



SCHOOL OF
ECONOMICS AND
MANAGEMENT

What drives language learners' to keep using Duolingo?

How do perceived learning value and perceived enjoyment influence learners' continuance intention to use Duolingo through satisfaction and habit, and how does relative challenge moderate these relationships?

by

Siti Humaira Binti Idrus
Hanyu Zheng

May 2026

Word Count: 24,852

Master's Programme in International Marketing and Brand Management

Supervisor: Javier Cenamor
Examiner: Burak Tunca

Abstract

This study investigates continuance intention to use Duolingo from a post-adoption perspective, drawing primarily on the Expectation-Confirmation Model (Bhattacharjee, 2001). In gamified mobile language learning, value is not achieved through initial adoption alone, but through repeated use that supports meaningful learning progress. Therefore, this study examines why active users intend to continue using Duolingo after prior app experience.

The study argues that perceived learning value and perceived enjoyment operate as a utilitarian–hedonic hybrid in gamified learning environments. These two app perceptions are proposed to influence continuance intention through two parallel mechanisms: satisfaction, which represents the evaluative pathway, and habit, which represents the behavioural pathway. Relative challenge, informed by flow theory, is included as a moderator to examine whether perceived learning difficulty changes these relationships.

A quantitative cross-sectional survey was conducted among active Duolingo users. After data screening, 186 valid responses were analysed using Partial Least Squares Structural Equation Modelling (PLS-SEM) in SmartPLS.

The findings show that satisfaction is the strongest mechanism explaining continuance intention, highlighting the importance of users' positive evaluation of the app rather than repeated use alone. Relative challenge operates selectively by strengthening the relationship between perceived enjoyment and satisfaction, as well as the indirect pathway from perceived enjoyment to continuance intention through satisfaction. The robustness check shows that streak significantly predicts continuance intention. However, unlike habit, streak is not significantly driven by perceived learning value, suggesting that psychological habit and visible behavioural records represent different forms of repeated use.

The study contributes to continuance research by extending the utilitarian–hedonic hybrid perspective to gamified mobile language learning. Practically, the findings suggest that gamified learning platforms should not focus only on making users return, but on making repeated use feel enjoyable, educationally meaningful, and connected to visible learning progress.

Acknowledgements

We would like to thank our supervisor, Javier Cenamor, for his guidance, constructive feedback, and support during the development of this thesis. His comments and supervision helped us strengthen the direction, structure, and academic quality of the study.

Finally, we would like to thank everybody who contributed in any way to this thesis, including those who participated in the survey, offered feedback, shared encouragement, or supported us during the writing process. Your help and kindness are deeply appreciated.

Table of Contents

Abstract.....	1
Acknowledgements.....	2
1 Introduction.....	7
1.1 Background.....	7
1.2 Research Problem.....	11
1.3 Research Gap and Research Question.....	15
1.4 Research Objective.....	19
2 Literature Review.....	20
2.1 The Utilitarian–Hedonic Hybrid In Gamified Learning Platforms.....	20
2.1.1 From Digital Education To Gamified Mobile Learning.....	21
2.2 Satisfaction And Habit As Post-Adoption Mechanisms.....	25
2.3 Flow Experience In Educational Psychology.....	28
3 Theoretical Framework and Hypothesis Development.....	32
3.1 Continuance Intention in Technology Post-Adoption.....	32
3.1.1 Foundations Of Consumer Behavioural Intention.....	32
3.1.2 Technological Acceptance And System Effectiveness.....	33
3.1.3 Pivoting To Post-Adoption.....	34
3.2 Hypothesis Development.....	37
3.2.2 Satisfaction As An Evaluative Post-Adoption Mechanism.....	37
3.2.3 Habit As A Behavioural Post-Adoption Mechanism.....	39
3.2.4 Relative Challenge As Learning Condition.....	40
3.2.5 Proposed research model.....	44
4 Methods.....	46
4.1 Research Philosophy and Approach.....	46
4.2 Research Design.....	47
4.3 Population and Sampling.....	48
4.4 Measurement of Constructs.....	51
4.4.1 Independent Variables: Perceived Learning Value and Perceived Enjoyment.....	52
4.4.2 Mediating Variables: Satisfaction and Habit.....	52
4.4.3 Moderating Variable: Relative Challenge.....	53

4.4.4 Dependent Variable: Continuance Intention.....	53
4.4.5 Streak as a Robustness-Check Variable.....	54
4.5 Data Collection.....	56
4.6 Data Analysis Process.....	57
4.7 Limitations and Mitigations.....	58
5 Findings and Results.....	59
5.1 Descriptive Statistics.....	59
5.2 Measurement Model Assessment.....	61
5.2.1 Item Refinement Prior to Final Measurement Model.....	61
5.2.2 Indicator Reliability, Construct Reliability and Convergent Validity.....	63
5.2.3 Discriminant Validity.....	65
5.3 Structural Model Assessment.....	67
5.3.1 Structural Model Diagnostics.....	67
5.3.2 Hypothesis Testing Results.....	71
5.4 Robustness Check (Replacing Habit with Streak).....	76
5.4.1 Robustness Check: Habit VS Streak.....	77
6 Discussion.....	79
6.1 Satisfaction-Led Continuance: Why Reflective Evaluation Still Matters.....	80
6.2 Duolingo Continuance Depends On Both Learning Value And Enjoyment.....	82
6.3 Relative Challenge: Why “Easy To Repeat” Is Not Always Enough.....	84
6.4 Habit And Streak Forms Of Repeated Use.....	86
6.5 Theoretical Implications.....	88
6.6 Practical Implications.....	90
6.7 Future Research Area.....	92
7 Conclusion.....	93
8 Use of AI.....	95
References.....	96
Appendices.....	105
Appendix A - Survey Questionnaire and Measurement Items.....	105
Appendix B - Structural Model Output.....	109
Appendix C - Robustness Check Using Streak.....	111
Appendix D - Pilot Survey and Questionnaire Refinement.....	113
Appendix E - Item Refinement.....	116

List of Tables

Table 2.1 Matrix of key empirical studies and constructs relevant to this study.....	33
Table 3.1 Summary of hypotheses.....	47
Table 3.2 Definition of constructs.....	49
Table 4.1 Summary of sample characteristics.....	54
Table 4.2 Summary of measurement of constructs (selected items).....	60
Table 5.1 Descriptive statistics of respondents.....	63
Table 5.2 Items removed during measurement model refinement.....	66
Table 5.3 Measurement model results (final loadings).....	68
Table 5.4 Discriminant Validity (HTMT).....	70
Table 5.5 Structural model diagnostics.....	72
Table 5.6 Summary of hypothesis testing (before moderation, one-tailed test).....	76
Table 5.7 Summary of hypothesis results (after moderation, two-tailed test).....	76
Table 5.8 Key comparison between habit model and streak robustness model.....	80
Table A1 Summary of Questionnaires.....	110
Table A2 Summary of Questionnaires (Chinese).....	112
Table B1 Collinearity statistics (VIF).....	114
Table B2 Model Fit.....	114
Table B3 Coefficient of determination.....	115
Table B4 Effect size results.....	115
Table C1 Summary of hypothesis testing for STREAK (before moderation, one-tailed test).....	116
Table C2 Summary of hypothesis results for STREAK (after moderation, two-tailed test).....	117
Table D1 Pilot Study HTMT results.....	118
Table D2 Summary of Questionnaires for Pilot Survey.....	119
Table E1 Cross Loadings (before refinement)Appen.....	121
Table E2 - Discriminant validity HTMT results (before refinement).....	122

List of Figures

Figure 1.2 Visual summary of research background (AI-generated).....	10
Figure 1.3 Visual summary of research problem in gamified mobile language-learning continuance (AI-generated).....	14
Figure 3.2 Summary section: Literature evolution from technology acceptance to continuance (AI-generated).....	36
Figure 3.3 The model framework.....	44
Figure 5.1 Bootstrapping results. Unmoderated (one-tailed) and moderated (two-tailed).....	71
Figure 6.1 Summary of key discussion insights from the model results (AI-generated).....	79

1 Introduction

This chapter introduces the research background, problem, gaps, research questions, and objectives of the study. It positions Duolingo as a relevant case of gamified mobile language learning, where continuance intention depends on both learning value and repeated engagement. The chapter also explains why a post-adoption perspective is necessary for understanding why users continue using the app after initial adoption.

1.1 Background

Learning a language requires consistent practice enough to make meaningful progress; so we wonder how this fares in the context of a language-learning app? (Essafi et al. 2024). Language-learning applications deliver value through repeated exposure, practice, feedback, and reinforcement, rather than a single-value transaction like other digital services. The modern methods adopted by language-learning applications no longer grounded by static language rules, rather, a broader, dynamic language acquisition process driven by a multilinear factor of cognition, emotion, environment, identity, agency and multilingual communication practices (Brown and Lee, 2025). This positions learners as active agents in the learning process, while technology and AI increasingly function as facilitators of personalised learning experiences. Consequently, to make meaningful progress in language learning, one relies not only on whether users adopt the app, but also whether they continue using it over time. This makes language-learning applications a valuable context for studying continuance intention, particularly in understanding why users keep returning to an app after initial adoption.

Before the expansion of digital education, learning was largely situated around physical settings such as classrooms, textbooks, and at learning institutions own learning management systems. The expansion of internet-based and mobile technologies changed how students accessed educational materials, making resources available beyond the physical classroom and institutional libraries, allowing educational resources to be distributed more widely. Mobile learning further expanded access to information and learning materials, by making learning more

flexible, personalised, and accessible beyond formal classroom settings (Ally, 2009; Traxler, 2009). This development later reached mobile learning applications, allowing users to engage with educational content in shorter, bite-sized formats supported by fun and interactive game features (Essafi et al., 2024; Freeman et al., 2023). However, this development posed another challenge for educational platforms. Their success depends not only on whether users adopt them initially, but also on whether users remain engaged over time (Brown and Lee, 2025). In information systems research, this issue is captured through the concept of continuance intention first started by Bhattacharjee's (2001) Expectation-Confirmation Model, referring to the intention to continue using a system after actual experience, and studied by numerous studies later (Cheng and Yuen, 2018; Li and Fang, 2019; Alshurideh et al., 2020).

Among the tenets of mobile learning applications, Duolingo is characteristically one of the most visible examples of the shift toward app-based language learning. Duolingo was founded with the mission “to develop the best education in the world and make it universally available”, by making learning “fun, free, and effective” (Duolingo, 2026). This approach is reflected in what the company describes as the “**Duolingo Method**”, which combines learning science, personalisation, focused content, motivation, and delight in app-based teaching and learning (Freeman et al., 2023). The method includes five principles: learn by doing, learn in a personalised way, focus on what matters, stay motivated, and feel the delight (Freeman et al., 2023). Its development across several technological milestones, including information systems, mobile applications, gamification, and artificial intelligence, is why we choose them as our model case study for continuance intention. In particular, Duolingo provides an appropriate context for studying why users continue using a gamified mobile language-learning app after initial adoption.

Duolingo positions itself as a leading global platform among education apps in terms of market share, supported by its growing paid subscribers by over 50% every year for the past five years (Business of Apps, 2026). In the third quarter of 2024, Duolingo reported 113.1 million monthly active users and 37.2 million daily active users, representing year-on-year growth of 36% and 54% respectively (Duolingo, 2024). By the third quarter of 2025, Duolingo reported that it had surpassed 50 million daily active users (Duolingo, 2025a). Nevertheless, their user engagement

statistics remained promising, making the investigation on their users' evaluational and behavioural dynamics behind users' continued use more important. Essentially, what makes the users return to the app repeatedly, enough for learning value to meaningfully be obtained?

In 2025, Duolingo announced the launch of 148 new language courses developed through generative AI, shared content systems, and internal tooling, enabling the company to scale course development more rapidly than before (Duolingo, 2025b). Earlier, they launched another premium feature called Duolingo Max consisting of AI-powered features such as Video Call with Lily, Roleplay and Explain My Answer to support close-to-realistic conversational practices with AI-powered characters, who also provide explanations and feedback (Duolingo, 2023). These developments suggest the learning environment is increasingly assimilating AI-supported interaction, making the study of understanding user continuance more important relating to these AI features, affecting course content, language instruction, gamification, personalisation, and feedbacks. On the contrary, the personalization and accessibility offered by the AI-driven language-learning tools raises questions about how users perceive the learning and experiential value of the app, given unpopular ethical reputations such as reduced human interaction and deficient sensitivity in capturing contextual or cultural nuances (De la Vall and Araya, 2023).

Freeman et al. (2023) describes Duolingo as a combination form of both education and gamification. This approach can be understood through Expectancy-Value Theory, suggesting that learners' decision to continue using is shaped by their belief in their likelihood of successfully achieving their goal and how much they value the goal (Wigfield & Eccles, 2000). In this sense, Duolingo's gamification is mainly goal-oriented, incorporating features such as short lessons, progress indicators, XP, streaks, rewards, reminders, and feedback to condition users' learning experience (Deterding et al. 2011; Dominguez et al. 2013). These features also interestingly brought the discussion of Duolingo within the broader development of gamified mobile learning or "edutainment", where educational content is combined with engaging design to sustain motivation and repeated engagement. This is especially reflected in the "stay motivated" and "feel the delight" principles of the Duolingo Method (Freeman et al. 2023).

From a theoretical perspective, it suggests that Duolingo operates as a utilitarian-hedonic hybrid platform. It must provide ample learning value by supporting users' language-learning goals,

while also remaining enjoyable, rewarding and easy to repeat. This dual structure discussion has been mentioned within the broader continuance research, where prior studies that explored continuance generously across different contexts, showing that users' continued engagement with digital platforms is shaped by both usefulness-related and enjoyment-related factors. The most relevant context to our Duolingo case is aligned within digital education, which later narrows down to gamified learning platforms. In other adjacent digital contexts, studies have also extended the understanding of the antecedents of continuance effects through utilitarian and hedonic perceptions, through the mediators and/or moderators such as personality types, satisfaction, habit, attitude, flow condition, demographics and trust (Han and Zo, 2023; Jo, 2022; Mishra et al. 2023; Ngo et al. 2025; Tseng et al. 2024). This is particularly relevant to the Duolingo Method's multifaceted approach to app-based language teaching and learning similarly with other language-learning apps, which must be useful enough to support progress, but also engaging enough to encourage repeated practice.

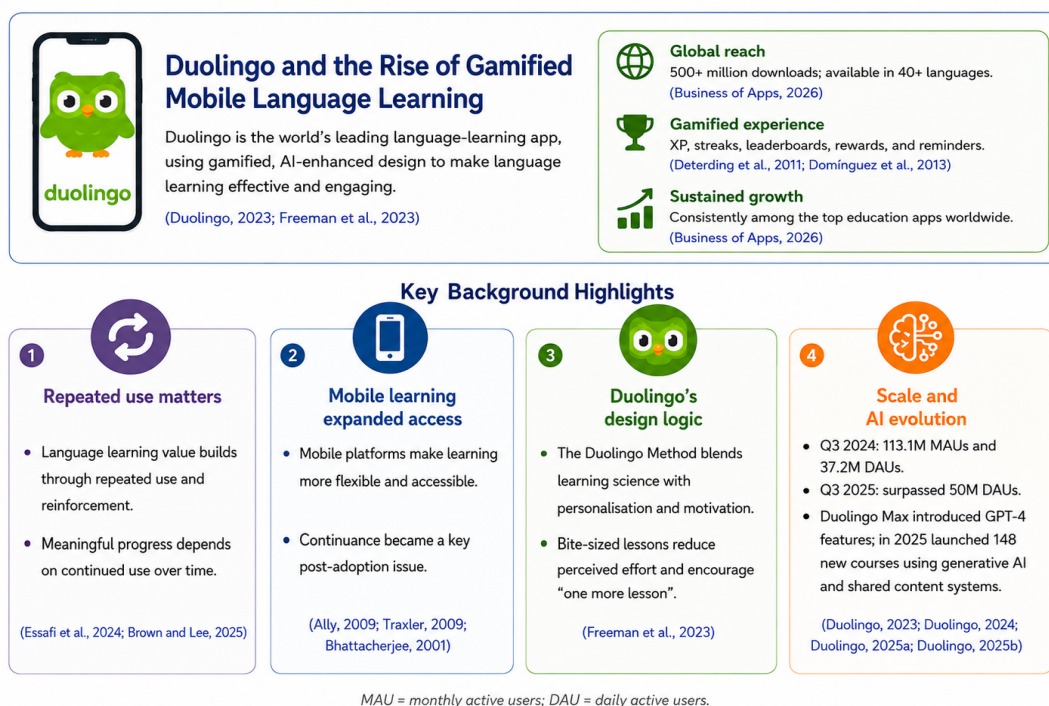


Figure 1.2 Visual summary of research background (AI-generated).

1.2 Research Problem

Digital education is essentially similar to traditional forms of education in offering learning values, where both become valuable only when learners remain meaningfully engaged over time. Especially in the context of mobile learning platforms, the rapid growth of app-based learning systems have intensified competition and fragmented users' attention, making sustaining user engagement a new, key system adoption challenge (Ally, 2009; Traxler, 2009; Bhattacharjee, 2001; Yan et al. 2021). The importance of engagement sustenance is especially emphasized in language learning, where real progress relies on repeated exposure, practice, feedback and reinforcement, rather than initial adoption. Within information systems research, researchers frequently studied post-adoption behaviour to understand the underlying drivers for continuous intention to use as they have found it is more important for long-term success, than initial technological acceptance (Bhattacharjee, 2001; Limayem et al. 2007; Mishra et al. 2023). In mobile language learning, discontinuance is therefore not merely an app-retention issue; it may also interrupt the learning process itself.

Duolingo, and other gamified language-learning mobile applications offer a meaningful context for examining this problem because they are designed to combine educational content, supported by game-like features. Despite Duolingo having a large user base worldwide and a highly popular medium for language learning, they are insufficient in explaining why individual users continue using the app. Past continuance research argued continued use is driven by multiple factors rather than a linear causal relationship. For example, the theory of Expectation-Confirmation Model (ECM) by Bhattacharjee (2001) explains continuance intention is associated with post-use perceived usefulness and satisfaction. Whereas, in consumer technology research, Venkatesh et al. (2012) extended the factors by including performance expectancy, hedonic motivation and habit. Systematic review on continuance studies also show that continuance intention may be shaped by psychological, technological, social, and behavioural antecedents, as well as by mediating and moderating mechanisms (Yan et al. 2021; Franque et al. 2021; Hsiao et al. 2016). These insights suggest that Duolingo users may continue

not only because the app is useful for learning, but also because it is enjoyable, satisfying, and capable of becoming part of their routine.

First, a key problem concerns how users evaluate gamified mobile learning applications through both learning-oriented and enjoyment-oriented perceptions. General prior mobile applications research, utilitarian and hedonic values are often treated as competing explanations. Utilitarian value is classically referring to the functional or goal-oriented benefits, while hedonic value is referring to the sense of pleasure, enjoyment and experiential gratification (Babin, Darden and Griffin, 1994; Batra and Ahtola, 1991). However in reality, these values are increasingly conjoined and interdependent in modern digital platforms. Recent studies show that both dimensions are important in post-adoption contexts. For example, Han and Zo (2023) found that both utilitarian and hedonic value influence satisfaction in mobile health app continuance, while Jo (2022) examined both value dimensions as antecedents of continuance intention toward AI personal assistants. In digital education, Tseng et al. (2024) similarly framed continuance of online English-learning resources through utilitarian and hedonic perspectives. Consequently, this blurs the clarity of either perception evaluation by Duolingo users and how that translates to their usage behaviour: whether the app helps them learn languages effectively, or whether it has a high experiential value such as enjoyable, rewarding and motivating enough to repeat. Therefore, continuance in gamified mobile language learning should not be explained through learning value or enjoyment alone, but through how these perceptions operate together.

Second, positive perceptions on the app experience do not automatically translate into continuance intention. Previous research shows that the relationship between user perceptions and continued use is explained through different post-adoption mechanisms. However, these mechanisms appear fragmented across contexts, sometimes emphasising evaluative mechanisms, sometimes behavioural mechanisms, and sometimes both. For example, the original ECM positions confirmation, perceived usefulness, and satisfaction as key constructs in post-adoption continuance (Bhattacharjee, 2001), while later extensions have incorporated perceived ease of use, attitude, trust, and enjoyment (Mishra et al. 2023). In the context of mobile social apps, Hsiao et al (2016) positioned both satisfaction, habit and customer value as parallel mediators of the effects of perceived usefulness and perceived enjoyment on continuance intention. Other

studies had also examined attachment, trust, emotional engagement, functional performance, compatibility, self-efficacy, task-technology fit, and system quality as mechanisms explaining continued use across app, AI-service, and e-learning contexts (Lin and Wang, 2011; Chang et al. 2015; Li and Fang, 2019; Ngo et al. 2025). This suggests the study of continuance is not a simple direct outcome of positive perceptions, instead these perceptions need to be translated through post-adoption mechanisms before becoming continued intention or behaviour. Therefore, the present study treats continuance intention in Duolingo as a post-adoption process through both evaluative and behavioural mechanisms (more explanation in Literature Review).

Third, a further problem concerns the tension between the learning designed for repeated engagement and the significant reduction of learning effort in Duolingo. Flow theory suggests that engagement is shaped by the relationship between challenge and skill (Csikszentmihalyi, 1990; Nakamura and Csikszentmihalyi, 2002). However, Freeman et al (2023), Duolingo uses bite-sized lessons by splitting learning materials into smaller and manageable sections so users are encouraged to complete “one more lesson” as a result of reduced perceived effort. This design may support motivation, enjoyment, and repeated engagement because learning feels easier and more achievable. However, it also raises an important learning-related question: does making learning easier and more manageable necessarily strengthen perceived learning value? Although language learning requires repetition, it also needs to be challenging enough for users to feel they are making meaningful learning progress (Engeser and Rheinberg, 2008; Løvoll and Vittersø, 2014). If lessons feel too easy, users may enjoy the app but question whether it provides enough learning value. If lessons feel too difficult, users may recognise its educational value but feel frustrated or less willing to continue. In other words, reduced effort may support repeated engagement, but it may also weaken users’ sense of meaningful learning progress, thereby affecting the evaluative and behavioural pathways leading to continuance intention.

To conclude, the study of continuance intention in gamified mobile language learning is inadequate if it relies only on initial adoption or positive perceptions of usefulness or enjoyment as direct explanations. Continuance research shows that post-adoption is influenced by multiple mechanisms including satisfaction, habit, trust, attachment, system quality and other context-specific factors. However, these mechanisms are often fragmented across different

digital contexts. In Duolingo, this problem becomes more specific because the app combines learning value, enjoyment, gamification, repetition, and reduced perceived effort. This creates a more complex continuance problem: users may return because the app is useful, enjoyable, satisfying, routine-forming, or simply easy to repeat, but these explanations may not operate in the same way for all users. Therefore, the research problem is not only why users continue using Duolingo, but how its learning, entertainment, behavioural, and difficulty-related conditions together shape continued use after initial adoption.

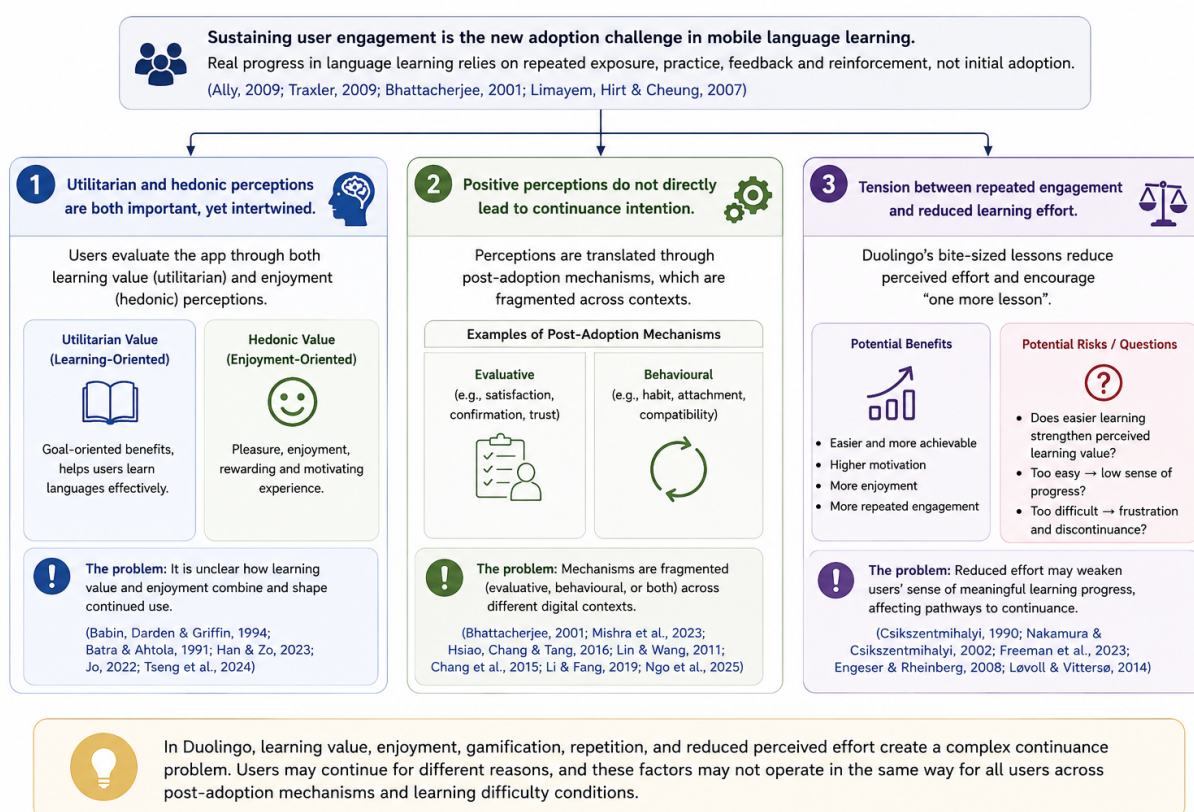


Figure 1.3 Visual summary of research problem in gamified mobile language-learning continuance (AI-generated).

1.3 Research Gap and Research Question

Initial adoption or generally positive app perceptions are not sufficient to explain continuance intention in gamified mobile language learning. If Duolingo's platform design is to make learning feels easy and enjoyable enough for users to repeat usage, the key question is whether the app experience is able to shape meaningful continuance, which is crucial in sustaining app engagement. The concept of edutainment, commonly understood as the blending of education and entertainment, is useful for understanding gamified mobile language-learning platforms because such platforms sit between learning and leisure (Anikina and Yakimenko, 2015; Okan, 2003, cited in Anikina and Yakimenko, 2015). Gamification extends this logic by using game design elements in non-game contexts to improve user experience and engagement (Deterding et al. 2011; Domínguez et al. 2013). Therefore, Duolingo must provide learning value while also keeping the experience enjoyable. Particularly, there is still a gray area in understanding which user perceptions matter, which post-adoption mechanisms translate these perceptions into continuance intention, and which learning conditions may strengthen or weaken these relationships.

Systematic literature reviews articles on continuance research show that continuance intention is shaped by a broad range of antecedents, mediators, and moderators across digital contexts. For example, Franque et al. (2021), in a meta-analysis of information systems continuance, identified affective commitment, attitude, satisfaction, hedonic value, flow, and perceived usefulness as among the strongest predictors of continuance intention. Similarly, Yan, et al. (2021) showed that online technology continuance studies commonly organise antecedents into psychological, technological, social, and behavioural factors, while also examining mediating and moderating mechanisms. These reviews suggest that continuance intention is not explained by a single pathway, but by context-specific combinations of perceptions, mechanisms, and conditions. Building from this broader evidence, this study identifies three gaps in gamified mobile language-learning continuance.

The first gap concerns the applicability of treating utilitarian and hedonic values separately in the context of gamified mobile learning, given that broader continuance reviews identify both

perceived usefulness and hedonic value as important predictors of continuance intention (Franque et al. 2021). Generally, other adjacent digital contexts such as mobile apps, AI services, and digital education have established that usefulness-related and enjoyment-related values can both influence continued use (Hsiao et al. 2016; Li and Fang, 2019; Jo, 2022; Han and Zo, 2023; Mishra et al. 2023; Tseng et al. 2024; Ngo et al. 2025). However in a gamified language-learning context, app-based teaching and learning is designed to combine both values, or as we like to call it, a “third” value comprising a hybrid of both. The consequence of this hybrid makes the analysis of the direct relationship between user perceptions and continuance intention becomes intertwined or unclear under quantitative principles, which requires further interventions to stagger the individual impact. Therefore, this study conceptualizes perceived learning value as the utilitarian perception and perceived enjoyment as the hedonic perception, while examining them together as antecedents of continuance intention in Duolingo.

The second gap which forms one of the central focuses of this study, concerns how these perception values are being translated into continuance intention through a double-mechanism process in gamified mobile learning. This is consistent with systematic literature reviews showing that continuance research has moved beyond direct antecedent–intention relationships toward models involving psychological, technological, social, behavioural, mediating, and moderating factors (Yan, et al. 2021). Many prior continuance studies have adopted satisfaction as a mediator, inspired from technology continuance theory, especially the Expectation - Confirmation Model. In ECM, satisfaction is a conscious, post-use evaluation that connects users’ system experience with their intention to continue using a system (Bhattacharjee, 2001). However, in a gamified learning environment we argue satisfaction may not be the only mechanism explaining post-adoption behaviour. Gamified app environments such as Duolingo incorporate game-like elements (e.g. streaks, XP, feedbacks, etc.) to encourage frequent and repeated use, gradually becoming habitual (Freeman et al. 2023). Habit-based continuance research argues that repeated use can become increasingly automatic and less dependent on conscious evaluation over time (Limayem et al. 2007). UTAUT2 also recognises habit as a key factor in consumer technology use, suggesting that continuance may be shaped not only by reflective evaluation but also by routinized behaviour (Venkatesh et al. 2012).

Presumably, this distinction is important because users may continue because they are satisfied with an app and, at the same time, because its use has become habitual. Satisfaction explains continuance through an evaluative route: users continue because they are pleased with the experience and judge it as worth repeating, while habit explains continuance through a behavioural route: users continue because app use has become part of their routine and requires less conscious deliberation. For example, Hsiao et al. (2016) found that both satisfaction and habit mediated the effects of perceived usefulness and perceived enjoyment on continuance intention in mobile social apps. Similarly, Wang and Lin (2021) examined mobile learning app continuance by integrating expectation-confirmation, flow-related variables, and habit, further supporting the relevance of studying continuance through both evaluative and repeated-use mechanisms in mobile learning contexts. Therefore, this study identifies satisfaction as the evaluative mediator and habit as the behavioural mediator linking perceived learning value and perceived enjoyment to continuance intention in Duolingo.

The third and final gap is the conceptualization of flow-related learning conditions specifically in gamified mobile language-learning context, as flow has been identified as an important continuance-related factor, but its role may vary depending on context and measurement approach (Franque et al. 2021; Yan, et al. 2021). In classic flow research, perceived challenge and perceived skill are often treated as separate dimensions, where researchers compare or calculate the difference between the two to examine whether balance or imbalance is associated with flow experience (Csikszentmihalyi, 1990; Nakamura and Csikszentmihalyi, 2002). Prior studies also treated challenge and skill as antecedent conditions, while flow experience itself is measured separately through dimensions such as enjoyment, focused attention, control, engagement, time distortion, and autotelic experience (Shin, 2006; Oliveira et al. 2022). However, the present study is not concerned with measuring full flow experience. Firstly, due to Duolingo's already reduced perceived learning effort, measuring challenge and skill separately, then calculating a difference score may be unnecessary and conceptually less suitable. Secondly, there is a likelihood that existing users may not interpret either of them in isolation, where users might self-report perceived challenges relative to what they feel capable of doing. Moreover, later flow research cautions that perfect balance is not always the only desirable condition, because some degree of challenge or imbalance may sustain interest and motivation (Engeser

and Rheinberg, 2008; Løvoll and Vittersø, 2014). Therefore, this study adopts relative challenge as a narrower flow-informed moderator, referring to the perceived difficulty of Duolingo's learning activities relative to users' current language ability.

Based on these gaps, the present study develops an integrated post-adoption framework for Duolingo continuance. The framework examines how perceived learning value and perceived enjoyment influence continuance intention through satisfaction and habit, while also testing whether relative challenge moderates the relationships between app perceptions and the two post-adoption mechanisms. This framework responds to the need for a more context-sensitive explanation of continuance intention in gamified mobile language learning.

The main research question is:

How do perceived learning value and perceived enjoyment influence users' continuance intention to use Duolingo through satisfaction and habit, and how does relative challenge moderate these relationships?

The study is guided by the following sub-research questions:

- **RQ1:** How do perceived learning value and perceived enjoyment influence satisfaction with Duolingo?
- **RQ2:** How do perceived learning value and perceived enjoyment influence habit in using Duolingo?
- **RQ3:** How do satisfaction and habit influence continuance intention to use Duolingo?
- **RQ4:** Do satisfaction and habit mediate the relationships between perceived learning value, perceived enjoyment, and continuance intention?
- **RQ5:** Does relative challenge moderate the relationships between perceived learning value, perceived enjoyment, satisfaction, and habit?

1.4 Research Objective

The main objective of this study is to explain continuance intention to use Duolingo from a post-adoption perspective. Based on the research gaps identified in the previous section, this study focuses on three issues. First, it examines how users' learning-oriented and enjoyment-oriented perceptions contribute to continuance intention in a gamified mobile language-learning context. Second, it investigates whether these perceptions are translated into continuance intention through two parallel post-adoption mechanisms: satisfaction and habit. Third, it examines whether these relationships vary according to users' perceived relative challenge, given that language-learning activities may feel too easy, appropriately demanding, or too difficult depending on users' current ability.

The main objectives are:

1. To examine the effect of perceived learning value and enjoyment on satisfaction and habit with Duolingo.
2. To examine whether satisfaction mediates the relationship between perceived learning value, perceived enjoyment, and continuance intention.
3. To examine whether habit mediates the relationship between perceived learning value, perceived enjoyment, and continuance intention.
4. To examine whether relative challenge moderates the relationships between perceived learning value and the two post-adoption mechanisms, namely satisfaction and habit.
5. To examine whether relative challenge moderates the relationships between perceived enjoyment and the two post-adoption mechanisms, namely satisfaction and habit.

By addressing these objectives, this study aims to provide a more integrated explanation of continuance intention in gamified mobile language learning. In particular, it seeks to explain Duolingo continuance as a process shaped by app perceptions, post-adoption mechanisms, and learning-condition differences rather than by initial adoption or positive perceptions alone.

2 Literature Review

This chapter reviews the literature on continuance intention, utilitarian–hedonic value, satisfaction, habit, gamification, and flow-related learning conditions. It explains how prior studies have approached continued use across digital education, mobile apps, AI services, and gamified learning. The chapter identifies the theoretical gaps that justify examining perceived learning value, perceived enjoyment, satisfaction, habit, and relative challenge in the Duolingo context.

2.1 The Utilitarian–Hedonic Hybrid In Gamified Learning Platforms

Earlier digital platform studies, researchers often treated utilitarian and hedonic values as binary categories. This separation reflected the dominant focus of early information systems (ISSM) and technology acceptance (TAM) research, where systems were commonly evaluated according to their functional effectiveness, usefulness, ease of use, system quality, information quality, and user satisfaction (Davis, 1989; DeLone and McLean, 1992, 2003). Utilitarian value refers to functional, instrumental, and practical benefits, while hedonic value refers to enjoyment, pleasure, and experiential gratification (Babin et al. 1994; Batra and Ahtola, 1991). It is crucial to discern the dynamics of both perspectives because while a system may help users complete a task, the elements of fun, appealing experience and interactiveness also play a significant role for continued engagement in the digital platforms (Venkatesh et al. 2012; Hsiao et al. 2016; Han and Zo, 2023).

However, recent digital platform studies suggest that modern applications have increasingly operated as utilitarian-hedonic hybrids rather than purely functional or purely enjoyable systems (Han and Zo, 2023; Jo, 2022; Ngo et al. 2025; Mpinganjira et al. 2025). Rather than asking whether utility or enjoyment matters more, the discussion is now steering towards finding ways to balance both values in different digital contexts (Babin, Darden and Griffin, 1994; Han and Zo, 2023; Tseng et al. 2024). This is especially relevant in continuance research because users’

decision to continue using a platform is often shaped by accumulated experience rather than initial adoption alone.

The hybridity of these values being studied is visible in recent app and AI-service continuance studies. For example by Han and Zo (2023), in the context of mobile health applications continued use, they conceptualized the perceptions of value as multidimensional rather than linear: both utilitarian and hedonic values suggested far stronger predictors of satisfaction and continuance than much less influenced social value. Similarly, in the study of AI personal assistants by Jo (2022), his approach in treating utilitarian and hedonic values as key antecedents of continuance intention, reflecting the importance of creating enjoyable and experiential, that functionality alone may not fully explain continued use of interactive digital services. A study on continued use of AI-powered digital voice assistants by Mpinganjira et al. (2025), had found both utilitarian and hedonic are significant predictors together, rather than purely functional. Ngo et al. (2025) demonstrate that continuance intention toward AI chatbot customer service is shaped by both functional performance and emotional engagement. Together, these studies reinforce the view that post-adoption digital behaviour is rarely driven by a single value dimension.

2.1.1 From Digital Education To Gamified Mobile Learning

The discussion on the relationship between utilitarian and hedonic values becomes more specific in digital education. Earlier educational technology designs were conventional and static, such as course management systems and e-learning platforms, primarily focusing on content deliverability, accessibility and instructional efficiency of learning (DeLone and McLean, 1992, 2003; Lin and Wang, 2011; Stone and Baker-Eveleth, 2013). As a result, early digital learning-related studies explained continuance usage mainly from the functional lens such as usefulness, ease of use, system quality and task-technology fit with academic tasks (Davis, 1989; DeLone and McLean, 2003; Goodhue and Thompson, 1995; Lin and Wang, 2011). For example, research on e-learning continuance has shown that perceived usefulness and satisfaction significantly predict continued use intention, suggesting that learners continue when a system

provides practical support for study-related activities (Lin and Wang, 2011; Stone and Baker-Eveleth, 2013).

As digital learning environments evolved, studies now shifted the focus from adoption to sustain engagement. This shift is especially important when many learning environments are voluntary and recurring, therefore understanding how to sustain learners' engagement as continued use is essential for effective language learning. In such contexts, usefulness alone may not be sufficient if the learning experience feels boring, unrewarding, or difficult to sustain over time. Shin,(2006), Chen and Tu (2021), and Tseng et al. (2024) suggest that learner engagement is shaped not only by functional learning support, but also by the quality of the learning experience. This is particularly relevant in language learning, where repeated practice is necessary but difficult to maintain if the platform does not remain engaging.

The study of online English-learning resources (OELR) by Tseng et al. (2024) provides a useful example of this shift.. Their research model explicitly examined continuance through utilitarian and hedonic perspectives, showing that language-learning technologies should not be understood only through usability or functional learning effectiveness. Instead of separating the main perspectives, they presumed hybridity, and focused more on comparing the underlying drivers of each perspective. Their key contradictory findings on the negative relationship between ease of use to continuance intention is showing that language-learning technologies should not be understood only through usability or learning effectiveness, but also through enjoyment-related experience. This means in language learning, making a system easier to use does not automatically guarantee sustained engagement if the experience does not feel meaningful, enjoyable, or motivating. This is not a definitive conclusion, but such findings are particularly relevant because stronger emphasis on functionality in language learning may deter than attract users.

Gamification in learning systems embraces the utilitarian-hedonic hybrid nature more suitably. Unlike conventional e-learning systems that mainly deliver content, gamified learning systems integrate game-like features, interactive feedback and motivational structures such as points, goals, levels, competition, progress indicators and rewards into the learning process (Deterding et al. 2011; Domínguez et al. 2013; Oliveira et al. 2022; Chen and Tu, 2021). Gamification is

commonly understood as the use of game design elements in non-game contexts to improve user experience and engagement (Deterding et al. 2011). In education, gamification has been used to increase student motivation and engagement in learning environments (Domínguez et al. 2013). This aligns in recent studies such as Oliveira et al. (2022) who explained that gamification aims to provide game-like experiences and motivations that influence user behaviour in learning environments. Similarly, Chen and Tu (2021) also described digital game-based learning as involving elements such as fun, play, goals, competition, and problem-solving, and their findings showed that digital game-based learning can make learning feel more interesting and effective. These studies suggest that gamified learning systems are not merely educational tools; they are also designed experiences that encourage learners to remain engaged.

The edutainment literature supports this hybrid interpretation. Edutainment is broadly understood as the blending of education and entertainment, where learning is made more engaging through informal, interactive, and enjoyable formats (Anikina and Yakimenko, 2015; Okan, 2003, cited in Anikina and Yakimenko, 2015). This perspective is relevant to gamified mobile language learning because such platforms sit between learning and leisure. They must provide educational value while also keeping the experience enjoyable enough to encourage repeated use. Therefore, hedonic value in gamified learning should not be treated as mere entertainment. Instead, enjoyment may help sustain repeated engagement with learning activities, especially when learning depends on consistent practice over time.

Presumably, this creates an important distinction between gamified learning and general consumer apps. In gamified learning, the expressions of utilitarian and hedonic values in gamified learning are not necessarily mutually exclusive or competing with each other, rather, they are often designed to reinforce one another (Venkatesh et al. 2012; Hsiao et al 2016; Oliveira et al. 2022). In other words, the learning app needs to effectively and adequately support users' learning outcomes such as learning progress, convenience, efficiency and educational usefulness, while also making the experience enjoyable, rewarding, and motivating enough to continue (Wang and Lin, 2021; Stone and Baker-Eveleth, 2013; Tseng et al. 2024). In this context, enjoyment is not separate from learning; it may function as part of the mechanism that sustains repeated learning behaviour.

In our study, Duolingo is a strong representation of this hybrid structure within the tenets of mobile gamified learning apps. On the utilitarian side, it offers language-learning courses that support vocabulary, grammar, reading, listening, speaking, and, in some cases, writing practice. On the hedonic side, its design integrates gamified features, such as streaks, XP, leaderboards, progress indicators, playful Duo personas, and immediate feedback. The Duolingo Method also reflects this dual logic by combining learning science with principles such as motivation and delight (Freeman et al. 2023). In addition, Duolingo's recent integration of AI-supported features and AI-generated course development suggests that the platform continues to develop both its instructional and experiential dimensions (Duolingo, 2023). Although the high usage rates (Duolingo, 2025a; 2025b) indicate users value enjoyment more than course quality, statistics alone should not be interpreted as evidence. Rather, these developments make it more important to examine how users themselves perceive Duolingo's learning value and enjoyment, and how these perceptions relate to continuance intention.

Based on this dialogue, the main gap is not simply that gamified mobile learning is under-researched. Rather, the more focused gap is that prior studies have examined utilitarian and hedonic value across mobile apps, AI services, online learning, and gamified systems, but less is known about how this dual-value logic applies specifically to gamified mobile language-learning continuance. Duolingo represents a context where learning value and enjoyment are deeply intertwined: users are both learners pursuing practical language goals and app users engaging with a designed, gamified experience.

Therefore, this study conceptualises Duolingo as a utilitarian–hedonic hybrid platform. Perceived learning value represents the utilitarian dimension, capturing whether users believe Duolingo supports their language-learning goals. Perceived enjoyment represents the hedonic dimension, capturing whether users find the app fun, pleasant, and engaging to use. This framing allows the study to examine continuance intention not as the result of usefulness or enjoyment alone, but as an outcome shaped by both learning-oriented and enjoyment-oriented perceptions.

2.2 Satisfaction And Habit As Post-Adoption Mechanisms

Building on the trajectory from utilitarian–hedonic hybridity in gamified mobile learning platforms, a further gap emerges in post-adoption research. It is no longer sufficient to examine only the direct relationship between users' perceptions and continuance intention. A more important question is how these perceptions are transformed into sustained use over time. In the context of gamified learning, this study proposes satisfaction and habit as two complementary post-adoption mechanisms.

Prior post-adoption literature had treated satisfaction as a key evaluative mechanism, although it is conceptualized in multiple ways over the years across different study contexts. In ECM, satisfaction is understood as a post-use judgement formed after users compare their actual system experience with prior expectations (Bhattacharjee, 2001). In consumer and mobile app research, satisfaction is also described as an affective or emotional response to whether the experience meets or exceeds expectations (Oliver, 1980; Giese and Cote, 2000, as cited in Ngo et al. 2025; Oliver and Desarbo, 1988, as cited in Hsiao et al. 2016). In branded app research, satisfaction may also reflect the quality of the user–app relationship, where positive experiences reinforce attachment and future use (Li and Fang, 2019). Therefore, this study defines satisfaction as users' overall post-use evaluation of whether Duolingo provides a worthwhile, positive, and rewarding learning experience.

However, another stream of post-adoption research argues satisfaction may not adequately explain continued technology use, especially when use becomes frequent, and eventually becomes habitual. While satisfaction represents an evaluative approach, habit represents a behavioural approach. Habit refers to the extent to which behaviour becomes automatic through learning and repetition, reducing the need for conscious deliberation each time the behaviour is performed (Limayem, et al. 2007). This distinction is important because users do not always return to a system after actively reassessing its value. Instead, repeated use may become part of their routine, triggered by reminders, notifications, streaks, past repetition, or environmental cues (Limayem et al. 2007; Venkatesh et al. 2012; Hsiao et al. 2016).

This is relevant in gamified mobile learning applications such as Duolingo. The use experience is often short, frequent, and repeated, supported by features such as bite-sized lessons, XP, streak days and goals. They are designed to encourage users to return and eventually, and may help app use become part of their daily routine (Oliveira et al. 2022; Chen and Tu, 2021; Freeman et al. 2023). In this sense, continuance in gamified learning may not only depend on whether users are satisfied with the experience, but also on whether using the app has become habitual.

Even for educational technology research, satisfaction and habit are increasingly discussed as mechanisms together. Sabani et al. (2026), in the context of generative AI use in education, integrates habit into the ECM-based framework and proposed habit can play multiple roles in shaping post-adoption behaviour. This suggests the contribution of habit not merely as another predictor, but also becoming increasingly important as behavioural mechanism in understanding post-adoption behaviour (Limayem et al. 2007; Venkatesh et al. 2012; Wang and Lin, 2021; Sabani et al. 2026). At the same time, this creates a tension: once use becomes habitual, continuance may depend less on fresh satisfaction-based evaluation and more on automatic repetition. This means that habitual use does not necessarily indicate that the user is satisfied, reflective, or learning effectively (Kim et al. 2005; Limayem et al. 2007; Sabani et al. 2026).

This tension does not mean that satisfaction and habit are competing mechanisms, rather they may operate in parallel. Hsiao et al. (2016), in mobile social apps context found that both satisfaction and habit significantly influenced continuance intention, and both mechanisms mediated the effects of perceived usefulness and perceived enjoyment. Their findings are important because they show that users may continue using an app both because they are satisfied with it and because its use has become automatic. Therefore, satisfaction and habit can be understood as two different but complementary routes to continuance: one evaluative and one behavioural.

Another recent technological context on AI-enabled education study by Faraon et al. (2025) further supports this interesting perspective, examining both behavioural intentions to use and actual use of ChatGPT as dependent variables. They found that performance expectancy, hedonic motivation, and habit predicted students' behavioural intention to use ChatGPT, while habit and behavioural intention predicted actual use. This is relevant for Duolingo in relation to

educational technology use, where this means that perceived learning value may motivate users to evaluate the app as useful for language learning, perceived enjoyment to offer a rewarding experience when using the app, hence this frequent practice may help transform occasional use into a learning routine. Over time, repeated practice may help transform occasional app use into a more routine learning behaviour.

The gamified learning context strengthens this argument further. A study by Chen and Tu (2021) shows that digital game-based learning features game-like elements (goals, streaks, quests, ranking) can positively improve learning motivation, boosts confidence and play performance. Conveniently, their findings about learning performance further suggests the significance of the utilitarian-hedonic hybrid nature in gamified learning contexts, as it supports the idea that game-based design is evidently making learning more engaging and effective (Deterding et al. 2011; Domínguez et al. 2013; Oliveira et al. 2022; Freeman et al. 2023). In relation to the present study, such design features may strengthen satisfaction by improving the perceived quality of the learning experience, and may strengthen habit by encouraging repeated participation through structured goals and rewards.

This section therefore positions satisfaction and habit as conceptually distinct but complementary. Satisfaction addresses the question of whether users evaluate Duolingo positively after using it. Meanwhile, habit answers a different question of whether using Duolingo will become part of the user's routine. These mechanisms may reinforce each other, but they are not the same. A user may be satisfied but not yet habitual, especially if they use the app only occasionally. Another user may be habitual but not deeply satisfied, especially if continued use is driven by streak preservation or routine rather than meaningful learning value. This distinction allows the present study to examine continuance intention more carefully.

In conclusion, the present study adopts satisfaction and habit as two post-adoption mechanisms linking perceived learning value and perceived enjoyment to continuance intention. Perceived learning value and perceived enjoyment represent what users perceive in the app experience. Satisfaction and habit explain how those perceptions may become continued use. Satisfaction captures the reflective, evaluative pathway, while habit captures the routinised, behavioural

pathway. This is especially suitable for Duolingo, where users must not only perceive learning benefits and enjoyment, but also return repeatedly for language learning to occur.

2.3 Flow Experience In Educational Psychology

In digital education, other than understanding continuance intention requires more than examining whether users perceive a platform as valuable or enjoyable. An important question in digital education: what is the learning which these perceptions most likely to become satisfaction and repeated use? This is especially relevant in language learning, as learners have different language ability, prior knowledge, objectives, personalities and tolerance (or acceptance) for difficulty. Some advanced learners need stimulation from challenging tasks, else the lesson may feel too boring for them, however this may be overwhelming for other beginner level users. Therefore, in addition to user perceptions and post-adoption mechanisms, it is important to consider how learners experience the difficulty of the learning activity itself.

Flow theory provides an important foundation for understanding this learner–task relationship. Csikszentmihalyi (1990) first founded the flow theory to study the mechanics of optimal experience, which refers to a mental state of deep involvement or immersion in the present moment. Later the definition expanded that flow occurs when individuals perceive opportunities for action that are appropriate to their capacities, supported by clear goals and feedback. (Nakamura and Csikszentmihalyi, 2002). The central conditions of flow come from perceived challenge and perceived skill, a state of mind when users are more likely to remain engaged when an activity stretches their abilities without making them feel either underused or overwhelmed. This is highly relevant to gamified learning because platforms such as Duolingo provide structured goals, progression, feedback, and repeated challenges.

However, the challenge-skill balance should not be assumed as a comprehensive concept, as simply balancing the two does not offer a complete explanation of optimized experience. Later refinement of flow model is when flow experience is achieved above the individual's both challenge and skill average levels (Massimini and Carli, 1988, cited in Nakamura and Csikszentmihalyi, 2002; Nakamura and Csikszentmihalyi, 2002). In other words, even though

both challenge and skill are well-matched, when both are low, the experience may not be motivating or meaningful. (Nakamura and Csikszentmihalyi, 2002). This is important for the present study because it suggests that Duolingo users may not be satisfied simply because lessons are easy. A productive learning experience may require challenges that are manageable but still sufficiently challenging to support perceived progress.

This critique is reinforced by later challenge-skill studies. Earlier study by Engeser and Rheinberg (2008) that challenges the classical approach of the flow theory relying on challenge-skill balance, is arguably conditional on other factors such as perceived importance of an activity and achievement motive. They also found viewing the challenge-skill balance would leave out explanations from the individual effects of challenge and skill, and it does not operate in isolation. A similar study by Fullagar et al. (2013) on student musicians over a semester, where well-balanced challenge and skill can generate flow states, as an effective intervention for the more experienced performers in reducing performance anxiety. This supports the idea that challenge-skill fit is especially relevant in learning or performance contexts where users must repeatedly face tasks that can either support optimal experience or contribute to performance anxiety. These studies suggest that challenge and skill matter, but their effects are shaped by context, motivation, and the nature of the learning task.

A more direct critique comes from Løvoll and Vittersø (2014) in their study of outdoor education students during wilderness excursions. Their findings show that challenge-skill balance is a questionable concept to produce optimal experience, and might even be associated with boredom, while interest and flow may peak when challenge optimally exceeds skill. This finding introduces the new idea of productive imbalance in the context of learning, where learners require some degree of challenge enough to stretch their current ability, in order to feel meaningful progress, mastery and growth. Therefore, this study does not assume that perfect balance is always ideal. Instead, it recognises that learners may respond differently depending on whether the activity feels under-challenging, appropriately demanding, or over-challenging relative to their current ability.

Flow theory has also been applied directly in digital learning environments. Shin (2006) conceptualized online learner's flow experience as a multidimensional construct involving

enjoyment, telepresence, focused attention, engagement, and time distortion. This is important because it shows that flow is not limited to sport, leisure, or gaming contexts, but can also explain engagement in online learning. However, Shin's study also argued that full flow experience is intervened by a broader range of factors and measures. Measuring full flow requires attention to several experiential dimensions, not simply whether learning activities feel difficult or easy.

More recent gamified learning research further supports this point. Oliveira et. al. (2022) examined flow in a gamified educational system and clearly positioned challenge-skill balance as one dimension of flow, alongside clear goals, feedback, concentration, control, and autotelic experience. Their study is useful because it connects flow directly to gamified educational environments, where learning activities are structured through goals, feedback, progression, and rewards. This aligns with prior studies that flow is a series of complex experiential conditions rather than a single measurement. This supports a more focused approach in the present study: rather than measuring full flow experience, the study examines one specific flow-informed condition that is most relevant to Duolingo's app design.

In conclusion, prior studies position challenge and skill in different ways. Even though classic flow theory measures them separately before comparing or combining them to examine flow, other studies treat them as one dimension within a wider flow experience construct. However, our central justification is in the context of gamified mobile language-learning apps, users' judgment of challenge is unlikely to be made in isolation from their perceived capability of doing. When users evaluate whether Duolingo's learning activities feel challenging, they are likely already comparing the task with what they feel capable of doing. Therefore, this study focuses on relative challenge, defined as the perceived difficulty of Duolingo's learning activities relative to users' current language ability.

Table 2.1 Matrix of key empirical studies and constructs relevant to this study

Literature	Context	Utilitarian	Hedonic	Satisfaction	Habit	Continuance	Flow Experience
Lin and Wang (2011)	E-learning continuance	✓		✓		✓	
Stone and Baker-Eveleth (2013)	Electronic textbook continuance	✓		✓		✓	
Shin (2006)	Online learning flow		✓			✓	✓
Hsiao, Chang and Tang (2016)	Mobile social apps	✓	✓	✓	✓	✓	
Wang and Lin (2021)	Mobile learning app continuance	✓	✓	✓	✓	✓	✓
Chen and Tu (2021)	Digital game-based learning	✓	✓				
Oliveira et al. (2022)	Gamified educational systems		✓			✓	✓
Han and Zo (2023)	Mobile health app continuance	✓	✓	✓		✓	
Jo (2022)	AI personal assistant continuance	✓	✓			✓	
Tseng et al. (2024)	Online English-learning resources	✓	✓			✓	
Ngo et al. (2025)	AI chatbot customer service continuance	✓	✓	✓		✓	
Mpinganjira et al. (2025)	AI-powered digital voice assistant continuance	✓	✓			✓	
Sabani et al. (2026)	Generative AI use in education	✓	✓	✓	✓	✓	
Faraon et al. (2025)	ChatGPT use in education	✓	✓		✓	✓	

3 Theoretical Framework and Hypothesis Development

This chapter develops the theoretical framework and hypotheses of the study. It explains how perceived learning value and perceived enjoyment are expected to influence continuance intention through satisfaction and habit. It also introduces relative challenge as a moderator that may shape the relationships between app perceptions and post-adoption mechanisms.

3.1 Continuance Intention in Technology Post-Adoption

This study is based on the fundamentals of technology continuance post-adoption theory. In technology research, it is argued long-term system success is more reliant on continuance after actual experience, not only initial acceptance (Bhattacharjee, 2001; Limayem, Hirt and Cheung, 2007; Mishra et al. 2023). This is because initial adoption is usually shaped by anticipatory beliefs formed before or during early use, whereas continuance intention is shaped by accumulated experience, post-use evaluation, and repeated interaction with the system (Davis, 1989; Venkatesh et al. 2003; Bhattacharjee, 2001). In the context of Duolingo, continuance intention reflects users' willingness to keep using the app after direct experience with its learning activities, feedback, gamified features, and repeated practice routines.

3.1.1 Foundations Of Consumer Behavioural Intention

Prior to technological continuance research, we argue the early roots of this field stemmed from early consumer behaviour literatures respectively. It started from the Theory of Reasoned Action (TRA) by Fishbein and Ajzen (1975) who argued behavioural intention is shaped by attitude and subjective norm. Later, the Theory of Planned Behaviour (TPB) by Ajzen (1991) extended TRA by arguing that perceived ability to perform the behaviour is also another strong determinant to behavioural intention as the former assumed behaviour is mostly under voluntary control, where people may not have full control over performing the behaviour (Ajzen, 1991). In another parallel stream around the same timeline, another key foundation is the Expectation-Confirmation Theory (ECT) by Oliver (1980) in consumer satisfaction research,

also known as expectation-disconfirmation theory, inspired literally from the consumers' experiential mechanics of achieving satisfaction, to explain why satisfaction happens post-consumption. This phase is important because it introduced two ideas that later shaped continuance research: first, intention can predict future behaviour; second, satisfaction is formed after users experience a product or service. However, these theories were developed to understand general consumer purchasing behaviour, not specifically on information systems or digital technology use.

3.1.2 Technological Acceptance And System Effectiveness

Afterwards, researchers founded more technology-specific models. The most prominent model is Technology Acceptance model (TAM) by Davis (1989) which is an influential theory in studying information systems research because it offers an explanation of why users accept or reject technology. It also became an important bridge that applied the logics from previous behavioural theories into the technological realm, proposing technological acceptance is driven mainly through perceived usefulness and perceived ease of use (Davis, 1989). For example, in the context of digital language learning, Duolingo-specific perceived usefulness could be translated to the ease of learning speaking, writing, listening and reading a new language. However, acceptance literature became too inconsistent and fragmented due to many emerging constructs to supplement the model (Venkatesh et al. 2003; Venkatesh et al. 2012). This phenomenon led to the development of the Unified Theory of Acceptance and Use of Technology (UTAUT) by Venkatesh et al. (2003), to unify those previous models, and added new concepts that capture how users form intentions toward technology. For example, performance expectancy, effort expectancy, social influence, and facilitating conditions as determinants of behavioural intention and use (Venkatesh et al. 2003). However, these models only explain the demand side or the users' perspectives, and they were also only touching a broad technology discussion, instead of different niches and variations, such as information systems (IS) which we argued is the earlier concept of digital platforms we know now (DeLone and McLean, 1992, 2003; Venkatesh et al. 2003).

Around the same timeline within the information system research stream, the Information Systems Success Model (ISSM) by DeLone and McLean (1992) contributed to the enrichment of understanding users' adoption beyond users' perceptions and beliefs from the information system perspective. ISSM introduced new concepts such as system quality, information quality, use, user satisfaction, individual impact and organisational impact, and later proceeded to add new constructs such as service quality and net benefits (DeLone and McLean, 2003). Previous research verified the advantage of ISSM in providing a more detailed explanation from the system's technical successes perspective (DeLone and McLean, 1992, 2003; Chang et al. 2015; Ngo et al. 2025), however the model does not adequately cover the psychological process through which users decide to continue using a system after actual experience.

3.1.3 Pivoting To Post-Adoption

Probably the most important shift in recorded technology research history is the Expectation-Confirmation Model (ECM) of IS Continuance by Bhattacharjee (2001) or also known as the post-acceptance model. The event which brought the shift to understanding consumers' behaviour post-adoption was in response to the limitation by previous technology acceptance theories and the problems faced by businesses on users discontinuing or disengaging with digital services (Bhattacharjee, 2001; Limayem et al. 2007; Mishra et al. 2023). Bhattacharjee argued the long-term success of IS is more closely reliant on users' loyalty in usage than initial acceptance. The elements of ECM consist of confirmation which influences perceived usefulness and satisfaction; while perceived usefulness and satisfaction influence continuance intention. This model is important because it combines the satisfaction logic from expectation-confirmation theory with the usefulness logic from technology acceptance research (Oliver, 1980; Davis, 1989; Bhattacharjee, 2001).

This is also why we chose this theory as a central theoretical foundation to understand the drivers behind users' continuance intention in using the Duolingo app. However, ECM remains relatively cognitive and utilitarian from perceived usefulness and satisfaction. This creates a gap for studying gamified mobile apps such as Duolingo, where continued use may also depend on

additional factors such as enjoyment and habit (Limayem et al. 2007; Venkatesh et al. 2012; Hsiao et al. 2016; Mishra et al. 2023).

Venkatesh later published an extended version of UTAUT2 by adding hedonic motivation, price value and habit (Venkatesh et al. 2012), as they argued consumer-centric technologies incorporate elements of fun and interactiveness, as well as incentivising habitual use through means of rewards, because perceived voluntariness is an important driver of continued usage (Limayem et al. 2007; Venkatesh et al. 2012; Hsiao et al. 2016). In other words, the research conversation moves from a mainly functional view of technology use toward a more hybrid view, where usefulness and enjoyment can coexist (Babin, Darden and Griffin, 1994; Batra and Ahtola, 1991; Han and Zo, 2023; Jo, 2022). Prior to this update, going back to information system research stream, Limayem et al. (2007) newly found that habit specifically in the form of repeated use, became increasingly important in IS continuance research. They argued users continue using a system not out of conscious deliberation, but also because it has become part of their routine. However, they also added this will create a paradox effect in limiting the predictive nature of intention as usage becomes more automatic (Limayem et al. 2007).

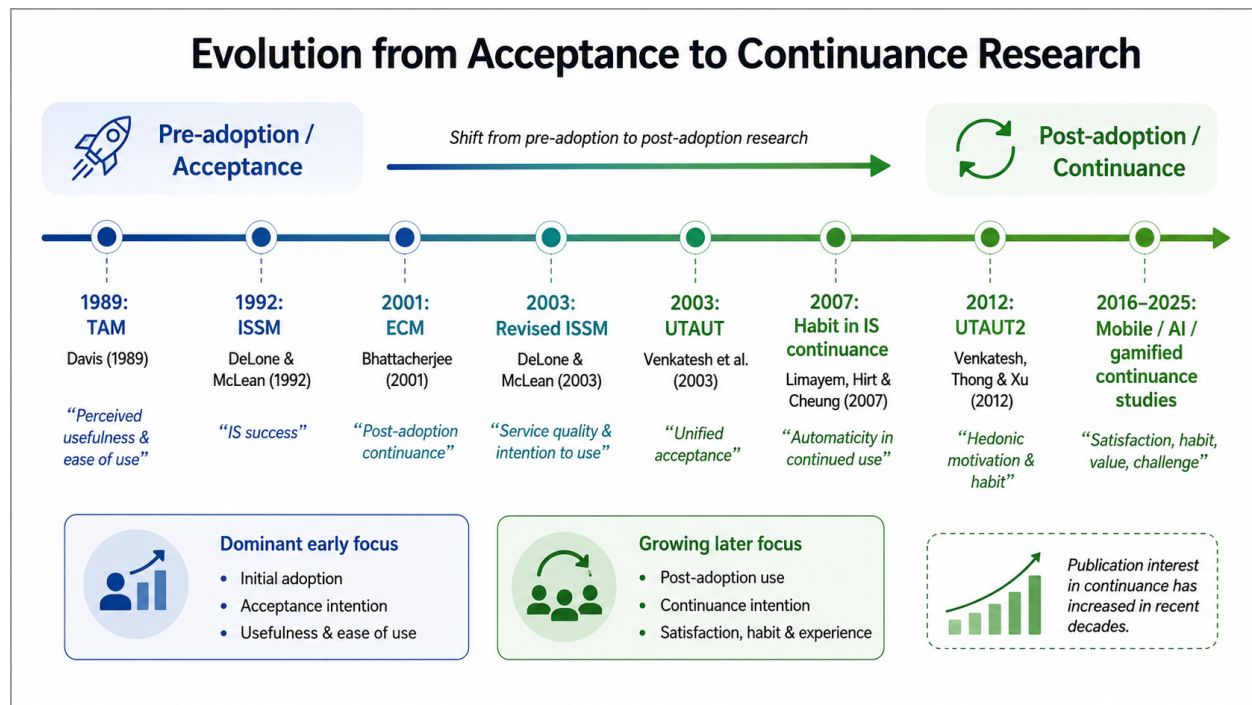


Figure 3.2 Summary section: Literature evolution from technology acceptance to continuance (AI-generated)

3.2 Hypothesis Development

3.2.2 Satisfaction As An Evaluative Post-Adoption Mechanism

The first pathway in this study is the evaluative pathway, mediated by satisfaction. Satisfaction here represents users' evaluation based on their prior experience using the system, whether or not their actual experience successfully confirms their prior expectations and created value for them. This aligns with the Expectation-Confirmation Model of IS continuance by Bhattacharjee (2001), making it a suitable model for this pathway as Duolingo continuance is not just about whether users initially adopted the app, but whether they evaluate the app positively after using it.

Past studies provide strong support for satisfaction as a predictor of continuance intention. In digital education, Stone and Baker-Eveleth (2013) found that satisfaction significantly influenced students' continuance intention to use electronic textbooks. In mobile app research, Hsiao, Chang and Tang (2016) showed that satisfaction mediated the effects of perceived usefulness and perceived enjoyment on continuance intention. In AI-enabled services, Ngo et al. (2025) also found that satisfaction strongly predicted continuance intention toward AI chatbot customer service. These studies consistently suggest that satisfaction operates as a post-use evaluative mechanism across different digital contexts.

However, prior findings also suggest that satisfaction may be shaped by different types of app perceptions. In ECM and TAM-based research, usefulness-related perceptions are important because users are more satisfied when a system helps them achieve their goals (Davis, 1989; Bhattacharjee, 2001). In UTAUT2 and consumer technology research, enjoyment is also important because voluntary consumer technologies are not used only for performance outcomes, but also for pleasurable experience (Venkatesh et al. 2012). This is particularly relevant to Duolingo because it combines learning content with gamified interaction.

Therefore, in this study, we positioned the user experience perceptions as the antecedents to positively influence satisfaction. The usefulness-related perception is represented by the *perceived learning value*, while pleasurable experience is represented by *perceived enjoyment*.

In the context of Duolingo, perceived learning value is adapted to refer to users' perception of whether or not Duolingo supports their language learning journeys, improves learning efficiency or in any way contributes to their learning progress. If users perceive their experience actually creates meaningful learning value, it will likely lead to positive experience evaluation, which increases satisfaction. At the same time, while they are learning, the app also offers enjoyable game features that rewards and motivates the users. Prior studies on mobile apps and digital services show that hedonic or enjoyment-related perceptions can positively influence satisfaction and continuance (Hsiao, Chang and Tang, 2016; Han and Zo, 2023; Jo, 2022). Therefore, perceived enjoyment should also increase satisfaction.

As a result, the perceptions mediated by satisfaction should positively influence continuance intention. In theory, the users who are more likely to perceive learning language in the app as worth continuing since they are highly satisfied, because their experience has been evaluated positively, from meeting their prior expectations. Therefore, satisfaction is expected to mediate the effects of perceived learning value and perceived enjoyment on continuance intention.

| *H1: Perceived learning value positively influences satisfaction with Duolingo.*

| *H2: Perceived enjoyment positively influences satisfaction with Duolingo.*

| *H3: Satisfaction positively influences continuance intention to use Duolingo.*

| *H4a: Satisfaction mediates the relationship between perceived learning value and continuance intention to use Duolingo.*

| *H4b: Satisfaction mediates the relationship between perceived enjoyment and continuance intention to use Duolingo.*

3.2.3 Habit As A Behavioural Post-Adoption Mechanism

The second pathway in this study is the behavioural pathway, mediated by habit. While satisfaction explains continuance through reflective evaluation, habit may also influence continuance intention to use when users repeatedly engage to learn language in the app until it becomes a part of the daily routine. In other words, habit refers to the extent to which behaviour becomes automatic through learning and repetition (Limayem et al. 2007). UTAUT2 also recognizes consumption can also be heuristic, where users may continue their usage unconsciously or automatically because use has become routine-based (Venkatesh et al. 2012).

Prior studies show that habit is relevant in post-adoption contexts, but its role is not always identical to satisfaction. Limayem, Hirt and Cheung (2007) argued that habit can limit the predictive power of intention because repeated behaviour may become automatic over time. In mobile app continuance, Hsiao, Chang and Tang (2016) found that habit significantly influenced continuance intention and mediated the effects of perceived usefulness and perceived enjoyment. In educational technology, Faraon et al. (2025) found that habit predicted behavioural intention and actual ChatGPT use among students. These studies suggest that habit can operate as a distinct behavioural mechanism in digital continuance.

Nevertheless, the distinction between habit and satisfaction is important in this study as Duolingo given the utilitarian-hedonic hybrid values. The mechanisms could provide the extension to explain users' reasons objectively for continuing to use the app as these two mechanisms may operate in parallel rather than sequentially (Limayem et al. 2007; Venkatesh, et al. 2012; Hsiao et al. 2016). The perceived learning antecedent is expected to positively influence habit as users believe in increasing the frequency in language practice to help them make meaningful progress (Brown and Lee, 2025), hence why they are more likely to keep returning to the app repeatedly which then strengthens automaticity and routine (Limayem et al. 2007; Gardner et al. 2012). Similarly, Duolingo's game features such as streaks, XP, rewards, leagues and badges are designed to further motivate users to engage with the app repeatedly. By doing so, the perceived enjoyment antecedent could positively influence habit by making language learning easier to repeat and more experiential, interesting enough that users make app use more routine-based.

In overall, this pathway is chosen as a suitable competing mechanism as it will positively influence continuance intention because unlike satisfaction, making Duolingo use more routine-based reduces the conscious, evaluative effort. Therefore, habit is expected to mediate the effects of perceived learning value and perceived enjoyment on continuance intention.

| *H5: Perceived learning value positively influences habit in using Duolingo.*

| *H6: Perceived enjoyment positively influences habit in using Duolingo.*

| *H7: Habit positively influences continuance intention to use Duolingo.*

| *H8a: Habit mediates the relationship between perceived learning value and continuance intention to use Duolingo.*

| *H8b: Habit mediates the relationship between perceived enjoyment and continuance intention to use Duolingo.*

3.2.4 Relative Challenge As Learning Condition

Finally, relative challenge is our final flow-inspired construct being included in the study as the moderator of the pathways. It refers to the information of users' perception of difficulty of Duolingo's language learning lessons relative to their current language ability. Relative challenge in this context of study is not to indicate full flow experience such as focused attention, engagement, control, time distortion and autotelic experience (Shin, 2006), but to capture a single aspect of learning condition: whether the app feels more or less challenging compared with what users believe they can currently do.

A unique distinction in this study when incorporating relative challenge is its contribution as a learning condition either higher or lower difficulty, in how it changes the users' app perceptions towards satisfaction and habit, and inherently continuance intention to use. The friction of this intervention is inspired by the opposing studies against classical flow theory (Csikszentmihalyi, 1990) where the relationship between perceived challenge and perceived skills is not straightforward in the context of gamified learning experience such, as they argued, the best experience sometime require the imbalance of the magnitude of challenge-skill antecedents, as a

sufficient degree of challenge may be necessary to sustain interest and motivation (Engeser and Rheinberg, 2008; Løvoll and Vittersø, 2014). If Duolingo feels insufficiently challenging, users may lose interest; if it feels overly difficult, users may avoid returning.

Duolingo's language lesson is designed to make learning manageable through short lessons, feedback, streaks, rewards, and progress indicators, at significantly reduced perceived learning effort (Freeman et al. 2023). For the evaluative pathway, perceived learning value and perceived enjoyment may translate into satisfaction differently depending on whether Duolingo feels sufficiently challenging or too easy. If the app feels too easy, learning value may be weakened because users may not feel that they are making meaningful progress. If the app feels more challenging, enjoyment may become more important because enjoyable design can help users experience difficulty as engaging rather than discouraging. For the behavioural pathway, relative challenge may also shape whether users return repeatedly. Habit formation may be harder when tasks feel either too easy to remain meaningful or too difficult to repeat comfortably. The result of our study might be able to provide an interesting insight on the extent of the utilitarian-hedonic hybridity being operationalized along the evaluative and behavioural mediation pathways.

Recent methodological studies further support this approach. Joessel, Pichon and Bavelier (2024) developed a video-game-based method to induce high and low flow by calibrating difficulty to participants, showing that personalised difficulty is increasingly important in flow research. Similarly, physiological studies using measures such as EEG show that researchers are still developing ways to detect flow states more objectively and in real time. These studies suggest that full flow is difficult to capture through a simple survey. Therefore, the present study does not claim to measure the entire flow experience. Instead, it examines relative challenge as a narrower and more contextually appropriate construct for a survey-based study of Duolingo continuance.

Because prior flow research presents mixed evidence on whether balance or imbalance is more beneficial, the moderation hypotheses are stated in non-directional form. This allows the study to test whether relative challenge changes the relationships between app perceptions and post-adoption mechanisms without assuming one universal direction.

Evaluative pathways:

| *H9: Relative challenge moderates the relationship between perceived learning value and satisfaction.*

| *H10: Relative challenge moderates the relationship between perceived enjoyment and satisfaction.*

| *H11a: Relative challenge moderates the indirect relationship between perceived learning value and continuance intention through satisfaction.*

| *H11b: Relative challenge moderates the indirect relationship between perceived enjoyment and continuance intention through satisfaction.*

Behavioural pathways:

| *H12: Relative challenge moderates the relationship between perceived learning value and habit.*

| *H13: Relative challenge moderates the relationship between perceived enjoyment and habit.*

| *H14a: Relative challenge moderates the indirect relationship between perceived learning value and continuance intention through habit.*

| *H14b: Relative challenge moderates the indirect relationship between perceived enjoyment and continuance intention through habit.*

Table 3.1 Summary of hypotheses

Hypothesis	Before moderation	Hypothesis	After moderation
Evaluative pathways			
H1	Perceived learning value positively influences satisfaction with Duolingo.	H9	Relative challenge moderates the relationship between perceived learning value and satisfaction.
H2	Perceived enjoyment positively influences satisfaction with Duolingo.	H10	Relative challenge moderates the relationship between perceived enjoyment and satisfaction.
H3	Satisfaction positively influences continuance intention to use Duolingo.		
H4a	Satisfaction mediates the relationship between perceived learning value and continuance intention.	H11a	Relative challenge moderates the indirect relationship between perceived learning value and continuance intention through satisfaction.
H4b	Satisfaction mediates the relationship between perceived enjoyment and continuance intention.	H11b	Relative challenge moderates the indirect relationship between perceived enjoyment and continuance intention through satisfaction.
Behavioural pathways			
H5	Perceived learning value positively influences habit in using Duolingo.	H12	Relative challenge moderates the relationship between perceived learning value and habit.
H6	Perceived enjoyment positively influences habit in using Duolingo.	H13	Relative challenge moderates the relationship between perceived enjoyment and habit.
H7	Habit positively influences continuance intention to use Duolingo.		
H8a	Habit mediates the relationship between perceived learning value and continuance intention.	H114a	Relative challenge moderates the indirect relationship between perceived learning value and continuance intention through habit.
H8b	Habit mediates the relationship between perceived enjoyment and continuance intention.	H14b	Relative challenge moderates the indirect relationship between perceived enjoyment and continuance intention through habit.

3.2.5 Proposed research model

This integrated framework allows the study to capture continuance through both evaluative and behavioural routes while accounting for the role of the learning experience itself.

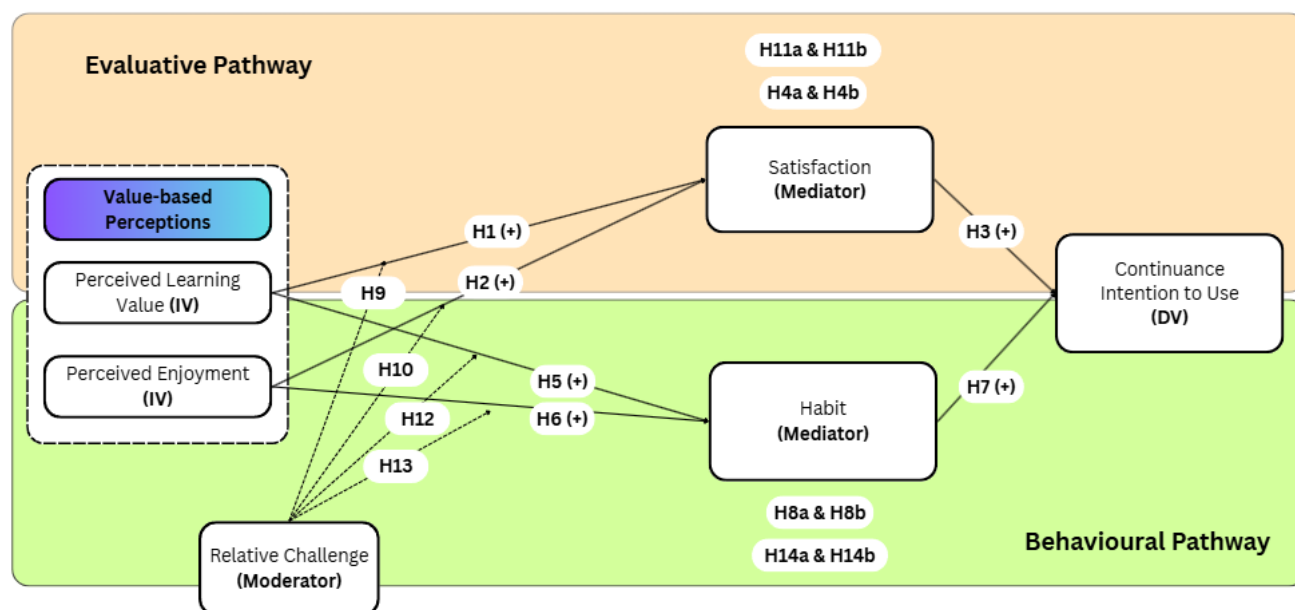


Figure 3.3 The model framework

Table 3.2 Definition of constructs

Construct	Direct quotation / source definition	Source
Perceived learning value (PLV)	Perceived learning value is adapted from usefulness/performance-related beliefs. It is the belief that technology improves users' experiences.	Adapted from (Davis, 1989; Bhattacharjee, 2001; Venkatesh et al. 2012; Faraon et al. 2025)
Perceived enjoyment (PE)	Perceived enjoyment is the extent to which using IT is enjoyable "in its own right." or "intrinsically pleasurable."	(Venkatesh and Davis, 2000, cited in Jo, 2022; Venkatesh and Davis, 2000, cited in Ngo et al. 2025)
Satisfaction (SAT)	Satisfaction is explained as users' emotional response after consumption; in ECM, it arises when outcomes "meet or exceed expectations." or a judgement that the experience is worthwhile.	(Oliver, 1980; Giese and Cote, 2000, cited in Ngo et al. 2025; Bhattacharjee, 2001)
Habit (HABIT)	Habit is defined as "the extent to which people tend to perform behaviors automatically because of learning."	(Limayem et al. 2007)
Continuance intention (CI)	Continuance intention refers to users' "ongoing intention to use a system beyond initial adoption."	(Bhattacharjee, 2001, Ngo et al. 2025)
Relative challenge (RC)	Relative challenge captures whether Duolingo activities are more or less challenging relative to current individual's ability.	(Nakamura and Csikszentmihalyi, 2002; Løvoll and Vittersø, 2014)

4 Methods

This chapter explains the methodological approach used to conduct the study. It presents the research philosophy, research design, population and sampling strategy, measurement of constructs, data collection procedure, and data analysis process. The chapter also describes the steps taken to ensure survey quality, including pilot testing, translation, screening, and measurement validation.

4.1 Research Philosophy and Approach

The purpose of the study is to statistically measure via quantitative methods the theoretical derivations of selected constructs namely perceived learning value, perceived enjoyment, satisfaction, habit, relative challenge and continuance intention, by the measure of statistical significance. **Epistemologically**, this study is driven by a positivist position through systematic approach such as observations, measurement and empirical testing (Saunders, Lewis and Thornhill, 2019; Easterby-Smith et al. 2021). Although constructs such as satisfaction and habit are not directly observable, they are measured indirectly that represent latent variables (Hair et al. 2022). **Ontologically**, this study follows a realist approach where users' perceptions and behavioural affinities exist as the indirect observable phenomena. For example, the satisfaction construct is not physically observable but they can be retrieved through self-reported answers from the survey, operationalized as indicators (Hair et al. 2022; Easterby-Smith et al. 2021).

From the synchronization of both epistemological and ontological stance, the study proceeds to make a deductive logic based on hypotheses, developed from prior empirical findings and existing theories such as ECM. These hypotheses are then tested via deductive approach to examine whether the proposed relationships are supported in a specific empirical context.

Overall by using a quantitative, deductive and positivist approach, the study is able to test the proposed moderated mediation model and examine how users' app perceptions, post-adoption mechanisms, and perceived learning difficulty relate to continuance intention.

4.2 Research Design

This study adopts a cross-sectional survey research design, by collecting data from respondents at a single point of time with the aim to investigate relationships among variables within a group, which in this case the current, active users of Duolingo (Saunders, Lewis and Thornhill, 2019; Bryman, 2016).

In this study, we are examining the users' perceptions, satisfaction, habit, relative challenge, and continuance intention based on their existing app experience and no variables are manipulated as the study is non-experimental. This approach is suitable because the focus of the study is to conduct statistical analysis of proposed relationships among the latent constructs in a real post-adoption context. Survey-based designs are commonly used in information systems and continuance research because they allow researchers to measure users' beliefs, evaluations, habits, and intentions after actual technology use (Bhattacharjee, 2001; Venkatesh et al. 2012; Hsiao, Chang and Tang, 2016).

The research also involves testing moderated effects in both direct and indirect pathways. Perceived learning value and perceived enjoyment are positioned as independent variables, satisfaction and habit are positioned as parallel mediators, continuance intention is the dependent variable, and relative challenge is included as a moderator of the relationships between the independent variables and the two mediators. Therefore, the design is appropriate following the aim of this design is to study the difference in user perceptions influence continuance intention when mediated by satisfaction and habit under different learning conditions from perceived relative challenge (Hayes, 2018).

The questionnaire was designed as a bilingual survey in English and Chinese. The English version was distributed through Google Forms, while the Chinese version was distributed through Wenjuanxing. A back-translation procedure was used to support linguistic equivalence between the two versions. This was important because the study measures latent psychological constructs which may be sensitive to wording differences across languages.

The study strictly adhered to standard ethical practice for survey-based research. All respondents provided consent of voluntary participation, were briefed about the academic purpose of the study and no personal details were collected. Responses were treated anonymously and used only for academic research purposes. Respondents could stop completing the survey at any point.

4.3 Population and Sampling

The target respondents of this study are the current and active Duolingo users who have substantial experience using the app, as the study examines continuance intention in a gamified mobile language-learning context. This was important because inactive users or discontinued users are not a reliable source of post-use evaluations of the app, as their perspectives will not be compatible with the operationalizations of the constructs. For example, active users are more suitable for examining habit because habit represents repeated usage with the app. Therefore, respondents were screened to ensure that they had Duolingo usage experience via verification check, and an active streak at the time of response.

To meet this requirement, our sampling strategy consists of non-probability, convenience and purposive sampling. Convenience sampling because the respondents were recruited through nearest accessible channels both online and offline networks, including social media, messaging groups, university-related channels and field recruitment mainly around the compounds of Lund University. Purposive sampling strategy was to ensure the respondents meet the survey criteria who possess relevant Duolingo usage experience and are still actively using. This approach is suitable for theory-testing research where the aim is to examine relationships among constructs within a relevant user group rather than to estimate population-level proportions.

Through convenience sampling, the respondents were of international respondents, conveniently recruited through Sweden- and China-based channels, but the study does not require country-level or ethnicity-based comparisons, and they are not treated as representatives of the main ethnicities of these two environments. This is reflected in the final sample in Table 4.1, which is regionally, composed of respondents from Europe and Asia, followed by North America, Latin America and the Caribbean, Africa, Oceania, and other or unspecified regions.

A total of 255 responses, after data screening, 186 valid responses were retained for analysis. The final sample consisted of 127 female respondents, 55 male respondents, 2 non-binary respondents, and 2 respondents who preferred not to say. Most respondents were aged 18–34, with 70 respondents aged 18–24 and 77 respondents aged 25–34. In terms of education, 109 respondents held a bachelor's degree, 35 held a master's degree or above, and 42 had college-level education or below. The final sample also showed active Duolingo engagement. All retained respondents reported an active streak, with 34 respondents reporting a streak of 1–7 days, 22 reporting 8–30 days, 50 reporting 31–100 days, and 80 reporting more than 100 days.

Table 4.1 Summary of sample characteristics

Variable	Category	Grand Total, N (%)
Sample	Total	186 (100%)
Gender	Female	127 (68%)
	Male	55 (30%)
	Non-binary	2 (1%)
	Prefer not to say	2 (1%)
Age	18–24	70 (38%)
	25–34	77 (41%)
	35–44	25 (13%)
	45+	14 (8%)
Education	Master’s or above	35 (19%)
	Bachelor’s	109 (59%)
	College or below	42 (23%)
Streak	1 - 7 days	34 (18%)
	8 - 30 days	22 (12%)
	31 - 100 days	50 (27%)
	More than 100 days	80 (43%)
Region	Europe	75 (40%)
	Asia	46 (25%)
	North America	9 (5%)
	Latin America & Caribbean	8 (4%)
	Africa	7 (4%)
	Oceania	1 (4%)
	Other/Unspecified	40 (22%)

Note: Percentages are based on respective column total.

4.4 Measurement of Constructs

Based on the proposed research model, the selected questionnaires (refer Appendix A) were designed to measure the latent construct effectively. The main model includes two independent variables, namely perceived learning value and perceived enjoyment; two mediating variables, namely satisfaction and habit; one moderating variable, relative challenge; and one dependent variable, continuance intention. In addition, the current Duolingo streak category was collected for the robustness check, where streak was used as an alternative behavioural indicator to compare against the psychological habit construct.

All constructs except Streak, were measured using a seven-point Likert scale ranging from 1 = strongly disagree to 7 = strongly agree, fully adapted from prior studies across relevant contexts such technology acceptance, continuance, mobile learning, gamified learning, habit and flow. Nevertheless, the items were adapted to suit the Duolingo case inspired by the prior empirical studies who adopted a similar approach, in an attempt to preserve the conceptual meaning of the original constructs. This approach was used to support content validity and to ensure that the study builds on established measurement traditions.

Prior to finalizing the measurements, we conducted a three-day pilot survey (refer Appendix D) with a small sample to assess the feasibility of items in terms of clarity and discriminant validity. This step was a cautionary step due to high chance of closely conceptually related constructs within the gamified learning context, especially by satisfaction to the perception constructs and continuance intention. The pilot results indicated discriminant validity concerns, especially around the satisfaction construct. Therefore, several items were dropped, replaced, or added to improve construct clarity. For example, the broad satisfaction items originally from the ECM original paper by Bhattacharjee (2001) were revised into more focused items that captured users' post-use evaluation of Duolingo and the Duolingo company (Singh and Singh, 2024; as cited in Ngo et al. 2025). This refinement helped distinguish satisfaction from perceived learning value and perceived enjoyment.

4.4.1 Independent Variables: Perceived Learning Value and Perceived Enjoyment

Perceived learning value refers to users' perception that Duolingo supports their language-learning goals and contributes to their learning progress. It was measured using items adapted from studies on learning value, mobile learning, and perceived value in digital services, including Tseng and Teng (2014, as cited in Tseng et al. 2024), Ooi and Lee (2018), and Han and Zo (2023). The final items focused on whether Duolingo helps users learn faster, improves learning efficiency, and is useful for language learning.

Perceived enjoyment refers to the extent to which users find Duolingo enjoyable, fun, entertaining, and engaging to use. It was measured using items adapted from Venkatesh et al. (2012) and Shin (2006). The final items captured users' enjoyment of Duolingo as a gamified and interactive language-learning application.

4.4.2 Mediating Variables: Satisfaction and Habit

Satisfaction refers to users' overall post-use evaluation of Duolingo. It captures whether users feel positively about their experience with the app and whether the experience is perceived as worthwhile. Satisfaction was measured using items adapted from Singh and Singh (2024), as cited in Ngo et al. (2025). The final satisfaction items were made more specific to users' evaluation of Duolingo and the Duolingo company, including positive impression, feeling valued as a consumer, and overall satisfaction with the app experience.

Habit refers to the extent to which using Duolingo has become automatic and routine-based. In this study, habit is treated as a psychological automaticity construct rather than a direct measure of repeated behaviour. The construct was measured using items adapted from the Self-Report Behavioural Automaticity Index developed by Gardner et al. (2012). These items focus on whether using Duolingo occurs automatically

4.4.3 Moderating Variable: Relative Challenge

Relative challenge is the moderating variable in this study. It refers to users' perception of the difficulty of Duolingo's learning activities relative to their current language ability (Shin, 2006). The construct is grounded in flow-related challenge and skill literature, but it is not intended to measure full flow experience. Instead, it captures a narrower learning condition that may shape whether perceived learning value and perceived enjoyment translate into satisfaction and habit.

During the pilot stage, perceived challenge and perceived skill were initially considered as separate measures. However, the pilot results indicated measurement concerns, including conceptual overlap and discriminant validity issues. As a result, perceived skill items were removed, and the final questionnaire measured relative challenge directly. This direct operationalisation was considered more suitable because users are likely to judge the difficulty of Duolingo's learning activities in relation to what they feel capable of doing. Therefore, rather than calculating relative challenge as the difference between separate challenge and skill scores, the final study measured perceived difficulty relative to current language ability directly.

4.4.4 Dependent Variable: Continuance Intention

Continuance intention is the dependent variable in this study. It refers to users' intention to continue using Duolingo in the future. Since the study adopts a post-adoption perspective, continuance intention is appropriate because it captures whether users intend to keep using the app after actual experience. Continuance intention was measured using items adapted from Bhattacharjee (2001), Tseng et al. (2024), Lin and Wang (2011), and related continuance intention research. The items captured users' willingness to use Duolingo again, intention to continue rather than discontinue use, preference for Duolingo in future language learning, and overall intention to continue using the app for language learning.

4.4.5 Streak as a Robustness-Check Variable

The current Duolingo streak category was collected as a behavioural indicator of repeated app engagement, as a robustness check to compare the results of the main habit model with an alternative model using streak. Habit and streak are related but conceptually different. Habit refers to psychological automaticity, meaning that users feel they use Duolingo automatically or routinely. Streak, on the other hand, reflects an invisible behavioural record of repeated use within the app. By comparing habit and streak, the study examines whether the behavioural pathway remains consistent when repeated use is represented by a self-reported psychological construct versus an app-based behavioural indicator.

Table 4.2 Summary of measurement of constructs (selected items)

Construct	Role	Operational Definition	N	Source
Perceived Learning Value (PLV)	Independent variable	Users' perception that Duolingo supports their language-learning goals, improves learning efficiency, and contributes to learning progress.	3	Tseng and Teng (2014), as cited in Tseng et al. (2024); Ooi and Tan (2016) as cited in Ooi and Lee (2018); Han and Zo (2023)
Perceived Enjoyment (PE)	Independent variable	Users' perception that using Duolingo is enjoyable, fun, entertaining, and engaging.	4	Venkatesh et al. (2012); Shin (2006)
Satisfaction (SAT)	Mediating variable	Users' overall post-use evaluation of whether Duolingo provides a positive and worthwhile app experience.	3	Singh and Singh (2024), as cited in Ngo et al. (2025)
Habit (HABIT)	Mediating variable	The extent to which using Duolingo has become automatic and routine-based.	4	Gardner et al. (2012)
Relative Challenge (RC)	Moderating variable	Users' perception of the difficulty of Duolingo's learning activities relative to their current language ability.	4	Shin (2006)
Continuance Intention (CI)	Dependent variable	Users' intention to continue using Duolingo for future language learning.	4	Bhattacharjee (2001); Chang (2010) and Vatanasombut et al. (2008) as cited in Liu and Wang (2011); Tseng et al. (2015), as cited in Tseng et al. (2024)
Current Duolingo Streak Category	Control variable	Users' current behavioural engagement with Duolingo, represented by their active streak category.	1	Self-reported behavioural data

4.5 Data Collection

The data collection period was over the course of two weeks via online survey platforms: Google Forms was for the Sweden-based channels, while Wenjuanxing was for China-based channels. The reason for the two different platforms due to the accessibility requirements in both countries.

The survey began with a brief introduction such as the purpose of the study and voluntary participation. Next the respondents will be screened by double-verification questions: regarding their streak history and a simple knowledge test, to filter respondents who have substantial use experience in Duolingo to support the validity of the study. This was followed by the main construct items measuring perceived learning value, perceived enjoyment, satisfaction, habit, relative challenge, and continuance intention. The final section collected demographic information, including age, gender, education level, and country of recruitment.

Next, from the collected data, responses were excluded if they were incomplete, failed the screening criteria, did not report an active Duolingo streak, answered verification questions incorrectly, or showed invalid response patterns such as straight-lining. After applying these screening procedures, out of 255 responses, 186 valid responses remained. The use of data screening improves the credibility of the analysis by ensuring that respondents were relevant to the research context and that responses were not careless or invalid.

The final dataset was then prepared for analysis. Scale items were coded so that higher values represented higher levels of the construct being measured. Reverse coding was applied where necessary. Composite scores and latent variable indicators were prepared for analysis in SmartPLS 4.

4.6 Data Analysis Process

This study analysed the collected data through Partial Least Squares Structural Equation Modelling (PLS-SEM) in SmartPLS 4. PLS-SEM is suitable for prediction-oriented research and for analysing complex models involving direct and indirect relationships among latent variables (Hair et al. 2022; Sarstedt, Ringle and Hair, 2022). Prior empirical studies using PLS-SEM similarly justify the method when the model includes indirect effects, moderation, and bootstrapping-based significance testing (Cenamor et al. 2019).

The analysis followed the standard two-stage approach in PLS-SEM. First, the measurement model was assessed to evaluate whether the survey items measured their intended constructs reliably and validly. This included examining outer loadings, Cronbach's alpha, rho_A, composite reliability, average variance extracted, and HTMT. Second, the structural model was assessed to test the hypothesised relationships among the constructs. This included collinearity assessment, path coefficients, R^2 , f^2 , and hypothesis testing: mediation effects, moderation effects, and moderated mediation effects.

Bootstrapping was used to test the statistical significance of the path coefficients and indirect effects. Bootstrapping is appropriate in PLS-SEM because it is a non-parametric resampling procedure that does not require normally distributed data and is commonly used to assess the significance of direct, indirect, and interaction effects (Preacher and Hayes, 2008; Hair et al. 2022). In this study, one-tailed testing was applied to the main direct and mediation hypotheses because these hypotheses predicted positive relationships. Two-tailed testing was applied to moderation and moderated mediation hypotheses because the direction of the relative challenge effect was not assumed. Lastly, the study adopted a significance threshold of $p < 0.05$.

4.7 Limitations and Mitigations

In this study, several limitations need to be clarified throughout the operationalization of the study, and the steps that were taken to manage them.

The cross-sectional survey design only analyzes all variables at one point of time. The findings and analyses are best interpreted as correlations of one another, and do not sufficiently explain causal relationships. Nevertheless, the theoretical framework driving the research model was developed from established theories and prior empirical findings, aligning with our research positivist and realist approach.

The self-reporting nature of data does not leave out potential bias from the respondents in recall bias and mood at the point of answering the survey. We manage these concerns by ensuring the questionnaires used are well-established measurement items from prior studies. Nevertheless, even before the formal blasting, we ran a pilot test to test those questions which resulted in a second-round validation on the operationalization of the questions based on the pilot test results. To improve reliability and validity of data, we screened the respondents by adopting eligibility questions to only include current and active users of Duolingo, not those who already discontinued the app. Finally, although the non-probability sampling from Sweden and China platform may potentially reduce generalizability, our respondents consist of multiple nationalities and demographics. The final sample was suitable because the priority was to find actual active Duolingo users who are capable of evaluating the app.

Finally, some constructs are conceptually close around satisfaction such as by perceived learning value, perceived enjoyment and continuance intention. To address this, we have adopted items refinement process during the analysis and passed the measurement model assessment. Although this closeness should still be considered when interpreting the results, it is a more genuine reflections of the hybrid nature of Duolingo as both a learning and entertainment platform.

5 Findings and Results

This chapter presents the empirical findings of the study. It reports the respondent profile, descriptive statistics, measurement model assessment, structural model diagnostics, hypothesis testing, and robustness check. The chapter focuses on whether the proposed relationships are statistically supported before the findings are discussed in the next chapter.

5.1 Descriptive Statistics

Table 5.1 Descriptive statistics of respondents

	Mean	Median	SD	Skewness
PLV	4.92	5.013	1.318	-0.573
PE	5.122	5.483	1.319	-1.014
HABIT	4.412	4.890	1.615	-0.401
SAT	5.053	5.307	1.281	-1.037
CI	5.312	5.716	1.389	-0.932
RC	2.361	2.012	1.211	1.428

Table 5.1 presents the descriptive statistics for the main constructs in this study. Overall, the results show that respondents evaluated Duolingo positively, particularly in terms of continuance intention, perceived enjoyment, satisfaction, and perceived learning value.

The first highlight is **continuance intention** with the highest mean and median score ($M = 5.312$, $Mdn = 5.716$, $SD = 1.389$), indicating a strong willingness in continuance intention to use Duolingo, which supported the ECM theory, reflecting users' willingness to continue using outcome. **Perceived enjoyment** results support the relevance of hedonic motivation in consumer technology use (Venkatesh et al. 2012) by obtaining a high mean score ($M = 5.122$, $SD = 1.319$), showing most respondents found Duolingo entertaining. This aligns with the argument of

game-like features in learning environments can help boost engagement and motivation (Deterding et al. 2011; Domínguez et al. 2013; Oliveira et al. 2022). **Satisfaction**, a central construct in the Expectation-Confirmation Model (Bhattacharjee, 2001), showed a similarly positive mean score ($M = 5.053$, $SD = 1.281$), suggesting that respondents have a positive experience in Duolingo. **Perceived learning value** also recorded a positive mean score ($M = 4.920$, $SD = 1.318$). This implies respondents generally perceived Duolingo as useful for language-learning. This result aligns with utilitarian value theory, where users value a system when it meets the functional expectations of a system (Babin et al. 1994; Batra and Ahtola, 1991). **Habit** recorded a moderate mean score ($M = 4.412$, $SD = 1.615$) which indicates a fragmented behaviour based on the high standard deviations. Potentially some users developed automatic use patterns, while others may still use Duolingo more persistently. This is consistent with habit-continuance research, where repeated use can become automatic over time, but that habit strength may vary depending on users' prior repetition and experience (Limayem et al. 2007; Gardner et al. 2012). Finally, the **relative challenge** recorded the expected lowest mean score ($M = 2.361$, $SD = 1.211$) as well high positive skewness, suggesting that respondents generally did not perceive Duolingo's learning activities as highly difficult relative to their current language ability. This is reflective of Duolingo Method learning design in making learning easy (Freeman et al. 2023).

The overall descriptive statistics provide reasonable reflection of the study context: Duolingo is generally perceived as enjoyable, satisfying, useful for learning, and worth continuing, while its learning activities are significantly easier and manageable, rather than highly challenging.

5.2 Measurement Model Assessment

5.2.1 Item Refinement Prior to Final Measurement Model

Prior to interpreting the final measurement model, item refinement was conducted to improve discriminant validity among closely related constructs. The initial measurement model showed high HTMT values between perceived learning value and satisfaction, suggesting that some items may have overlapped conceptually and empirically. To identify the source of overlap, cross-loadings (refer Appendix E) were examined. The priority of removal falls on items with closest loadings, therefore PLV1 and SAT1 were first removed as perceived learning value and satisfaction were relatively close, indicating a failure in conceptual distinction between the items. Next, PE2 was removed for the similar reason being too associated with perceived learning value to ensure both constructs are sufficiently distinguished under the threshold of 0.9.

PLV1 and SAT1 were removed because their loadings were relatively close across perceived learning value and satisfaction, indicating insufficient empirical distinction between the items. PE2 was also removed because its cross-loading pattern showed close association with perceived learning value, suggesting that the item did not clearly distinguish perceived enjoyment from learning value.

This refinement is theoretically consistent with the utilitarian–hedonic hybridity discussed in this study. In gamified mobile language learning, learning value and enjoyment are expected to be closely related because the app is designed to make learning useful and enjoyable at the same time. However, for measurement purposes, the constructs must remain empirically distinguishable. Therefore, the removal of PLV1, SAT1 and PE2 was not intended to deny the conceptual hybridity of Duolingo, but to ensure that the final measurement model captured perceived learning value, perceived enjoyment and satisfaction as related but distinct constructs.

After removing these items, the final measurement model achieved acceptable reliability, convergent validity and discriminant validity. The retained items are presented in Table 5.2, while the item refinement process and cross-loading results are provided in Appendix E.

Table 5.2 Items removed during measurement model refinement

Removed item	Construct	Reason for removal	Decision
PLV1	Perceived learning value	Cross-loading pattern showed overlap with satisfaction, contributing to high HTMT between PLV and SAT.	Removed
SAT1	Satisfaction	Cross-loading pattern showed overlap with perceived learning value, contributing to high HTMT between PLV and SAT.	Removed
PE2	Perceived enjoyment	Cross-loading pattern showed close association with perceived learning value, reducing distinction between enjoyment and learning value.	Removed

5.2.2 Indicator Reliability, Construct Reliability and Convergent Validity

In PLS-SEM, outer loadings above 0.70 are acceptable to indicate proportionally sufficient variance within the latent construct (Hair et al. 2022). As shown in Table 5.3, all retained items exceeded the recommended threshold, with loadings ranging from 0.787 to 0.907. The lowest loading was observed for CHA1 (0.787), while the highest loading was observed for CI4 (0.907). Therefore, all indicators positioned beyond the allowed threshold indicate adequate representation of their respective constructs.

The items within each construct demonstrated acceptable to strong internal consistency reliability. The constructs exceeded the recommended threshold of 0.70. Cronbach's alpha values ranged from 0.822 to 0.891, rho_A values ranged from 0.827 to 0.893, and composite reliability values ranged from 0.893 to 0.925.

All constructs also achieved acceptable convergent validity., AVE values above 0.50 indicate that a construct explains more than half of the variance in its indicators (Fornell and Larcker, 1981; Hair et al. 2022). As shown in Table 5.3, all AVE values exceeded this threshold, ranging from 0.688 for relative challenge to 0.756 for perceived learning value.

The overall measurement model results suggest that the retained items reliably and validly measured their intended constructs. Perceived learning value, perceived enjoyment, satisfaction, habit, continuance intention, and relative challenge all met the recommended criteria for indicator reliability, internal consistency reliability, and convergent validity.

Table 5.3 Measurement model results (final loadings)

Item	Item Wording	Load.	α	pA	C.R.	AVE
Perceived Learning Value			0.838	0.838	0.903	0.756
PLV2	Duolingo enables me to learn new languages faster.	0.839				
PLV3	Using Duolingo increases my learning efficiency	0.900				
PLV4	Using Duolingo is very useful for my language learning	0.868				
Perceived Enjoyment			0.891	0.893	0.925	0.754
PE1	Using Duolingo is truly fun.	0.868				
PE3	Using Duolingo is very entertaining.	0.897				
PE4	I find Duolingo's interactive features enjoyable to use.	0.855				
PE5	Duolingo's learning activities are not boring to me.	0.853				
Satisfaction			0.822	0.827	0.893	0.737
SAT2	My experience leaves me with a positive impression of the Duolingo company.	0.871				
SAT3	I feel valued as a consumer when using Duolingo to learn a language.	0.840				
SAT4	Overall, I am satisfied with my experience with Duolingo.	0.864				
Habit			0.88	0.886	0.917	0.734
HABIT1	Using Duolingo is something I do automatically.	0.831				
HABIT2	Using Duolingo is something I do without having to consciously remember.	0.887				
HABIT3	I often start using Duolingo without consciously deciding to do so.	0.818				
HABIT4	Using Duolingo is something I do without thinking.	0.890				
Continuance Intention			0.882	0.883	0.919	0.739
CI1	Next time I am willing to use Duolingo again to learn a new language.	0.824				
CI2	I intend to continue using Duolingo rather than discontinue its use.	0.855				
CI3	For future language learning, I would choose Duolingo.	0.851				
CI4	Overall, I intend to continue using the Duolingo app for language learning.	0.907				
Relative Challenge			0.849	0.862	0.898	0.688
RC1	Duolingo's learning activities are too demanding for me.	0.787				
RC2	It is difficult for me to understand Duolingo's lesson content.	0.834				
RC3	It is difficult for me to keep up with Duolingo's learning activities.	0.833				
RC4	It is hard for me to complete the tasks required in Duolingo.	0.861				

5.2.3 Discriminant Validity

In this section, is the measurement of discriminant validity based on the heterotrait-monotrait ratio (HTMT) over the threshold of 0.9, meaning that HTMT values below 0.90 indicate acceptable discriminant validity.. HTMT is recommended in PLS-SEM because it provides a sensitive assessment of whether constructs are empirically distinct from one another (Henseler et al. 2015). After the refinement of items, all HTMT values were below the 0.90 threshold, as shown in Table 5.4. The constructs in the measurement model all operated within acceptable discriminant validity and were adequately distinct from each other.

However, several values were close to the threshold and should be interpreted with caution. The HTMT value between satisfaction and continuance intention was 0.882, while the values between satisfaction and perceived learning value, and satisfaction and perceived enjoyment were 0.881 and 0.880 respectively. These values remain below the 0.90 threshold, but they indicate conceptual closeness among the constructs. This is theoretically reasonable because users who perceive Duolingo as enjoyable are also likely to report higher satisfaction and stronger intention to continue using the app.

Finally, referring to the interaction terms, the HTMT value between $RC \times PE$ and $RC \times PLV$ was 0.877, which is close to the threshold but still acceptable discriminant validity. This outcome is expected as both terms were driven by the same moderator RC. Despite this, the value remains below 0.90, implying that the interaction constructs were still empirically distinguishable.

Table 5.4 Discriminant Validity (HTMT)

	CI	HABIT	PE	PLV	RC	SAT	RC x PE
CI							
HABIT	0.568						
PE	0.732	0.706					
PLV	0.810	0.672	0.725				
RC	0.463	0.285	0.519	0.376			
SAT	0.882	0.664	0.880	0.881	0.503		
RC x PE	0.473	0.314	0.553	0.468	0.641	0.587	
RC x PLV	0.337	0.269	0.507	0.362	0.625	0.476	0.877

Overall, the HTMT results support the discriminant validity of the measurement model. Although some construct pairs were closely related, all values remained below the accepted 0.90 threshold. Therefore, the constructs were considered sufficiently distinct for the structural model assessment.

5.3 Structural Model Assessment

5.3.1 Structural Model Diagnostics

Prior to interpreting the hypothesis testing results, the structural model was evaluated through several diagnostic checks consisting of collinearity, model fit, explanatory power and effect size (Hair et al. 2022; Sarstedt, Ringle and Hair, 2022). This approach is relevant based on prior PLS-SEM studies that report diagnostic indicators before interpreting hypothesis results, particularly when the model includes mediation and moderation effects (Henseler et al. 2009; Hair et al. 2019; Cenamor et al. 2019). For example in Cenamor et al. (2019), PLS-SEM is suitable for evaluation of models involving indirect and moderation effects. Full outputs are provided in Appendix B.

Firstly, the model is evaluated in terms of collinearity test through variance inflation factor (VIF) especially given the existence of moderation effects, as high overlap among predictor constructs can inflate standard errors and reduce the stability of path coefficient estimates (Hair et al. 2022). As shown in Table 5.5, the VIF values for the main-effect predictors ranged from 1.488 to 1.981, which is well below the recommended threshold of 5.0. The interaction terms recorded higher VIF values, with $RC \times PE$ reaching 4.947 and $RC \times PLV$ reaching 4.545. Although these values are close to the threshold, they remain below 5.0 and are therefore considered acceptable (Hair et al. 2022). Therefore, multicollinearity was not considered severe enough to prevent interpretation of the structural paths, although the moderation results were interpreted with caution.

Table 5.5 Structural model diagnostics

Assessment	Main Highlights	Result	Threshold	Interpretation
Collinearity	Main-effect VIF range	1.488–1.981	VIF < 5.0	No collinearity concern among main predictors
	Interaction-term VIF range	4.545–4.947	VIF < 5.0	Acceptable, but close to threshold
Model fit	SRMR, saturated model	0.065	SRMR < 0.08	Acceptable approximate fit
	SRMR, estimated model	0.071	SRMR < 0.08	Acceptable approximate fit
Explanatory power	R ² for satisfaction	0.702	Higher values indicate stronger explanation	Substantial explanatory power
	R ² for continuance intention	0.580		Moderate-to-substantial explanatory power
	R ² for habit	0.460		Moderate explanatory power
Effect size	Largest f ² effect	SAT → CI = 0.757	0.02 small, 0.15 medium, 0.35 large	Satisfaction had a large effect on continuance intention
	Key f ² effects on satisfaction	PE → SAT = 0.313; PLV → SAT = 0.300	0.15 medium; 0.35 large	PE and PLV had medium-to-large effects on satisfaction
	Habit effect on continuance intention	HABIT → CI = 0.022	0.02 small	Habit had a very small effect on continuance intention
	Interaction effect on satisfaction	RC × PE → SAT = 0.036	0.02 small	Moderation effect was small

In terms of model fit, the test evaluates this using the standardized root mean square residual (SRMR). SRMR assesses the difference between the observed correlations and the model-implied correlations, with values below 0.08 generally indicating acceptable approximate fit (Henseler et al. 2014; Hair et al. 2022). In this study, the saturated model produced an SRMR value of 0.065, while the estimated model produced an SRMR value of 0.071. Both values are below the recommended threshold, suggesting acceptable approximate model fit. However, because PLS-SEM is primarily variance-based and prediction-oriented, model fit indices should be treated as supplementary rather than as the main criterion for evaluating model adequacy (Hair et al. 2019; Sarstedt, Ringle and Hair, 2022). Therefore, the SRMR result was interpreted alongside measurement quality, explanatory power, effect sizes, and path significance.

Next, the model is assessed in terms of its explanatory power using the coefficient of determination (R^2). R^2 indicates the proportion of variance in an endogenous construct that is explained by its predictors (Hair et al. 2022). The model explained 70.2% of the variance in satisfaction, 58.0% of the variance in continuance intention, and 46.0% of the variance in habit, showing the model has significant explanatory power. In particular, the high R^2 for satisfaction suggests that perceived learning value, perceived enjoyment, relative challenge, and the interaction terms jointly explain a substantial proportion of users' satisfaction with Duolingo. The R^2 for continuance intention also indicates that the model explains a meaningful proportion of users' intention to continue using the app.

Finally, referring to the effect size using f^2 . The results show that satisfaction had the strongest contribution to continuance intention ($f^2 = 0.757$), indicating a large effect. Perceived enjoyment and perceived learning value also made meaningful contributions to satisfaction, with f^2 values of 0.313 and 0.300 respectively. f^2 shows the contribution of each predictor to an endogenous construct when that predictor is removed from the model (Cohen, 1988; Hair et al. 2022). By contrast, habit had only a very small contribution to continuance intention ($f^2 = 0.022$). The interaction between relative challenge and perceived enjoyment had a small effect on satisfaction ($f^2 = 0.036$), suggesting that although the moderation effect is statistically significant, but the contribution magnitude is quite low.

Looking across the structural model diagnostics checks, overall results indicate that the model is fit for hypothesis testing. The VIF results suggest that collinearity is not severe, the SRMR values indicate acceptable approximate fit, and the R^2 values show that the model explains a reasonable amount of variance in the endogenous constructs. At the same time, the f^2 results suggest that the main explanatory strength of the model lies in the satisfaction pathway rather than the habit pathway. Therefore, the hypothesis testing results should be interpreted with particular attention to the relative strength of the evaluative and behavioural mechanisms.

5.3.2 Hypothesis Testing Results

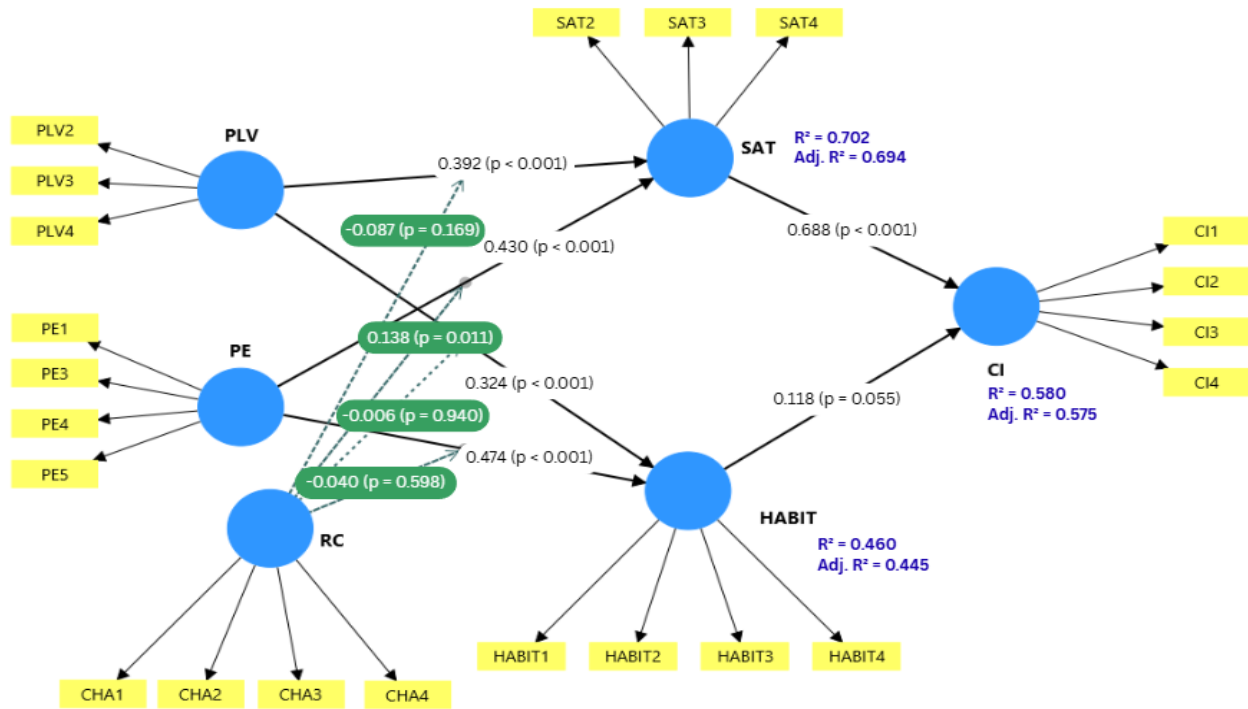


Figure 5.1 Bootstrapping results. Unmoderated (one-tailed) and moderated (two-tailed)

Table 5.6 summarizes the direct and mediation hypothesis prior to moderation effects of relative challenge, while Table 5.7 refers to the summary of post-moderation hypothesis results.. For directional pathways, the hypothesis testing is based on one-tailed tests, while moderation pathways hypothesis testing is based on two-tailed tests, and both are evaluated against a 5% significance level.

Table 5.6 Summary of hypothesis testing (before moderation, one-tailed test)

Effects	Hypothesis	Path	β	SD	t-values	p-value	Conclusion
Evaluative pathways							
Direct effects	H1	PLV $\rightarrow$ SAT	0.392	0.062	6.298	<0.001	Supported
	H2	PE $\rightarrow$ SAT	0.430	0.062	6.878	<0.001	Supported
	H3	SAT $\rightarrow$ CI	0.688	0.059	11.732	<0.001	Supported
Indirect effects	H4a	PLV $\rightarrow$ SAT $\rightarrow$ CI	0.270	0.049	5.483	<0.001	Supported
	H4b	PE $\rightarrow$ SAT $\rightarrow$ CI	0.295	0.05	5.883	<0.001	Supported
Behavioural pathways							
Direct effects	H5	PLV $\rightarrow$ HABIT	0.324	0.078	4.153	<0.001	Supported
	H6	PE $\rightarrow$ HABIT	0.474	0.082	5.782	<0.001	Supported
	H7	HABIT $\rightarrow$ CI	0.118	0.074	1.596	0.110	Not supported
Indirect effects	H8a	PLV $\rightarrow$ HABIT $\rightarrow$ CI	0.038	0.028	1.376	0.169	Not supported
	H8b	PE $\rightarrow$ HABIT $\rightarrow$ CI	0.056	0.035	1.581	0.114	Not supported

Table 5.7 Summary of hypothesis results (after moderation, two-tailed test)

Effects	Hypothesis	Path	β	SD	t-values	p-value	Conclusion
Evaluative pathways							
Moderation effects	H9	RC $\times$ PLV $\rightarrow$ SAT	-0.087	0.063	1.376	0.169	Not supported
	H10	RC $\times$ PE $\rightarrow$ SAT	0.138	0.054	2.532	0.011	Supported
Moderated mediation effects	H11a	RC $\times$ PLV $\rightarrow$ SAT $\rightarrow$ CI	-0.060	0.044	1.349	0.177	Not supported
	H11b	RC $\times$ PE $\rightarrow$ SAT $\rightarrow$ CI	0.095	0.039	2.432	0.015	Supported
Behavioural pathways							
Moderation effects	H12	RC $\times$ PLV $\rightarrow$ HABIT	-0.006	0.08	0.075	0.940	Not supported
	H13	RC $\times$ PE $\rightarrow$ HABIT	-0.040	0.075	0.528	0.598	Not supported
Moderated mediation effects	H14a	RC $\times$ PLV $\rightarrow$ HABIT $\rightarrow$ CI	-0.001	0.012	0.061	0.952	Not supported
	H14b	RC $\times$ PE $\rightarrow$ HABIT $\rightarrow$ CI	-0.005	0.01	0.449	0.654	Not supported

A. Evaluative Pathway

Aligning with previous structural model assessment results, the outcome for evaluative pathways show really strong significance. The results provide strong support for the evaluative pathway. Perceived learning value had a positive and significant effect on satisfaction ($\beta = 0.392$, $t = 6.298$, $p < 0.001$), supporting H1. Perceived enjoyment also had a positive and significant effect on satisfaction ($\beta = 0.430$, $t = 6.878$, $p < 0.001$), supporting H2. These findings favourably support the utilitarian-hedonic hybridity in gamified learning. Users are more satisfied with Duolingo when they perceive the app as useful for language learning and enjoyable to use.

Satisfaction had a strong positive effect on continuance intention ($\beta = 0.688$, $t = 11.732$, $p < 0.001$), supporting H3. This was the strongest direct effect in the model, suggesting that satisfaction is the main immediate predictor of users' intention to continue using Duolingo.

The mediation results further support the satisfaction effects. Perceived learning value had a significant indirect effect on continuance intention through satisfaction ($\beta = 0.270$, $t = 5.483$, $p < 0.001$), supporting H4a. Perceived enjoyment also had a significant indirect effect on continuance intention through satisfaction ($\beta = 0.295$, $t = 5.883$, $p < 0.001$), supporting H4b. Therefore, the results show that perceived learning value and perceived enjoyment influence continuance intention through users' satisfaction with Duolingo. Overall, the evaluative post-adoption pathway shows solid significant results in predicting continuance intention.

B. Behavioural pathway

Behavioural pathways results share the same significance results as evaluative for perceived learning value and perceived enjoyment to habit formation. Perceived learning value had a positive and significant effect on habit ($\beta = 0.324$, $t = 4.153$, $p < 0.001$), supporting H5. Perceived enjoyment also had a positive and significant effect on habit ($\beta = 0.474$, $t = 5.782$, $p < 0.001$), supporting H6. These findings suggest that users are more likely to develop habitual Duolingo use when they also perceive the app is useful for learning and enjoyable.

However, habit did not significantly predict continuance intention ($\beta = 0.118$, $t = 1.596$, $p = 0.110$), so H7 was not supported. This indicates that although users may develop a sense of routine or automaticity in using Duolingo, it did not significantly translate into continuance intention in the main model.

The mediation results through habit were also not significant. The indirect effect of perceived learning value on continuance intention through habit was not significant ($\beta = 0.038$, $t = 1.376$, $p = 0.169$), so H8a was not supported. Similarly, the indirect effect of perceived enjoyment on continuance intention through habit was not significant ($\beta = 0.056$, $t = 1.581$, $p = 0.114$), so H8b was not supported. Therefore, the behavioural pathway is only partially supported, where perceived learning value and perceived enjoyment predicted habit, but habit did not significantly carry these effects forward to continuance intention.

C. Moderation effects of relative challenge

The moderation results show that relative challenge played a selective role in the model. Relative challenge did not significantly moderate the relationship between perceived learning value and satisfaction ($\beta = -0.087$, $t = 1.376$, $p = 0.169$), so H9 was not supported. However, relative challenge significantly moderated the relationship between perceived enjoyment and satisfaction ($\beta = 0.138$, $t = 2.532$, $p = 0.011$), supporting H10. This suggests that the effect of perceived enjoyment on satisfaction varies depending on users' perceived relative challenge.

The moderation effects on the behavioural pathway were not significant. Relative challenge did not moderate the relationship between perceived learning value and habit ($\beta = -0.006$, $t = 0.075$, $p = 0.940$), so H12 was not supported. It also did not moderate the relationship between perceived enjoyment and habit ($\beta = -0.040$, $t = 0.528$, $p = 0.598$), so H13 was not supported. These findings suggest that relative challenge does not significantly shape the way perceived learning value or perceived enjoyment translates into habitual use.

D. Moderated mediation effects

The moderated mediation results further show that relative challenge significantly moderated only one indirect pathway. The moderated indirect effect of perceived enjoyment on continuance intention through satisfaction was significant ($\beta = 0.095$, $t = 2.432$, $p = 0.015$), supporting H11b. This indicates that the enjoyment–satisfaction–continuance pathway depends on users' perceived relative challenge.

In contrast, the moderated indirect effect of perceived learning value on continuance intention through satisfaction was not significant ($\beta = -0.060$, $t = 1.349$, $p = 0.177$), so H11a was not supported. The moderated mediation effects through habit were also not significant. Relative challenge did not significantly moderate the indirect effect of perceived learning value on continuance intention through habit ($\beta = -0.001$, $t = 0.061$, $p = 0.952$), so H14a was not supported. It also did not significantly moderate the indirect effect of perceived enjoyment on continuance intention through habit ($\beta = -0.005$, $t = 0.449$, $p = 0.654$), so H14b was not supported.

In summary, the hypothesis testing results show that the evaluative pathway was strongly supported, while the behavioural pathway was only partially supported. Perceived learning value and perceived enjoyment both increased satisfaction and habit, but only satisfaction significantly predicted continuance intention. Relative challenge played a limited but meaningful role: it moderated the relationship between perceived enjoyment and satisfaction, and also moderated the indirect effect of perceived enjoyment on continuance intention through satisfaction. This suggests that relative challenge is most relevant in the hedonic evaluative pathway, rather than in the learning-value pathway or the habit-based pathway.

5.4 Robustness Check (Replacing Habit with Streak)

This robustness check was conducted to examine whether the main conclusions remain consistent when the behavioural pathway is represented by an app-based behavioural measure instead of self-reported automaticity.

Table 5.8 summarises the main differences between the original habit model and the alternative streak model. The full direct and indirect effect results for the streak model are provided in Appendix C.

Table 5.8 Key comparison between habit model and streak robustness model

Key comparison	HABIT	STREAK	Highlight
PLV → SAT	Significant	Significant	Learning value consistently increases satisfaction.
PE → SAT	Significant	Significant	Enjoyment consistently increases satisfaction.
SAT → CI	Significant	Significant	Satisfaction remains the strongest continuance pathway.
PLV → HABIT/STREAK	Significant	Not significant	Learning value supports psychological habit, but not streak behaviour.
PE → HABIT/STREAK	Significant	Significant	Enjoyment supports both habit formation and streak engagement.
HABIT/STREAK → CI	Not significant	Significant	Streak predicts continuance better than self-reported habit.
PLV/PE → HABIT/STREAK → CI	Not significant	Not significant	Neither habit nor streak works as a mediator.
PLV/PE → SAT → CI	Significant	Significant	Satisfaction remains the key mediation mechanism.
RC × PE → SAT	Significant	Significant	Relative challenge consistently moderates enjoyment → satisfaction.
RC-moderated mediation	Not significant	Not significant	Relative challenge does not shape habit/streak pathways.

5.4.1 Robustness Check: Habit VS Streak

The streak model central finding is consistent with the main habit model, in which the evaluative pathway through satisfaction remains supported. In both models, perceived learning value and perceived enjoyment significantly predicted satisfaction, and satisfaction significantly predicted continuance intention. This shows that satisfaction remains the strongest and most consistent mechanism explaining continuance intention, regardless of whether repeated use is measured as habit or streak

The first key similarity is that perceived enjoyment remains important in both models. In the habit model, perceived enjoyment significantly predicted habit. In the streak model, perceived enjoyment also significantly predicted streak ($\beta = 0.217$, $p = 0.031$). This suggests that enjoyable app experiences may encourage repeated engagement, whether this is reflected psychologically as habit or behaviourally as streak maintenance.

The second similarity is that the satisfaction mediation pathway remains significant. In the streak model, both $PLV \rightarrow SAT \rightarrow CI$ ($\beta = 0.285$, $p < 0.001$) and $PE \rightarrow SAT \rightarrow CI$ ($\beta = 0.314$, $p < 0.001$) were supported. This is consistent with the habit model, where satisfaction also mediated both relationships. Therefore, the robustness check reinforces the conclusion that perceived learning value and perceived enjoyment influence continuance intention mainly through satisfaction.

Another important similarity is that the indirect effects through streak were not supported. $PLV \rightarrow STREAK \rightarrow CI$ was not significant ($\beta = 0.014$, $p = 0.172$), and $PE \rightarrow STREAK \rightarrow CI$ was also not significant ($\beta = 0.024$, $p = 0.097$). This means that although streak directly predicts continuance intention, it does not operate as a mediation mechanism linking perceived learning value or perceived enjoyment to continuance intention. In other words, streak appears to function more as an independent behavioural driver than as a pathway through which app perceptions become continuance intention.

The main difference concerns the behavioural pathway. In the habit model, both perceived learning value and perceived enjoyment significantly predicted habit. However, habit did not

significantly predict continuance intention. In the streak model, the pattern changed: perceived learning value did not significantly predict streak ($\beta = 0.127$, $p = 0.107$), but streak significantly predicted continuance intention ($\beta = 0.113$, $p = 0.020$). This suggests that users' perception of learning value may contribute to psychological habit formation, but it does not necessarily translate into streak maintenance. By contrast, once users have an active streak, this behavioural record appears to be associated with stronger continuance intention.

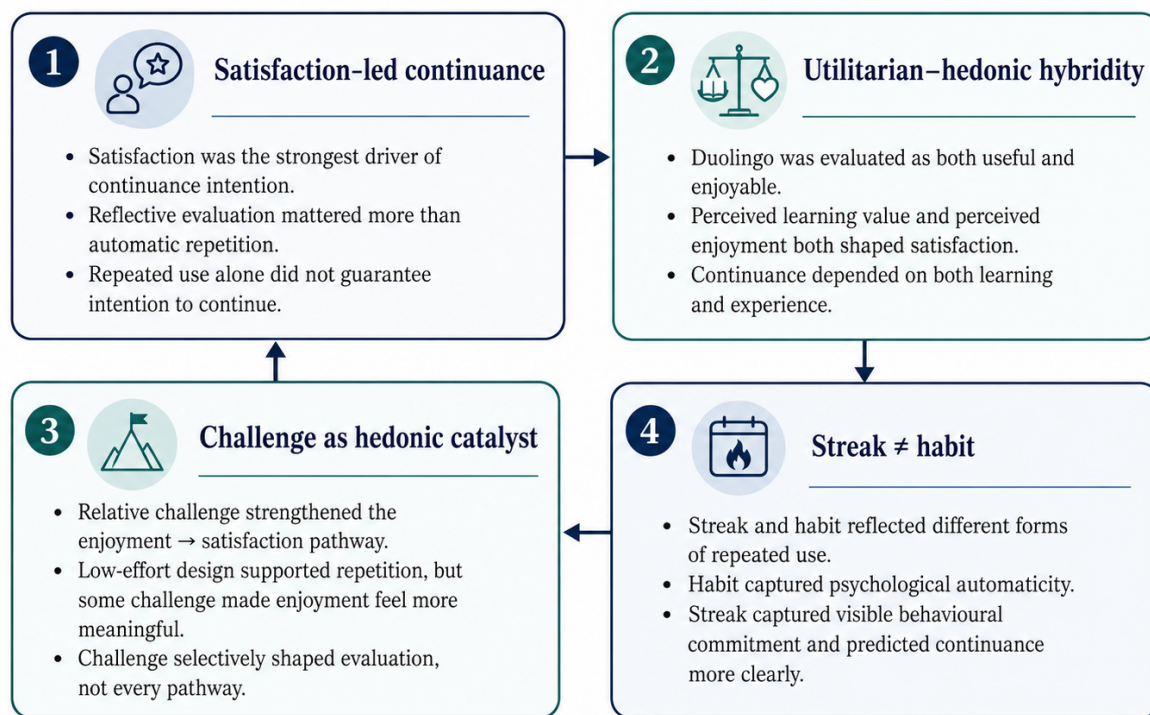
The moderation results also remained selective. Similar to the habit model, relative challenge significantly moderated only the relationship between perceived enjoyment and satisfaction ($\beta = 0.139$, $p = 0.010$). However, the moderated mediation effect through satisfaction was no longer supported in the streak model. Relative challenge did not significantly moderate the indirect effect of perceived enjoyment on continuance intention through satisfaction ($\beta = -0.007$, $p = 0.608$), nor did it moderate any streak-based indirect pathway. Therefore, the role of relative challenge remains limited and should be interpreted cautiously.

To conclude, the robustness check shows that replacing habit with streak does not change the significant role of satisfaction which means it remains that Duolingo continuance is primarily explained by users' satisfaction with the app experience. However, the streak model reveals another important insight: behavioural streak is a stronger direct predictor of continuance intention than self-reported habit, while perceived learning value does not significantly drive streak maintenance. This suggests that habit and streak capture different aspects of repeated use: habit reflects psychological automaticity, whereas streak reflects behavioural commitment or streak-preservation behaviour.

6 Discussion

This chapter discusses the meaning of the empirical findings in relation to the research objectives and prior literature. It explains the theoretical and practical implications of the study, including the roles of satisfaction, habit, relative challenge, and streak in Duolingo continuance. The chapter also outlines the study's limitations, future research directions, and overall conclusion.

Key discussion insights from the final model



Source: Developed from the study findings.

Figure 6.1 Summary of key discussion insights from the model results (AI-generated)

6.1 Satisfaction-Led Continuance: Why Reflective Evaluation Still Matters

The main objective of this study concerns the debate between satisfaction as an evaluative mechanism and habit as a behavioural mechanism in explaining continuance intention. This distinction is grounded in post-adoption research, where satisfaction reflects users' conscious evaluation after system use, while habit captures routinised or automatic use developed through repetition (Bhattacharjee, 2001; Limayem et al. 2007; Venkatesh et al. 2012). In overview, the hypothesis testing showed satisfaction strongly predicted continuance intention than habit. However the non-significant effects of habit do not necessarily reject the importance of routine. Both perceived learning value and perceived enjoyment still showed significant predictive power to habit, securing their important roles as antecedents to continuance. Nonetheless, Duolingo's fun learning design around game features showed more effects in aiding a user's evaluation whether the experience is worth continuing or not. The key point is, Duolingo continuance is not merely a matter of behavioural inertia. It remains dependent on users' overall evaluation of the app experience.

Another possible reason for the non-significant results, habit might be highly context-specific and may not always translate into future intention in the same way across other similar mobile app studies. For example Hsiao et al. (2016) found that both satisfaction and habit mediated continuance intention in mobile social apps. Similarly Wang and Lin (2021) also integrated habit into mobile learning app continuance, suggesting that repeated-use mechanisms can matter in educational technologies. In gamified language learning, forming a routine may not be enough; the routine must also be perceived as educationally meaningful before it becomes a strong reason to continue.

Next, understanding the difference is also important because learning a language involves intention to learn, expectations of tangible growth, and personal goals (Brown and Lee, 2025). Users may repeat Duolingo lessons, but they may still evaluate whether this repetition contributes to meaningful learning. A recent qualitative study by Başaran (2025) found that

repetitive and decontextualized language-learning exercises, limited opportunities for productive language use, and outdated instructional methods might contribute to learner fatigue and hinder deeper acquisition. This also aligns with Sabani et al.'s (2026) distinction between reflective use and automatic repetition in education technology contexts. If users are more reflective, especially in a sample largely composed of educated young adults, automatic use may not be sufficient to form a strong future intention. Satisfaction becomes the point where repeated experience is judged as worthwhile or not.

Therefore, the contribution of this finding is not implying that satisfaction is “better” than habit, nor diminishing the importance of repeated use. Rather, it shows that in gamified mobile language learning, repeated behaviour may need to pass through a positive evaluative judgement before it becomes a stable intention to continue. Gamification studies commonly position repeated engagement as a favourable outcome of game-like features: where points, rewards, goals, feedback, and progress indicators are used to encourage continued participation (Deterding et al. 2011; Domínguez et al. 2013; Oliveira et al. 2022). However, according to present findings, this assumption is not entirely supported, therefore should not be interpreted as successful continuance. Perhaps in the context of language learning, although users frequently return to the app, continuance intention resonates more if the repetitive experience is evaluated as worthwhile and meaningful in helping the users progress in their language skills.

6.2 Duolingo Continuance Depends On Both Learning Value And Enjoyment

Early technology acceptance and continuance models often emphasised usefulness, performance expectancy, or functional value (Davis, 1989; DeLone and McLean, 1992, 2003; Bhattacharjee, 2001). However, later consumer technology and mobile app research increasingly shows that hedonic motivation and enjoyment are also important, especially when use is voluntary and repeated (Venkatesh, et al. 2012; Hsiao, et al. 2016). Apparently, the strong significance of both perceived learning value and perceived enjoyment in predicting satisfaction and habit aligns with this hybrid interpretation. This shows Duolingo users do not evaluate the app only as a language-learning tool, but also as an enjoyable digital experience.

This aligns with the utilitarian–hedonic value perspective. Utilitarian value refers to functional, instrumental, and goal-oriented benefits, while hedonic value refers to pleasure, enjoyment, and experiential gratification (Babin, et al. 1994; Batra and Ahtola, 1991). In this study, perceived learning value represents the utilitarian dimension because it captures whether users believe Duolingo supports their learning progress. Perceived enjoyment represents the hedonic dimension because it captures whether users find the app fun, entertaining, and engaging. The results extended this dual-value principle into gamified mobile language learning, where both dimensions are intentionally built into the app experience. This is also aligning with the Duolingo Method: the app design is to make language learning effective and enjoyable (Freeman et al. 2023).

The results are also consistent with recent technology continuance studies across different contexts. Han and Zo (2023) showed that utilitarian and hedonic values both shape satisfaction in mobile health app continuance. Jo (2022) examined both value dimensions in AI personal assistant continuance, while Tseng et al. (2024) framed online English-learning continuance through utilitarian and hedonic perspectives. Together, these studies suggest that continued use of modern digital services is rarely driven by usefulness or enjoyment alone. In Duolingo, this

hybridity is especially important because users must perceive the app as educationally useful, but also enjoyable enough to support repeated practice.

However, the non-significant relationship of habit-to-continuance reveals an important caveat and changes how gamified learning is understood. While game features such as XP, streaks, goals and interactive feedback may increase enjoyment and incentivise repeated engagement (Deterding et al. 2011; Domínguez et al. 2013), their strongest contribution to continuance in this study operated through satisfaction or positive user's post-use evaluation. This means, that gamification should not be evaluated only by whether it increases engagement. It should also be based on meeting users' expectations in the overall learning experience.

Therefore, the results demonstrate Duolingo is a gamified learning platform that relies on the hybridity of both utilitarian and hedonic values. Enjoyment helps make repeated practice more attractive, while learning value gives repetition a credible educational purpose. If Duolingo is useful but not enjoyable, repeated engagement may be difficult to sustain, which supports the strongest explanation of continuance therefore comes from the interaction of both values, translated through satisfaction.

6.3 Relative Challenge: Why “Easy To Repeat” Is Not Always Enough

The Duolingo Method described by Freeman et al. (2023) explains that breaking learning into small, manageable chunks lowers the barrier to starting, hence making it easier for learners to complete “one more lesson”. This reduced effort approach within the design principle is to make repetition easy, as mastering a language requires repetition (Brown and Lee, 2025). The descriptive results are consistent with this principle, as respondents generally reported low relative challenge, suggesting that Duolingo was perceived as manageable rather than highly difficult. However, the moderated and moderated mediation findings provide a more selective insight. Relative challenge did not broadly shape all continuance pathways, rather, it only strengthened the relationship between perceived enjoyment and satisfaction, and also the indirect effect from enjoyment to continuance intention through satisfaction.

This finding suggests that within a low-challenge condition, users who perceived relatively more challenge may convert enjoyment into satisfaction. This partially aligns with the classic flow theory which argues that optimal flow experience is achieved depending on the relationship between perceived challenge and perceived skill (Csikszentmihalyi, 1990; Nakamura and Csikszentmihalyi, 2002). Similarly, in recent flow-related studies suggesting overly low challenge may reduce interest, while sufficient challenge may help sustain motivation and perceived progress (Engeser and Rheinberg, 2008; Løvoll and Vittersø, 2014).

This result is connected to the tension within Duolingo's app-based teaching and learning design. Although the reduced effort may help support repeated engagement, it may also risk diminishing some degree of challenge to make the learning experience feel meaningful. In other words, perhaps users may continue using Duolingo briefly because it is enjoyable to repeat, but satisfaction may be negatively impacted if users do not see actual progress in their language-learning such as improvements in listening, speaking, reading and writing. Conversely, if the challenge becomes excessive, enjoyment may decline. The significant results between enjoyment and challenge might propose that the optimal pedagogy design is not about making

learning easy, but to make it enjoyable enough while feeling sufficiently challenged to feel worthwhile.

The non-significant moderation of the learning-value pathway also offers an important perspective. This may suggest users evaluate learning value more directly from Duolingo's lesson plans. For users who believe the perceived learning value is high, it will increase satisfaction regardless of challenge level. Here, challenge is more significant for enjoyment because it may be considered as an implicit trait concerning the experiential quality of the task. A lesson is still useful when it is easy, but the enjoyment may become more satisfying as it helps boost motivation, and provides a sense of progress or achievement from the challenge.

The non-significant moderated mediation of habit also suggests relative challenge is not a strong condition for habit formation. Habit may rely more on routine-enforcement features such as repetition, notification nudges or cues (Limayem et al. 2007; Wood and Neal, 2007; Venkatesh et al. 2012). Users may return to Duolingo automatically because it has become part of their daily routine, not because each lesson feels optimally challenging. In this case, learning conditions may shape users' post-use evaluation more than their automatic use patterns.

Overall, the relative challenge finding contributes to flow-informed continuance research by showing that challenge does not need to represent a full flow experience to matter. In this study, relative challenge functions as a narrow learning condition with a selective effect. Although Duolingo was generally perceived as low in challenge, a slightly higher perceived relative challenge strengthened the pathway from enjoyment to satisfaction and then to continuance intention. This suggests that in gamified mobile language learning, ease may help users return, but some meaningful challenge may make enjoyable repetition feel more satisfying.

6.4 Habit And Streak Forms Of Repeated Use

The key insights from the robustness check is the difference effects between habit and streak. In the main model, habit did not significantly predict continuance intention. However, when habit was replaced with streak, streak significantly predicted continuance intention. This difference suggests that habit and streak are not equivalent forms of repeated use. Streak is a form of visible record of repeated use accumulated over days to incentivise users to continue returning to the app, while habit is a psychological form of repeated use based on experience of automaticity. The records from streak days are mechanised to be accumulated, protected and will be lost if users discontinue using the app, while habit is about reduced conscious effort. This makes a key distinction between implicit repeated use, represented by habit, and explicit repeated use, represented by streak.

A reason for the contrast in significance to continuance results can be observed through the concept of loss aversion and virtual asset. Bulduklu (2019) argues that in the present digital environments incorporate elements of virtual assets to demonstrate accumulated achievements to evoke a sense of perceived value in using the a system. For example the streak days and score leaderboard features in Duolingo as a form of symbolic achievement. The continuance intention significance might be explained by users' behaviour of fear of losing accumulated progress. Users may continue not only because the app has become automatic, but because losing the streak would feel like losing accumulated effort.

A second key difference is the non-significant results of perceived learning predicting the streak in the robustness model, whereas perceived enjoyment predicted the streak more clearly. This insight also reveals a potential disconnect between users' learning progress and gamified persistence. Users may believe Duolingo is useful, but it does not support the reason for maintaining the streak. This situation is aligned with Chen and Tu's (2021) argument that external motivation in online learning can be double-edged. Even though the game features such as rewards, competition, badges or any forms of symbolic progress can encourage users' engagement, they may shift the intention from learning to commitment to maintaining the numbers. In other words, in Duolingo, streaks can motivate users to return to the app, but if

streaks become disconnected from perceived learning value, they may encourage continuity without necessarily strengthening meaningful learning.

Nevertheless, the robustness check should not be interpreted as a definite evidence that streak is a better mechanism than habit. Rather, it shows that repeated use has multiple meanings. While satisfaction represents the reflective evaluation pathway clearly, habit represents psychological automaticity, and streak represents visible behavioural commitment: they are related but not interchangeable. This is potentially the reason why the behavioural pathway in the main model was weak: self-reported automaticity may not explain continuance intention in this sample, but visible streak engagement does. Future continuance research in gamified apps should therefore distinguish between psychological habit, behavioural frequency, streak maintenance, and reward-driven persistence.

Summary

In conclusion, our findings show that Duolingo continuance is best explained through a post-use evaluative pathway, significantly supported by utilitarian-hedonic hybrid perceptions and selectively shaped by the gamified learning conditions along the enjoyment-satisfaction moderated mediation. Although Duolingo effectively creates both perceived learning value and perceived enjoyment, these perceptions matter most when they are translated into satisfaction. This supports ECM while extending it into a gamified mobile language-learning context where hedonic experience is essential.

Other findings also demonstrate contrasting results from the oversimplistic view of gamified retention. Duolingo's design appears capable of supporting routine formation, but habit did not necessarily translate into continuance intention. At the same time, streak showed stronger predictive power in continuance, but did not mediate the effect of learning value. Relative challenge also played a selective role by strengthening the enjoyment-satisfaction pathway, but not every pathway in the model.

6.5 Theoretical Implications

This study contributes to continuance intention research in the context of gamified learning and flow-informed educational technology. The findings and results reinforce the role of satisfaction as a key mechanism of continuance intention, based on the pioneering post-adoption theory, Expectation-Confirmation Model (ECM) by Bhattacharjee 2001. Although Duolingo is designed with persistent game-like features such as streaks, XP, badges and other forms of digital values, satisfaction showed statistically significant results at least once in every direct, indirect, moderated and moderated mediation pathway. The key takeaway is satisfaction remains theoretically more important even when behavioural repetition is actively promoted in gamified environments.

The significance of satisfaction also extended further clarification in the dynamic between evaluative and behavioural post-adoption mechanisms. Earlier studies across different digital contexts demonstrated significant habit effects as predictors of continued (Limayem, et al. 2007; Venkatesh et al. 2012) . Mobile app studies have also found habit to be an important predictor or mediator of continuance intention (Hsiao et al. 2016; Wang and Lin, 2021). However in this study, unlike satisfaction, the self-reported habit did not significantly influence continuance intention. This suggests satisfaction and habit should not be treated as interchangeable mechanisms. In gamified mobile language learning, users may develop routines, but these routines do not necessarily translate into continuance intention unless the overall experience remains satisfying.

The study also confirms the utilitarian-hedonic hybridity in a gamified mobile learning environment, as both perceived learning value and perceived enjoyment jointly predict satisfaction. Earlier continuance models often emphasised usefulness or functional benefits, while later consumer technology models incorporated hedonic motivation (Bhattacharjee, 2001; Venkatesh et al. 2012). The present findings show Duolingo users evaluate the app not only as a platform to learn languages, but also an enjoyable digital experience. Theoretically, this supports the view that gamified learning platforms operate as utilitarian–hedonic hybrids.

Next, the moderation analysis findings added a deeper and more critical perspective to the debate between matching challenge and skill classical flow theory (Csikszentmihalyi, 1990; Nakamura and Csikszentmihalyi, 2002), and the later opposing studies. The later studies caution that perfect balance is not always the only desirable condition, as overly low challenge may reduce interest while some degree of challenge can sustain motivation and perceived progress (Engeser and Rheinberg, 2008; Løvoll and Vittersø, 2014) within the educational technology context. Duolingo was generally experienced as easy or manageable, but among users who perceived slightly more relative challenge, enjoyment became more satisfying. This shows that relative challenge does not influence all post-adoption mechanisms equally; instead, it operates selectively by strengthening the enjoyment–satisfaction pathway..

Finally, the robustness check analysis successfully distinguished key conceptual differences of habit as psychological state of automaticity, from streak as the visible daily records of continued use. The findings showed the concept of repeated use is layered therefore it should not be treated as a single construct. Habit reflects users' subjective automaticity, while streak reflects a platform-visible record of continued use. Since streak predicted continuance intention while habit did not, future continuance research should distinguish between psychological habit, behavioural frequency, and platform-visible achievement.

6.6 Practical Implications

The findings provide several practical implications for Duolingo and similar gamified mobile learning platforms. First, the strong effect of satisfaction suggests that retention strategies should not focus only on bringing users back to the app. Although gamified features can encourage repeated use, users are more likely to intend to continue when they evaluate the app experience positively. Therefore, it should shift its focus from maintaining the streak days, to improving the quality of learning and feedback systems, learning progressions and the overall app experience. Despite the persistent gamified features in encouraging users to repeat use, the repetition needs to meaningfully show credible learning outcomes for users to evaluate the app experience positively.

Next, is concerning Duolingo's learning methods. Companies should ensure a careful balance between maintaining learning value and enjoyment that operate through gamification in learning progress. The game features such as XP, streaks and rewards even though they are great tools to encourage repeated use, Duolingo should ensure they serve a broader learning objective: helping users understand repeated practice contributes to language development. For example, rewards could be linked more clearly to vocabulary retention, grammar mastery, pronunciation improvement, listening comprehension, or personalised learning goals. This would help transform gamified motivation into perceived learning value. A mobile learning design that is entertaining but provides poor learning values may not sustain satisfaction, while a design that is pedagogically strong but feels boring may not support long-term engagement.

The third point is companies should avoid reducing difficulty too low only to boost engagement metrics. The role of relative challenge suggests that enjoyment becomes more satisfying when users perceive an appropriate level of difficulty. Bite-sized lessons may lower the barrier to participation, but lessons still need to feel meaningful and sufficiently challenging. Provided every user has different prior abilities, perceived importance in achieving goals and resilience, the same activity can feel either too easy or too difficult for another (Csikszentmihalyi, 1990; Nakamura and Csikszentmihalyi, 2002; Engeser and Rheinberg, 2008; Brown and Lee, 2025). Presently Duolingo already uses adaptive progression where lesson difficulty changes according

to users' course progress and performance. However, based on the findings, personalization of difficulty level should also help users interpret challenges as meaningful progress. Suggested intervention for example is offering optional challenge modes, progress diagnostics across different aspects of language mastery, and feedback that clearly summarizes learning instructions: what a learner has improved and what they are ready to attempt next. In this sense, challenge should be designed as a motivating learning condition rather than an obstacle to retention.

Finally, businesses create new forms of digital value that connect streak milestone to clear learning progress, such as words reviewed, skills strengthened, or grammar patterns practised. Reflective streak summaries could help users interpret streaks not merely as engagement commitment, but as evidence of actual learning commitment and progress. In this way, streaks would function not only as a retention mechanism, but also as a meaningful representation of language-learning development.

6.7 Future Research Area

New future research can be potentially explored in several areas. Firstly, given the different forms of repeated use between habit and streak. Future studies could examine both habit and streak as mechanisms in parallel, paired with the same utilitarian-hedonic antecedents. Potential outcome areas that can be expanded are their effects on streak maintenance, use frequency, loss aversion and behavioural persistence together. This allows us to understand better the underlying reasons for continuance: to protect accumulated progress or to form a routine.

Another future research area is to explore additional post-adoption mechanism pairings beyond the evaluative and behavioural routes examined in this study. Although the present study shows that habit was a weaker mechanism than satisfaction, gamified mobile learning may still be explained through other social, technological, relational, or experiential pathways. This is consistent with broader continuance research showing that continued use is shaped by psychological, technological, social, and behavioural factors, as well as by different mediating and moderating mechanisms (Yan et al. 2021; Franque et al. 2021). Future studies could therefore test alternative mediator pairings, such as satisfaction and trust, satisfaction and attachment, social presence and community support, feedback quality and system quality, personalisation and task-technology fit, or trust in AI and perceived responsiveness. Exploring such alternative mechanism pairings may help explain how utilitarian and hedonic perceptions translate into continued use in ways not fully captured by satisfaction and habit alone.

A third potential future research area studies could develop on the optimal magnitude of the challenge-skill gap against the classical flow theory in the areas of gamified learning. This research do not exclude the importance of well-matched challenge and skill to moderate the perceived learning and enjoyment pathways, however, an optimally lower or higher challenge could have significant effects on satisfaction. Researchers could also examine nonlinear effects, where very low and very high challenge may reduce satisfaction, while moderate challenge produces the strongest engagement and continuance.

Finally, future research can also be explored bridging other digital contexts with education, such as AI-assisted digital learning environments. Platforms such as Duolingo are increasing adopting AI features to grow the platforms at scale, such as AI-generated course, chatbots, adaptive lessons and personalized feedback. On top of gaming features, these AI features may also affect users' evaluation of learning value and enjoyment through satisfaction on continuance intention to use. Other potential constructs such as AI trust, data privacy, authenticity and transparency all could extend continuance research within the mobile language learning platforms.

Overall, future research should move beyond simple retention models and examine how evaluative, behavioural, social, technological, and learning-condition mechanisms interact. This would provide a richer understanding of why users continue using gamified mobile learning platforms and how such platforms can support both engagement and meaningful learning progress.

7 Conclusion

By drawing from continuance theory, habit theory, gamified learning literatures and flow concepts, this study aims to investigate continuance intention to use Duolingo from a post-adoption perspective. The study is driven by the main research question: how perceived learning value and perceived enjoyment influence continuance intention through satisfaction and habit, and whether these relationships vary according to users' perceived relative challenge.

Current findings show that evaluative pathway, represented by satisfaction results the highest significance in explaining continuance intention in Duolingo. Although both perceived learning value and perceived enjoyment significantly influence satisfaction and habit formation, satisfaction's effects remain consistent along either unmoderated or moderated models. These results strongly support the important role of satisfaction as the central mechanism in post-adoption continuance ECM theory. In other words, users remain using the app mainly due to their positive evaluation over the learning experience in Duolingo.

Based on the dual-significant results by perceived learning value and perceived enjoyment on satisfaction and habit, this highlights the hybrid nature of user perceptions in gamified mobile language learning from both utilitarian hedonic perspectives. This notion suggests the long-term success of language-learning platforms is from balancing education quality with engaging user experience. The findings therefore support the argument that gamified learning platforms position themselves within a unique space between education and entertainment, where both forms of value are necessary for sustained engagement.

Next, the moderation pathways driven by the role of relative challenge provided a deeper perspective on the classical flow theory. Duolingo's learning condition reveals a double-edged phenomena. Although low relative challenge can effectively promote repeated use, but some degree of challenge is also required for users to feel meaningful learning progress. Findings showed relative challenge selectively strengthened the relationship between perceived enjoyment and satisfaction, but it did not sufficiently support perceived learning value pathway. The findings therefore contribute to the growing view that effective gamified learning is not simply about reducing effort, but about maintaining a balance between accessibility, engagement, and meaningful challenge.

The robustness check offers an additional contribution by distinguishing psychological habit from behavioural streak engagement in predicting continuance intention. While self-reported habit did not significantly predict continuance intention, streak did. This suggests that visible behavioural commitments and psychological automaticity may represent different forms of repeated use. Consequently, continuance intention in Duolingo appears to be influenced not only by users' evaluations of the learning experience, but also by platform mechanisms that encourage ongoing engagement.

By integrating perceived learning value, perceived enjoyment, satisfaction, habit, relative challenge, and streak-based engagement into a single framework, this study contributes to a more comprehensive understanding of continuance intention in Duolingo and similar gamified learning environments. As mobile learning platforms continue to evolve through gamification and artificial intelligence, understanding why users continue learning over time will remain an important challenge for both researchers and practitioners.

8 Use of AI

This study made responsible use of artificial intelligence tools during the writing and refinement process. AI was used mainly to support language clarity, paragraph structure, academic tone, and the organisation of arguments. It was also used to help identify areas where wording could be improved or where additional citation support might be needed. Additionally, AI was also used to support the generation of some visuals intended to provide summary highlights of several sections throughout the report.

AI was not used to replace the originality of the study. The research topic, research problem, theoretical framework, construct selection, survey design, data collection, data analysis, interpretation of findings, and final conclusions were developed through academic judgement.

AI was also not used to fabricate data, generate survey responses, manipulate results, or conduct the statistical analysis independently. The empirical analysis, including data screening, SmartPLS analysis, hypothesis testing, and robustness check, was conducted and verified by the researchers. Any AI-assisted text was reviewed, revised, and accepted only when it accurately reflected the study's intended meaning, theoretical position, and empirical findings.

Because AI tools can produce inaccurate claims, unsupported statements, or incorrect references, all AI-assisted suggestions were treated critically. Sources, arguments, and interpretations were checked and revised where necessary. The final responsibility for the content, referencing, analysis, and conclusions of this thesis remains with the researchers.

References

- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, vol. 50, no. 2, pp. 179–211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- Ally, M. (ed.) (2009). *Mobile Learning: Transforming the Delivery of Education and Training*. Edmonton: AU Press.
- Alshurideh, M., Al Kurdi, B. and Salloum, S.A. (2020) ‘Examining the main mobile learning system drivers’ effects: A mix empirical examination of both the Expectation-Confirmation Model (ECM) and the Technology Acceptance Model (TAM)’, in Hassanien, A.E., Shaalan, K. and Tolba, M.F. (eds.) *Proceedings of the International Conference on Advanced Intelligent Systems and Informatics 2019*. Cham: Springer, pp. 406–417. https://doi.org/10.1007/978-3-030-31129-2_37.
- Anikina, O.V. and Yakimenko, E.V. (2015). Edutainment as a modern technology of education. *Procedia - Social and Behavioral Sciences*, vol. 166, pp. 475–479. <https://doi.org/10.1016/j.sbspro.2014.12.558>
- Babin, B.J., Darden, W.R. and Griffin, M. (1994). Work and/or fun: measuring hedonic and utilitarian shopping value. *Journal of Consumer Research*, vol. 20, no. 4, pp. 644–656. <https://doi.org/10.1086/209376>
- Başaran, S. (2025). Learning Languages With Duolingo: Myth or Reality? *Mesopotamia Journal of Interdisciplinary Studies*, 5(1), 37–47. DOI: <http://dx.doi.org/10.29228/mjis.82323>
- Batra, R. and Ahtola, O.T. (1991). Measuring the hedonic and utilitarian sources of consumer attitudes. *Marketing Letters*, vol. 2, no. 2, pp. 159–170. <https://doi.org/10.1007/BF00436035>
- Bhattacharjee, A. (2001). Understanding information systems continuance: an expectation-confirmation model. *MIS Quarterly*, vol. 25, no. 3, pp. 351–370. <https://doi.org/10.2307/3250921>
- Brown, H.D. and Lee, H. (2025). *Principles of Language Learning and Teaching: A Course in Second Language Acquisition*. 7th edn. New York: Routledge. <https://doi.org/10.4324/9781003494294>
- Bryman, A. (2016) *Social Research Methods*. 5th edn. Oxford: Oxford University Press.

- Bulduklu, Y. (2019). Mobile games on the basis of uses and gratifications approach: A comparison of the mobile game habits of university and high school students. *New Media & Society*, vol. 25, no. 5-6, pp. 901–917. <https://doi.org/10.1177/135485651774815>
- Business of Apps (2026). Duolingo Revenue and Usage Statistics (2026). Available at: <https://www.businessofapps.com/data/duolingo-statistics/> (Accessed: 29 May 2026).
- Cenamor, J., Parida, V. and Wincent, J. (2019) ‘How entrepreneurial SMEs compete through digital platforms: The roles of digital platform capability, network capability and ambidexterity’, *Journal of Business Research*, vol. 100, pp. 196–206. <https://doi.org/10.1016/j.jbusres.2019.03.035>.
- Chang, H.H. (2010). Task-technology fit and user acceptance of online auction. *International Journal of Human-Computer Studies*, vol. 68, no. 1–2, pp. 69–89. <https://doi.org/10.1016/j.ijhcs.2009.09.010>
- Chang, C.-C., Liang, C., Shu, K.-M. and Chiu, Y.-C. (2015) ‘Alteration of influencing factors of e-learning continued intention for different degrees of online participation’, *The International Review of Research in Open and Distributed Learning*, vol. 16, no. 4, pp. 33–61. <https://doi.org/10.19173/irrodl.v16i4.2084>.
- Chen, C.-C. and Tu, H.-Y. (2021). The effect of digital game-based learning on learning motivation and performance under social cognitive theory and entrepreneurial thinking. *Frontiers in Psychology*, vol. 12, article 750711. <https://doi.org/10.3389/fpsyg.2021.750711>
- Cheng, M. and Yuen, A.H.K. (2018). Student continuance of learning management system use: a longitudinal exploration. *Computers & Education*, vol. 120, pp. 241–253. <https://doi.org/10.1016/j.compedu.2018.02.004>
- Cohen, J. (1988) *Statistical Power Analysis for the Behavioral Sciences*. 2nd edn. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*. New York: Harper & Row.
- Davis, F. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*. 13. 319–. 10.2307/249008
- De la Vall, R.R.F. and Araya, F.G. (2023). Exploring the benefits and challenges of AI-language learning tools. *International Journal of Social Sciences and Humanities Invention*, vol. 10, no. 1, pp. 7569–7576, <https://doi.org/10.18535/ijsshi/v10i01.02>

- DeLone, W.H. and McLean, E.R. (1992). Information systems success: the quest for the dependent variable. *Information Systems Research*, vol. 3, no. 1, pp. 60–95.
<https://doi.org/10.1287/isre.3.1.60>
- DeLone, W.H. and McLean, E.R. (2003). The DeLone and McLean model of information systems success: a ten-year update. *Journal of Management Information Systems*, vol. 19, no. 4, pp. 9–30. <https://doi.org/10.1080/07421222.2003.11045748>
- Deterding, S., Dixon, D., Khaled, R. and Nacke, L. (2011). From game design elements to gamefulness: defining “gamification”. In: *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*. New York: ACM, pp. 9–15.
<https://doi.org/10.1145/2181037.2181040>
- Domínguez, A., Saenz-de-Navarrete, J., de-Marcos, L., Fernández-Sanz, L., Pagés, C. and Martínez-Herráiz, J.-J. (2013). Gamifying learning experiences: practical implications and outcomes. *Computers & Education*, vol. 63, pp. 380–392.
<https://doi.org/10.1016/j.compedu.2012.12.020>
- Duolingo (2023). Duolingo Max Uses OpenAI’s GPT-4 For New Learning Features. Available at: <https://blog.duolingo.com/duolingo-max/> (Accessed: 1 May 2026).
- Duolingo (2024). Duolingo Achieves 54% DAU Growth and 40% Revenue Growth in Third Quarter 2024. Available at:
<https://investors.duolingo.com/news-releases/news-release-details/duolingo-achieves-54-dau-growth-and-40-revenue-growth-third> (Accessed: 1 May 2026).
- Duolingo (2025a). Duolingo Surpasses 50 Million Daily Active Users, Grows DAU 36% and Revenue 41% in Third Quarter 2025 Year over Year. Available at:
<https://investors.duolingo.com/news-releases/news-release-details/duolingo-surpasses-50-million-daily-active-users-grows-dau-36> (Accessed: 1 May 2026).
- Duolingo (2025b). Duolingo Launches 148 New Language Courses. Available at:
<https://investors.duolingo.com/news-releases/news-release-details/duolingo-launches-148-new-language-courses> (Accessed: 1 May 2026).
- Duolingo (2026) Duolingo: The world’s most popular way to learn. Available at:
<https://www.duolingo.com/> (Accessed: 1 May 2026).
- Easterby-Smith, M., Jaspersen, L.J., Thorpe, R. and Valizade, D. (2021) *Management and Business Research*. 7th edn. London: SAGE Publications.

- Engeser, S. and Rheinberg, F. (2008). Flow, performance and moderators of challenge-skill balance. *Motivation and Emotion*, vol. 32, no. 3, pp. 158–172.
<https://doi.org/10.1007/s11031-008-9102-4>
- Essafi, M., Belfakir, L. and Moubtassime, M. (2024). Investigating mobile-assisted language learning apps: Babbel, Memrise, and Duolingo as a case study. *Journal of Curriculum and Teaching*, vol. 13, no. 2, pp. 197–215. <https://doi.org/10.5430/jct.v13n2p197>
- Fang, Y.-H. (2017) ‘Beyond the usefulness of branded applications: insights from consumer–brand engagement and self-construal perspectives’, *Psychology & Marketing*, vol. 34, no. 1, pp. 40–58. <https://doi.org/10.1002/mar.20972>
- Faraon, M., Rönkkö, K., Milrad, M. and Tsui, E. (2025). International perspectives on artificial intelligence in higher education: an explorative study of students’ intention to use ChatGPT across the Nordic countries and the USA. *Education and Information Technologies*, vol. 30, no. 13, pp. 17835–17880. <https://doi.org/10.1007/s10639-025-13492-x>
- Fishbein, M. and Ajzen, I. (1975). *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*. Reading, MA: Addison-Wesley.
- Fornell, C. and Larcker, D.F. (1981) ‘Evaluating structural equation models with unobservable variables and measurement error’, *Journal of Marketing Research*, vol. 18, no. 1, pp. 39–50.
<https://doi.org/10.1177/002224378101800104>.
- Franque, F.B., Oliveira, T., Tam, C. and Santini, F.D.O. (2021). A meta-analysis of the quantitative studies in continuance intention to use an information system. *Internet Research*, vol. 31, no. 1, pp. 123–158. <https://doi.org/10.1108/INTR-03-2019-0103>
- Freeman, C., Kittredge, A., Wilson, H. and Pajak, B. (2023). *The Duolingo Method for App-Based Teaching and Learning*. Duolingo Research Report, pp. 1–19. Available at: https://duolingo-papers.s3.amazonaws.com/reports/Duolingo_whitepaper_duolingo_method_2023.pdf (Accessed: 29 May 2026).
- Fullagar, C.J., Knight, P.A. and Sovern, H.S. (2013) ‘Challenge/skill balance, flow, and performance anxiety’, *Applied Psychology*, vol. 62, no. 2, pp. 236–259.
<https://doi.org/10.1111/j.1464-0597.2012.00494.x>.
- Gardner, B., Abraham, C., Lally, P. and de Bruijn, G.-J. (2012). Towards parsimony in habit measurement: testing the convergent and predictive validity of an automaticity subscale of the Self-Report Habit Index. *International Journal of Behavioral Nutrition and Physical Activity*, vol. 9, article 102. <https://doi.org/10.1186/1479-5868-9-102>

- Giese, J.L. and Cote, J.A. (2000). Defining consumer satisfaction. *Academy of Marketing Science Review*, vol. 2000, no. 1, pp. 1–24, <https://proserv.nu/Docs/Defining%20Customer%20Satisfaction.pdf>
- Goodhue, D.L. and Thompson, R.L. (1995). Task-technology fit and individual performance. *MIS Quarterly*, vol. 19, no. 2, pp. 213–236. <https://doi.org/10.2307/249689>
- Hair, J.F., Hult, G.T.M., Ringle, C.M. and Sarstedt, M. (2022) *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. 3rd edn. Thousand Oaks, CA: SAGE Publications.
- Hair, J.F., Risher, J.J., Sarstedt, M. and Ringle, C.M. (2019) ‘When to use and how to report the results of PLS-SEM’, *European Business Review*, vol. 31, no. 1, pp. 2–24. <https://doi.org/10.1108/EBR-11-2018-0203>.
- Han, K. and Zo, H. (2023). Understanding the mobile healthcare applications continuance: the regulatory focus perspective. *International Journal of Medical Informatics*, vol. 177, article 105161. <https://doi.org/10.1016/j.ijmedinf.2023.105161>
- Hayes, A.F. (2018) *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. 2nd edn. New York: Guilford Press.
- Henseler, J., Dijkstra, T.K., Sarstedt, M., Ringle, C.M., Diamantopoulos, A., Straub, D.W., Ketchen, D.J., Hair, J.F., Hult, G.T.M. and Calantone, R.J. (2014) ‘Common beliefs and reality about PLS: Comments on Rönkkö and Evermann (2013)’, *Organizational Research Methods*, vol. 17, no. 2, pp. 182–209. <https://doi.org/10.1177/1094428114526928>.
- Henseler, J., Ringle, C.M. and Sarstedt, M. (2015) ‘A new criterion for assessing discriminant validity in variance-based structural equation modeling’, *Journal of the Academy of Marketing Science*, vol. 43, no. 1, pp. 115–135. <https://doi.org/10.1007/s11747-014-0403-8>.
- Henseler, J., Ringle, C.M. and Sinkovics, R.R. (2009) ‘The use of partial least squares path modeling in international marketing’, in Sinkovics, R.R. and Ghauri, P.N. (eds.) *Advances in International Marketing*. Bingley: Emerald Group Publishing Limited, vol. 20, pp. 277–319. [https://doi.org/10.1108/S1474-7979\(2009\)0000020014](https://doi.org/10.1108/S1474-7979(2009)0000020014).
- Hsiao, C.-H., Chang, J.-J. and Tang, K.-Y. (2016). Exploring the influential factors in continuance usage of mobile social apps: satisfaction, habit, and customer value perspectives. *Telematics and Informatics*, vol. 33, no. 2, pp. 342–355. <https://doi.org/10.1016/j.tele.2015.08.014>

- Jo, H. (2022). Continuance intention to use artificial intelligence personal assistant: type, gender, and use experience. *Heliyon*, vol. 8, no. 9, article e10662. <https://doi.org/10.1016/j.heliyon.2022.e10662>
- Kim, S.S., Malhotra, N.K. and Narasimhan, S. (2005). Two competing perspectives on automatic use: a theoretical and empirical comparison. *Information Systems Research*, vol. 16, no. 4, pp. 418–432. <https://doi.org/10.1287/isre.1050.0070>
- Li, C.-Y. and Fang, Y.-H. (2019). Predicting continuance intention toward mobile branded apps through satisfaction and attachment. *Telematics and Informatics*, vol. 43, article 101248. <https://doi.org/10.1016/j.tele.2019.101248>
- Liao, H.-L. and Lu, H.-P. (2008) ‘The role of experience and innovation characteristics in the adoption and continued use of e-learning websites’, *Computers & Education*, vol. 51, no. 4, pp. 1405–1416. <https://doi.org/10.1016/j.compedu.2007.11.006>
- Lin, W.-S. and Wang, C.-H. (2012) ‘Antecedences to continued intentions of adopting e-learning system in blended learning instruction: A contingency framework based on models of information system success and task-technology fit’, *Computers & Education*, vol. 58, no. 1, pp. 88–99. <https://doi.org/10.1016/j.compedu.2011.07.008>
- Limayem, M., Hirt, S.G. and Cheung, C.M.K. (2007). How habit limits the predictive power of intention: the case of information systems continuance. *MIS Quarterly*, vol. 31, no. 4, pp. 705–737. <https://doi.org/10.2307/25148817>
- Løvoll, H.S. and Vittersø, J. (2014). Can balance be boring? A critique of the “challenges should match skills” hypothesis in flow theory. *Social Indicators Research*, vol. 115, no. 1, pp. 117–136. <https://doi.org/10.1007/s11205-012-0211-9>
- Massimini, F. and Carli, M. (1988). The systematic assessment of flow in daily experience. In: Csikszentmihalyi, M. and Csikszentmihalyi, I.S. (eds.) *Optimal Experience: Psychological Studies of Flow in Consciousness*. Cambridge: Cambridge University Press, pp. 266–287.
- Mishra, A., Shukla, A., Rana, N.P., Currie, W.L. and Dwivedi, Y.K. (2023). Re-examining post-acceptance model of information systems continuance: a revised theoretical model using MASEM approach. *International Journal of Information Management*, vol. 68, article 102571. <https://doi.org/10.1016/j.ijinfomgt.2022.102571>
- Mpinganjira, M. (2025). Drivers of intentions towards continued use of digital voice assistants: the moderating role of service experience. *Journal of Organizational and End User Computing*, vol. 37, no. 1, article 387836. <https://doi.org/10.4018/JOEUC.387836>

- Nakamura, J. and Csikszentmihalyi, M. (2002). The concept of flow. In: Snyder, C.R. and Lopez, S.J. (eds.) *Handbook of Positive Psychology*. Oxford: Oxford University Press, pp. 89–105.
- Ngo, T.T.A., Phan, T.Y.N., Nguyen, T.K., Le, N.B.T., Nguyen, N.T.A. and Le, T.T.D. (2025). Understanding continuance intention toward the use of AI chatbots in customer service among generation Z in Vietnam. *Acta Psychologica*, vol. 259, article 105468. <https://doi.org/10.1016/j.actpsy.2025.105468>
- Okan, Z. (2003). Edutainment: is learning at risk? *British Journal of Educational Technology*, vol. 34, no. 3, pp. 255–264. <https://doi.org/10.1111/1467-8535.00325>
- Oliveira, W., Hamari, J., Joaquim, S., Toda, A.M., Palomino, P.T., Vassileva, J. and Isotani, S. (2022). The effects of personalized gamification on students' flow experience, motivation, and enjoyment. *Smart Learning Environments*, vol. 8, article 16, <https://doi.org/10.1186/s40561-022-00194-x>
- Oliver, R.L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research*, vol. 17, no. 4, pp. 460–469. <https://doi.org/10.2307/3150499>
- Oliver, R.L. and DeSarbo, W.S. (1988). Response determinants in satisfaction judgments. *Journal of Consumer Research*, vol. 14, no. 4, pp. 495–507. <https://doi.org/10.1086/209131>
- Ooi, K.-B. and Tan, G.W.-H. (2016) 'Mobile technology acceptance model: An investigation using mobile users to explore smartphone credit card', *Expert Systems with Applications*, vol. 59, pp. 33–46. <https://doi.org/10.1016/j.eswa.2016.04.015>.
- Joessel, F., Pichon, S. and Bavelier, D. (2024). A video-game-based method to induce states of high and low flow. *Behavior Research Methods*, vol. 56, pp. 5128–5160. <https://doi.org/10.3758/s13428-023-02251-w>
- Preacher, K.J. and Hayes, A.F. (2008) 'Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models', *Behavior Research Methods*, vol. 40, no. 3, pp. 879–891. <https://doi.org/10.3758/BRM.40.3.879>.
- Sabani, A., Nguyen, H.G., Le, D. and Farah, M. (2026). Unlocking students' continuance intention to use generative AI for educational purposes: the dual role of habit. *Journal of Trade Science*, pp. 1–21. <https://doi.org/10.1108/JTS-08-2025-0038>
- Sarstedt, M., Ringle, C.M. and Hair, J.F. (2022) 'Partial least squares structural equation modeling', in Homburg, C., Klarmann, M. and Vomberg, A. (eds.) *Handbook of Market Research*. Cham: Springer, pp. 587–632. https://doi.org/10.1007/978-3-319-57413-4_15.

- Saunders, M.N.K., Lewis, P. and Thornhill, A. (2019) *Research Methods for Business Students*. 8th edn. Harlow: Pearson.
- Shin, N. (2006). Online learner's "flow" experience: an empirical study. *British Journal of Educational Technology*, vol. 37, no. 5, pp. 705–720.
<https://doi.org/10.1111/j.1467-8535.2006.00641.x>
- Singh, P. and Singh, V. (2024) 'The power of AI: enhancing customer loyalty through satisfaction and efficiency', *Cogent Business & Management*, vol. 11, no. 1, article 2326107.
<https://doi.org/10.1080/23311975.2024.2326107>
- Stone, R.W. and Baker-Eveleth, L. (2013). Students' expectation, confirmation, and continuance intention to use electronic textbooks. *Computers in Human Behavior*, vol. 29, no. 3, pp. 984–990. <https://doi.org/10.1016/j.chb.2012.12.007>
- Traxler, J. (2009). Learning in a mobile age. *International Journal of Mobile and Blended Learning*, vol. 1, no. 1, pp. 1–12. <https://doi.org/10.4018/jmbl.2009010101>
- Tseng, F.-C., Cheng, T.C.E., Li, K.C. and Teng, C.-I. (2023). Using online English learning resources: utilitarian and hedonic perspectives. *Online Information Review*, vol. 48, no. 1, pp. 187–207. <https://doi.org/10.1108/OIR-03-2022-0157>
- Vatanasombut, B., Igbaria, M., Stylianou, A.C. and Rodgers, W. (2008). Information systems continuance intention of web-based applications customers: The case of online banking. *Information & Management*, vol. 45, no. 7, pp. 419–428.
<https://doi.org/10.1016/j.im.2008.03.005>
- Venkatesh, V. and Davis, F.D. (2000). A theoretical extension of the Technology Acceptance Model: four longitudinal field studies. *Management Science*, vol. 46, no. 2, pp. 186–204.
<https://doi.org/10.1287/mnsc.46.2.186.11926>
- Venkatesh, V., Morris, M.G., Davis, G.B. and Davis, F.D. (2003). User acceptance of information technology: toward a unified view. *MIS Quarterly*, vol. 27, no. 3, pp. 425–478.
<https://doi.org/10.2307/30036540>
- Venkatesh, V., Thong, J.Y.L. and Xu, X. (2012). Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Quarterly*, vol. 36, no. 1, pp. 157–178. <https://doi.org/10.2307/41410412>
- Wang, Y.-T. and Lin, K.-Y. (2021). Understanding continuance usage of mobile learning applications: the moderating role of habit. *Frontiers in Psychology*, vol. 12, article 736051.
<https://doi.org/10.3389/fpsyg.2021.736051>

- Wigfield, A. and Eccles, J.S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, vol. 25, no. 1, pp. 68–81.
<https://doi.org/10.1006/ceps.1999.1015>
- Wood, W. and Neal, D.T. (2007) ‘A new look at habits and the habit-goal interface’, *Psychological Review*, vol. 114, no. 4, pp. 843–863.
<https://doi.org/10.1037/0033-295X.114.4.843>.
- Yan, M., Filieri, R. and Gorton, M. (2021). Continuance intention of online technologies: a systematic literature review. *International Journal of Information Management*, vol. 58, article 102315. <https://doi.org/10.1016/j.ijinfomgt.2021.102315>

Appendices

Appendix A - Survey Questionnaire and Measurement Items

Table A1 Summary of Questionnaires

Questions	Source	Scale	Status
<i>Perceived Learning Value</i>			
Duolingo helps to improve my listening, speaking, and reading	Adapted from Tseng and Teng (2014) as cited in Tseng et al (2024)	Strongly disagree → Strongly agree (1 to 7)	Dropped
Duolingo enables me to learn new languages faster.	Adapted from Ooi and Tan (2016) as cited in Ooi and Lee (2018)	Strongly disagree → Strongly agree (1 to 7)	
Using Duolingo increases my learning efficiency	Adapted from Han and Zo (2023)	Strongly disagree → Strongly agree (1 to 7)	
Using Duolingo is very useful for my language learning	Adapted from Han and Zo (2023)	Strongly disagree → Strongly agree (1 to 7)	
<i>Perceived Enjoyment</i>			
Using Duolingo is truly fun.	Adapted from Venkatesh et al. (2012)	Strongly disagree → Strongly agree (1 to 7)	
Using Duolingo is enjoyable.	Adapted from Venkatesh et al. (2012)	Strongly disagree → Strongly agree (1 to 7)	Dropped
Using Duolingo is very entertaining.	Adapted from Venkatesh et al. (2012)	Strongly disagree → Strongly agree (1 to 7)	
I find Duolingo's interactive features enjoyable to use.	Adapted from Shin (2006)	Strongly disagree → Strongly agree (1 to 7)	
Duolingo's learning activities are not boring to me.	Adapted from Shin (2006)	Strongly disagree → Strongly agree (1 to 7)	
<i>Satisfaction</i>			
Duolingo meets my expectations for online language learning.	Singh and Singh (2024); as cited in Ngo et al (2025)	Strongly disagree → Strongly agree (1 to 7)	Dropped
My experience using Duolingo leaves me with a positive impression of the Duolingo company.	Singh and Singh (2024); as cited in Ngo et al (2025)	Strongly disagree → Strongly agree (1 to 7)	
I feel valued as a consumer when using Duolingo to learn a language.	Singh and Singh (2024); as cited in Ngo et al (2025)	Strongly disagree → Strongly agree (1 to 7)	

Overall, I am satisfied with my experience with Duolingo.	Singh and Singh (2024); as cited in Ngo et al (2025)	Strongly disagree → Strongly agree (1 to 7)	
Habit			
Using Duolingo is something I do automatically.	Adapted from Gardner et al. (2012)	Not true for me → Very true for me (1 to 7)	
Using Duolingo is something I do without having to consciously remember.	Adapted from Gardner et al. (2012)	Not true for me → Very true for me (1 to 7)	
I often start using Duolingo without consciously deciding to do so.	Adapted from Gardner et al. (2012)	Not true for me → Very true for me (1 to 7)	
Using Duolingo is something I do without thinking.	Adapted from Gardner et al. (2012)	Not true for me → Very true for me (1 to 7)	
Perceived Challenge			
Duolingo's learning activities are too demanding for me.	Adapted from Shin (2006)	Not at all → Very much so (1 to 7)	
It is difficult for me to understand Duolingo's lesson content.	Adapted from Shin (2006)	Not at all → Very much so (1 to 7)	
It is difficult for me to keep up with Duolingo's learning activities.	Adapted from Shin (2006)	Not at all → Very much so (1 to 7)	
It is hard for me to complete the tasks required in Duolingo.	Adapted from Shin (2006)	Not at all → Very much so (1 to 7)	
Continuance Intention to Use			
Next time I am willing to use Duolingo again to learn a new language.	Adapted from Chang (2010); Vatanasombut et al. (2008) as cited in Lin and Wang (2011)	Strongly disagree → Strongly agree (1 to 7)	
I intend to continue using Duolingo rather than discontinue its use.	Adapted from Bhattacharjee (2001)	Strongly disagree → Strongly agree (1 to 7)	
For future language learning, I would choose Duolingo.	Adapted from Davis (1989)	Strongly disagree → Strongly agree (1 to 7)	
Overall, I intend to continue using the Duolingo app for language learning.	Adapted from Tseng et al (2015) as cited in Tseng et al (2024)	Strongly disagree → Strongly agree (1 to 7)	
Streak			
What is your CURRENT Duolingo streak? (excluding broken streak days)		No current streak 1 - 7 days 8 - 30 days 31 - 100 days More than 100 days	

Table A2 Summary of Questionnaires (Chinese)

Questions	Source	Scale	Status
Perceived Learning Value			
Duolingo 有助于提高我的听、说、读能力。	Adapted from Tseng and Teng (2014) as cited in Tseng et al (2023)	很不同意→ 很同意 (1 to 7)	Dropped
Duolingo 让我能更快地学习新语言	Adapted from Ooi and Tan (2016), as cited in Ooi and Lee (2018)	很不同意→ 很同意 (1 to 7)	
使用 Duolingo 提高了我的学习效率。	Adapted from Han and Zo (2023)	很不同意→ 很同意 (1 to 7)	
使用 Duolingo 对我的语言学习非常有帮助。	Adapted from Han and Zo (2023)	很不同意→ 很同意 (1 to 7)	
Perceived Enjoyment			
使用 Duolingo 非常有趣。	Adapted from Venkatesh et al. (2012)	很不同意→ 很同意 (1 to 7)	
使用 Duolingo 让我感到愉快。	Adapted from Venkatesh et al. (2012)	很不同意→ 很同意 (1 to 7)	Dropped
使用 Duolingo 很有娱乐性。	Adapted from Venkatesh et al. (2012)	很不同意→ 很同意 (1 to 7)	
我觉得 Duolingo 的互动特点很有趣。[单选题]	Adapted from Shin (2006)	很不同意→ 很同意 (1 to 7)	
Duolingo 的学习活动对我来说并不无聊。[单选题]	Adapted from Shin (2006)	很不同意→ 很同意 (1 to 7)	
Satisfaction			
Duolingo 满足了我对在线语言学习的期待。	Singh and Singh (2024); as cited in Ngo et al (2025)	很不同意→ 很同意 (1 to 7)	Dropped
使用 Duolingo 的经历让我对 Duolingo 公司有积极印象。。	Singh and Singh (2024); as cited in Ngo et al (2025)	很不同意→ 很同意 (1 to 7)	
使用 Duolingo 学习语言时, 我觉得自己作为用户受到重视。	Singh and Singh (2024); as cited in Ngo et al (2025)	很不同意→ 很同意 (1 to 7)	
总体而言, 我对使用 Duolingo 的体验感到满意。	Singh and Singh (2024); as cited in Ngo et al (2025)	很不同意→ 很同意 (1 to 7)	
Habit			

使用 Duolingo 是我自动会做的事情。	Adapted from Gardner et al. (2012)	很不符合 → 很符合 (1 to 7)	
使用 Duolingo 时, 我不需要刻意提醒自己。	Adapted from Gardner et al. (2012)	很不符合 → 很符合 (1 to 7)	
我经常在没有刻意决定的情况下就开始使用 Duolingo。	Adapted from Gardner et al. (2012)	很不符合 → 很符合 (1 to 7)	
使用 Duolingo 已经变成一种不假思索的行为。	Adapted from Gardner et al. (2012)	很不符合 → 很符合 (1 to 7)	
Perceived Challenge			
Duolingo 的学习活动对我来说要求太高。	Adapted from Shin (2006)	很不符合 → 很符合 (1 to 7)	
我很难理解 Duolingo 的课程内容。	Adapted from Shin (2006)	很不符合 → 很符合 (1 to 7)	
1我很难跟上 Duolingo 的学习活动。	Adapted from Shin (2006)	很不符合 → 很符合 (1 to 7)	
完成 Duolingo 的任务对我来说很困难。	Adapted from Shin (2006)	很不符合 → 很符合 (1 to 7)	
Continuance Intention to Use			
下次学习新语言时, 我愿意再次使用 Duolingo。	Adapted from Chang (2010); Vatanasombut et al. (2008) as cited in Lin and Wang (2011)	很不同意 → 很同意 (1 to 7)	
我打算继续使用 Duolingo, 而不是停止使用它。	Adapted from Bhattacharjee (2001) as cited in Fang (2016)	很不同意 → 很同意 (1 to 7)	
未来学习语言时, 我会选择 Duolingo。	Adapted from Davis (1989) as cited in Liao and Lu (2008)	很不同意 → 很同意 (1 to 7)	
总体而言, 我打算继续使用 Duolingo 进行语言学习。	Adapted from Tseng et al (2015) as cited in Tseng et al (2023)	很不同意 → 很同意 (1 to 7)	
Streak			
你目前的 Duolingo 连续使用天数是多少		○1-7 天 ○8-30 天 ○31-100 天 ○超过 100 天	

Appendix B - Structural Model Output

Table B1 Collinearity statistics (VIF)

Predictor	Endogenous construct	VIF
Habit	Continuance intention	1.488
Satisfaction	Continuance intention	1.488
Perceived enjoyment	Habit	1.981
Perceived enjoyment	Satisfaction	1.981
Perceived learning value	Habit	1.724
Perceived learning value	Satisfaction	1.724
Relative challenge	Habit	1.645
Relative challenge	Satisfaction	1.645
RC $\times$ PE	Habit	4.947
RC $\times$ PE	Satisfaction	4.947
RC $\times$ PLV	Habit	4.545
RC $\times$ PLV	Satisfaction	4.545

Table B2 Model Fit

	Saturated model	Estimated model
SRMR	0.065	0.071

Table B3 Coefficient of determination

Endogenous construct	R²	p-value
CI	0.580	0.000
HABIT	0.460	0.000
SAT	0.702	0.000

Table B4 Effect size results

Path	f²	p-value
SAT → CI	0.757	0.000
HABIT → CI	0.022	0.496
PE → SAT	0.313	0.003
PLV → SAT	0.300	0.010
PE → HABIT	0.210	0.012
PLV → HABIT	0.113	0.074
RC → SAT	0.005	0.645
RC → HABIT	0.000	0.951
RC × PE → SAT	0.036	0.225
RC × PLV → SAT	0.013	0.506
RC × PE → HABIT	0.002	0.863
RC × PLV → HABIT	0.000	0.997

Appendix C - Robustness Check Using Streak

Table C1 Summary of hypothesis testing for STREAK (before moderation, one-tailed test)

Effects	Hypothesis	Path	β	SD	t-value	p-value	Conclusion
Evaluative pathways							
Direct effects	H1	PLV $\rightarrow$ SAT	0.391	0.062	6.3	<0.001	Supported
	H2	PE $\rightarrow$ SAT	0.431	0.062	6.937	<0.001	Supported
	H3	SAT $\rightarrow$ CI	0.729	0.035	20.709	<0.001	Supported
Indirect effects	H4a	PLV $\rightarrow$ SAT $\rightarrow$ CI	0.285	0.047	6.024	<0.001	Supported
	H4b	PE $\rightarrow$ SAT $\rightarrow$ CI	0.314	0.049	0.049	<0.001	Supported
Behavioural pathways							
Direct effects	H5	PLV $\rightarrow$ STREAK	0.127	0.102	1.244	0.107	Not supported
	H6	PE $\rightarrow$ STREAK	0.217	0.116	1.869	0.031	Supported
	H7	STREAK $\rightarrow$ CI	0.113	0.055	2.06	0.020	Supported
Indirect effects	H8a	PLV $\rightarrow$ STEAK $\rightarrow$ CI	0.014	0.015	0.948	0.172	Not supported
	H8b	PE $\rightarrow$ STREAK $\rightarrow$ CI	0.024	0.019	1.301	0.097	Not supported

Table C2 Summary of hypothesis results for STREAK (after moderation, two-tailed test)

Effects	Hypothesis	Path	β	SD	t-value	p-value	Conclusion
Evaluative pathways							
Moderation effects	H9	$RC \times PLV \rightarrow SAT$	-0.09	0.063	1.428	0.153	Not supported
	H10	$RC \times PE \rightarrow SAT$	0.139	0.054	2.569	0.010	Supported
Moderated mediation effects	H11a	$RC \times PLV \rightarrow SAT \rightarrow CI$	-0.01	0.016	1.41	0.159	Not supported
	H11b	$RC \times PE \rightarrow SAT \rightarrow CI$	-0.007	0.014	2.516	0.608	Not supported
Behavioural pathways							
Moderation effects	H12	$RC \times PLV \rightarrow STREAK$	-0.09	0.121	0.743	0.457	Not supported
	H13	$RC \times PE \rightarrow STREAK$	-0.065	0.113	0.571	0.568	Not supported
Moderated mediation effects	H14a	$RC \times PLV \rightarrow STREAK \rightarrow CI$	-0.01	0.016	0.614	0.539	Not supported
	H14b	$RC \times PE \rightarrow STREAK \rightarrow CI$	-0.007	0.014	0.513	0.608	Not supported

Appendix D - Pilot Survey and Questionnaire Refinement

Table D1 Pilot Study HTMT results

	CHL	CI	HAB	PE	PLV	SAT	CHL x PE	CHL x PLV
CHL								
CI	0.25							
HAB	0.224	0.678						
PE	0.292	0.84	0.448					
PLV	0.33	0.526	0.416	0.502				
SAT	0.383	1.061	0.734	0.969	0.85			
CHL x PE	0.019	0.537	0.257	0.581	0.376	0.47		
CHL x PLV	0.4	0.183	0.145	0.315	0.427	0.24	0.692	

Table D2 Summary of Questionnaires for Pilot Survey

Questions	Source	Scale	Status	Remarks
Perceived Learning				
Duolingo helps to improve my language proficiency	Adapted from Tseng and Teng (2014) as cited in Tseng et al (2023)	Strongly disagree → Strongly agree (1 to 7)	Replaced	Less contextual to Duolingo
Duolingo helps to improve my vocabulary, pronunciation, and grammar	Adapted from Tseng and Teng (2014) as cited in Tseng et al (2023)	Strongly disagree → Strongly agree (1 to 7)	Replaced	Less contextual to Duolingo
Using Duolingo increases my learning efficiency	Adapted from Han and Zo (2023)	Strongly disagree → Strongly agree (1 to 7)	Remain	
Using Duolingo is very useful for my language learning.	Adapted from Han and Zo (2023)	Strongly disagree → Strongly agree (1 to 7)	Remain	
Perceived Enjoyment Value				
Using Duolingo is pleasureable	Adapted from Tseng and Teng (2014) as cited in Tseng et al (2023)	Strongly disagree → Strongly agree (1 to 7)	Replaced	Poor discriminant validity/too broad
Using Duolingo makes me happy	Adapted from Tseng and Teng (2014) as cited in Tseng et al (2023)	Strongly disagree → Strongly agree (1 to 7)	Replaced	Poor discriminant validity/too broad
Using Duolingo is very entertaining.	Adapted from Han and Zo (2023)	Strongly disagree → Strongly agree (1 to 7)	Remain	
Satisfaction				
My overall experience with Duolingo has been very dissatisfied / very satisfied.	Bhattacharjee (2001)	Very dissatisfied → Very satisfied (1 to 7)	Replaced	Poor discriminant validity/too broad
My overall experience with Duolingo has been very displeased / very pleased.	Bhattacharjee (2001)	Very displeased → Very pleased (1 to 7)	Replaced	Poor discriminant validity/too broad
My overall experience with Duolingo has been very frustrating / very contented.	Bhattacharjee (2001)	Very frustrating → Very contented (1 to 7)	Replaced	Poor discriminant validity/too broad
Habit				
Using Duolingo is something I do	Adapted from Gardner	No → Yes (1 to 7)	Remain	

automatically.	et al. (2012)			
Using Duolingo is something I do without having to consciously remember.	Adapted from Gardner et al. (2012)	No → Yes (1 to 7)	Remain	
Using Duolingo is something I start doing before I realize I'm doing it	Adapted from Gardner et al. (2012)	No → Yes (1 to 7)	Replaced	Unclear
Using Duolingo is something I do without thinking.	Adapted from Gardner et al. (2012)	No → Yes (1 to 7)	Remain	
Continuance Intention to Use				
I would continue to use Duolingo	Adapted from Tseng et al (2015) as cited in Tseng et al (2023)	Strongly disagree → Strongly agree (1 to 7)	Replaced	Too broad
I would recommend Duolingo to those who want to learn a language	Adapted from Tseng et al (2015) as cited in Tseng et al (2023)	Strongly disagree → Strongly agree (1 to 7)	Replaced	Too broad
I intend to continue using Duolingo rather than use alternative ways of learning languages (taking a physical class).	Adapted from Han and Zo (2023)	Strongly disagree → Strongly agree (1 to 7)	Replaced	Too broad
I would spend more time using Duolingo	Adapted from Tseng et al (2015) as cited in Tseng et al (2023)	Strongly disagree → Strongly agree (1 to 7)	Replaced	Too broad
I would consider subscribing to Duolingo Super/Max (whichever applies)	Adapted from Tseng et al (2015) as cited in Tseng et al (2023)	Strongly disagree → Strongly agree (1 to 7)	Removed	Not relevant
If I could, I would like to discontinue using Duolingo.	Adapted from Han and Zo (2023)	Strongly disagree → Strongly agree (1 to 7)	Removed	Not relevant
Challenge-Skill Fit				
How challenging do you find Duolingo's language exercises?	Adapted from Pearce et al. (2004)	Strongly disagree → Strongly agree (1 to 7)	Replaced	Changed to Relative Challenge
Is your current language level appropriate for handling Duolingo's language exercises?	Adapted from Pearce et al. (2004)	Strongly disagree → Strongly agree (1 to 7)	Replaced	anged to Relative Challen
Is your current language level appropriate for handling Duolingo's language exercises?	Adapted from Pearce et al. (2004)	Strongly disagree → Strongly agree (1 to 7)	Replaced	anged to Relative Challen

Appendix E - Item Refinement

Table E1 Cross Loadings (before refinement)

	CI	HABIT	PE	PLV	RC	SAT	RC x PLV	RC x PE
CHA1	-0.268	-0.192	-0.351	-0.232	0.791	-0.282	-0.45	-0.47
CHA2	-0.364	-0.175	-0.364	-0.332	0.835	-0.371	-0.494	-0.458
CHA3	-0.29	-0.243	-0.401	-0.276	0.831	-0.279	-0.501	-0.527
CHA4	-0.402	-0.233	-0.463	-0.315	0.859	-0.394	-0.513	-0.536
CI1	0.825	0.388	0.541	0.585	-0.393	0.631	0.236	0.333
CI2	0.852	0.489	0.602	0.559	-0.359	0.631	0.347	0.439
CI3	0.854	0.433	0.509	0.694	-0.322	0.685	0.287	0.394
CI4	0.905	0.448	0.59	0.596	-0.325	0.674	0.289	0.398
HABIT1	0.564	0.831	0.559	0.536	-0.31	0.559	0.245	0.319
HABIT2	0.463	0.887	0.534	0.553	-0.242	0.496	0.226	0.248
HABIT3	0.344	0.818	0.513	0.425	-0.057	0.456	0.165	0.175
HABIT4	0.345	0.89	0.55	0.447	-0.231	0.461	0.271	0.281
PE1	0.556	0.585	0.868	0.589	-0.405	0.631	0.404	0.456
PE2 (removed)	0.568	0.547	0.889	0.616	-0.495	0.679	0.493	0.518
PE3	0.556	0.613	0.88	0.549	-0.404	0.66	0.44	0.432
PE4	0.535	0.458	0.846	0.526	-0.378	0.63	0.469	0.485
PE5	0.605	0.519	0.848	0.631	-0.391	0.697	0.403	0.456
PLV1 (removed)	0.611	0.463	0.632	0.869	-0.367	0.76	0.428	0.491
PLV2	0.606	0.503	0.546	0.805	-0.292	0.652	0.344	0.42

PLV3	0.603	0.5	0.551	0.886	-0.298	0.694	0.294	0.365
PLV4	0.609	0.518	0.573	0.862	-0.244	0.668	0.277	0.347
SAT1 (removed)	0.596	0.46	0.566	0.727	-0.233	0.818	0.203	0.283
SAT2	0.621	0.464	0.639	0.673	-0.343	0.849	0.366	0.453
SAT3	0.593	0.479	0.601	0.62	-0.289	0.82	0.333	0.365
SAT4	0.721	0.528	0.721	0.682	-0.468	0.845	0.48	0.549
RC x PE	0.455	0.303	0.541	0.475	-0.601	0.501	0.903	1
RC x PLV	0.337	0.266	0.509	0.393	-0.591	0.42	1	0.903

Table E2 Discriminant validity HTMT results (before refinement)

	CI	HABIT	PE	PLV	RC	SAT	RC x PLV	RC x PE
CI								
HABIT	0.568							
PE	0.725	0.698						
PLV	0.805	0.652	0.749					
RC	0.463	0.285	0.538	0.402				
SAT	0.875	0.663	0.857	0.936	0.463			
RC x PLV	0.359	0.282	0.533	0.419	0.64	0.449		
RC x PE	0.484	0.318	0.566	0.506	0.651	0.536	0.903	