptions of artificial intelligence assisted language learning. *Computer Assisted Language Learning*, Advance online publication. https://doi.org/10.1080/09588221.2023.2246519

Andringa, S., & Godfroid, A. (2020). Sampling bias and the problem of generalizability in applied linguistics. *Annual Review of Applied Linguistics*, *40*, 134–142. https://doi.org/10.1017/S0267190520000033

Burston, J., & Giannakou, K. (2021). MALL language learning outcomes: A comprehensive meta-analysis 1994–2019. *ReCALL*, *34*, 147–168. https://doi.org/10.1017/S0958344021000240

Chiou, S. H., Xu, G., Yan, J., & Huang, C. Y. (2023). Regression modeling for recurrent events possibly with an informative terminal event using R package reReg. *Journal of Statistical Software*, *105*, 1–34. https://doi.org/10.18637/jss.v105.i05

Crowther, D., Kim, S., Lee, J., Lim, J., & Loewen, S. (2021). Methodological synthesis of cluster analysis in second language research. *Language Learning*, *71*(1), 99–130. https://doi.org/10.1111/lang.12428

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–339. https://doi.org/10.2307/249008

Feng, L., & Papi, M. (2020). Persistence in language learning: The role of grit and future self-guides. *Learning and Individual Differences*, *81*, Article 101904. https://doi.org/10.1016/j.lindif.2020.101904

Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, *2*, 156–168. https://doi.org/10.1177/2515245919847202

García Botero, G., Botero Restrepo, M. A., Zhu, C., & Questier, F. (2021). 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., Cincinnato, S., He, T., & Zhu, C. (2018). Acceptance and usage of mobile assisted language learning by higher education students. *Journal of Computing in Higher Education*, *30*, 426–451. https://doi.org/10.1007/s12528-018-9177-1

García Botero, G., Questier, F., & Zhu, C. (2019). 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

Godfroid, A., & Andringa, S. (2023). Uncovering sampling biases, advancing inclusivity, and rethinking theoretical accounts in second language acquisition: Introduction to the special issue SLA for All? *Language Learning*, *73*(4), 981–1002. https://doi.org/10.1111/lang.12620

Granić, A., & Marangunić, N. (2019). Technology acceptance model in educational context: A systematic literature review. *British Journal of Educational Technology*, *50*(5), 2572–2593. https://doi.org/10.1111/bjet.12864

Guggemos, J., Seufert, S., & Sonderegger, S. (2020). Humanoid robots in higher education: Evaluating the acceptance of Pepper in the context of an academic writing course using the UTAUT. *British Journal of Educational Technology*, *51*(5), 1864–1883. https://doi.org/10.1111/bjet.13006

Hiver, P., Al-Hoorie, A. H., & Mercer, S. (Eds.). (2021). *Student engagement in the language classroom.* Multilingual Matters. https://doi.org/10.21832/9781788923613

Hoi, V. N. (2020). Understanding higher education learners' acceptance and use of mobile devices for language learning: A Rasch-based path modeling approach. *Computers & Education*, *146*, Article 103761. https://doi.org/10.1016/j.compedu.2019.103761

Hoi, V. N., & Mu, G. M. (2021). Perceived teacher support and students' acceptance of mobile-assisted language learning: Evidence from Vietnamese higher education context. *British Journal of Educational Technology*, *52*(2). 879–898. https://doi.org/10.1111/bjet.13044

Hsu, H.-T., & Lin, C.-C. (2021). Extending the technology acceptance model of college learners' mobile-assisted language learning by incorporating psychological construct. *British Journal of Educational Technology*, *53*(2), 286–306. https://doi.org/10.1111/bjet.13165

Hsu, L. (2023). EFL learners' self-determination and acceptance of LMOOCs: The UTAUT model. *Computer Assisted Language Learning*, *36*(7), https://doi.org/10.1080/09588221.2021.1976210

Huang, C.-Y., & Wang, M.-C. (2004). Joint modeling and estimation for recurrent event processes and failure time data. *Journal of the American Statistical Association*, *99*(468), 1153–1165. https://doi.org/10.1198/016214504000001033

Hui, B. (2020). Processing variability in intentional and incidental word learning: An extension of Solovyeva and DeKeyser (2018). *Studies in Second Language Acquisition*, *42*(2), 327–357. https://doi.org/10.1017/S0272263119000603

Hwang, H., Coss, M., Loewen, S., & Tagarelli, K. (2024a). *MALL experience and UTAUT disposition questionnaire. Questionnaire from "Acceptance and engagement patterns of mobile-assisted language learning among non-conventional adult L2 learners: A survival analysis"* [Text/Materials]. IRIS Database, University of York, UK. https://doi.org/10.48316/kZMGP-5hAJp

Hwang, H., Coss, M., Loewen, S., & Tagarelli, K. (2024b). *MALL engagement survey and index analysis R code. Code from "Acceptance and engagement patterns of mobile-assisted language learning among non-conventional adult L2 learners: A survival analysis"* [Software]. IRIS Database, University of York, UK. https://doi.org/10.48316/eVFp4-hgS1m

Jeon, J. (2022). Exploring a self-directed interactive app for informal EFL learning: A self-determination theory perspective. *Education and Information Technologies*, *27*, 5767–5787. https://doi.org/10.1007/s10639-021-10839-y

Jung, Y., & Lee, J. (2018). Learning engagement and persistence in massive open online courses (MOOCS). *Computers & Education*, *122*, 9–22. https://doi.org/10.1016/j.compedu.2018.02.013

Kessler, M., Loewen, S., & Gönülal, T. (2023). Mobile-assisted language learning with Babbel and Duolingo: Comparing L2 learning gains and user experience. *Computer Assisted Language Learning.* Advance online publication. https://doi.org/10.1080/09588221.2023.2215294

Khechine, H., Raymond, B., & Augier, M. (2020). The adoption of a social learning system: Intrinsic value in the UTAUT model. *British Journal of Educational Technology*, *51*(6). 2306–2325. https://doi.org/10.1111/bjet.12905

Kim, T.-Y., & Kim, Y.-K. (2017). The impact of resilience on L2 learners' motivated behavior and proficiency in L2 learning. *Educational Studies*, *43*(1), 1–15. https://doi.org/10.1080/03055698.2016.1237866

Lai, C. (2013). A framework for developing self-directed technology use for language learning. *Language Learning & Teaching*, *17*(2), 100–122. http://doi.org/10125/44326

Lai, Y., Saab, N., & Admiraal, W. (2022). University students' use of mobile technology in self-directed language learning: Using the integrative model of behavior prediction. *Computers & Education*, *179*, Article 104413. https://doi.org/10.1016/j.compedu.2021.104413

Lakhal, S., Khechine, H., & Mukamurera, J. (2021). Explaining persistence in online courses in higher education: A difference-in-differences analysis. *International Journal of Educational Technology in Higher Education*, *18*, Article 19. https://doi.org/10.1186/s41239-021-00251-4

Lamb, M., & Arisandy, F. E. (2020). The impact of online use of English on motivation to learn. *Computer Assisted Language Learning*, *33*(1-2). https://doi.org/10.1080/09588221.2018.1545670

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

Loewen, S., Isbell, D. R., & Sporn, Z. (2020). The effectiveness of app-based language instruction for developing receptive linguistic knowledge and oral communicative ability. *Foreign Language Annals*, *53* (2), 209–233. https://doi.org/10.1111/flan.12454

Mair, P., & Wilcox, R. (2020). Robust statistical methods in R using the WRS2 package. *Behavior Research Methods*, *52*, 464–488. https://doi.org/10.3758/s13428-019-01246-w

Nielson, K. B. (2011). Self-study with language learning software in the workplace: What happens? *Language Learning & Technology*, *15*(3), 110–129. http://doi.org/10125/44265

Noels, K. A., Lou, N. M., Lascano, D. I. V., Chaffee, K. E., Dincer, A., Zhang, Y. S. D., & Zhang, X. (2019). Self-determination and motivated engagement in language learning. In M. Lamb, K. Csizér, A. Henry, & S. Ryan (Eds.), *The Palgrave handbook of motivation for language learning* (pp. 95–116). Palgrave Macmillan.

Peng, H., Jager, S., & Lowie, W. (2022). A person-centered approach to L2 learners' informal mobile language learning. *Computer Assisted Language Learning*, *35*(9), 2148–2169. https://doi.org/10.1080/09588221.2020.1868532

Plank, S. B., DeLuca, S., & Estacion, A. (2008). High school dropout and the role of career and technical education: A survival analysis of surviving high school. *Sociology of Education*, *81*(4), 345–370. https://doi.org/10.1177/003804070808100402

Plonsky, L. (2023). Sampling and generalizability in Lx research: A second-order synthesis. *Languages*, *8*(1), 75. https://doi.org/10.3390/languages8010075

Puebla, C., & García, J. (2022). Advocating the inclusion of older adults in digital language learning technology and research: Some considerations. *Bilingualism: Language and Cognition*, *25*, 398–399. https://doi.org/10.1017/S1366728921000742

Reinders, H., & Benson, P. (2017). Research agenda: Language learning beyond the classroom. *Language Teaching*, *50*(4), 561–578. https://doi.org/10.1017/S0261444817000192

Reinders, H., Lee, B. J., & Bonner, E. (2023). Tracking learner engagement in the L2 classroom with experience sampling. *Research Methods in Applied Linguistics*, *2*(2), Article 100052. https://doi.org/10.1016/j.rmal.2023.100052

Rumberger, R. W., & Rotermund, S. (2012). The relationship between engagement and high school dropout. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 491–513). Springer.

Stockwell, G. (Ed.). (2022). *Mobile assisted language learning: Concepts, contexts and challenges*. Cambridge University Press. https://doi.org/10.1017/9781108652087

Straub, E. T. (2009). Understanding technology adoption: Theory and future directions for informal learning. *Review of Educational Research*, *79*(2), 625–649. https://doi.org/10.3102/0034654308325896

Sudina, E., & Plonsky, L. (2024). The effects of frequency, duration, and intensity on L2 learning through Duolingo: A natural experiment. *Journal of Second Language Studies*, *7*(1), 1–4. https://doi.org/10.1075/jsls.00021.plo

van Berkel, N., Ferreira, D., & Kostakos, V. (2018). The experience sampling method on mobile devices. *ACM Computing Surveys*, *50*(6), 93. https://doi.org/10.1145/3161158

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, *27*(3), 425–478. https://doi.org/10.2307/30036540

Wilcox, R. (2005). Trimming and winsorization. In P. Armitage & T. Colton (Eds.), *Encyclopedia of biostatistics* (2nd ed). Wiley. https://doi.org/10.1002/0470011815.b2a15165

Winke, P., & Gass, S. M. (Eds.). (2019). *Foreign language proficiency in higher education*. Springer.

Zhang, D., Hennessy, S., & Pérez-Paredes, P. (2023). An investigation of Chinese EFL learners' acceptance of mobile dictionaries in English language learning. *Computer Assisted Language Learning*. Advance online publication. https://doi.org/10.1080/09588221.2023.2189915

Zou, C., Li, P., & Jin, L. (2022). Integrating smartphones in EFL classrooms: Students' satisfaction and perceived learning performance. *Education and Information Technologies*, *27*, 12667–12688. https://doi.org/10.1007/s10639-022-11103-7

Cite this article: Hwang, H.-B., Coss, M. D., Loewen, S., & Tagarelli, K. M. (2024). Acceptance and engagement patterns of mobile-assisted language learning among non-conventional adult L2 learners: A survival analysis. *Studies in Second Language Acquisition,* 1–27. https://doi.org/10.1017/S0272263124000354