Effects of Mobile-Assisted Language Learning on Proficiency and Motivation for Chinese ESL Students' Learning French on Duolingo

Hanlu Duan

MSc in Applied Linguistics and Second Language Acquisition, 2023

DECLARATION BY THE CANDIDATE AS AUTHOR OF THE DISSERTATION

1. I understand that I am the owner of this dissertation and that the copyright rests with me unless I specifically transfer it to another person.
2. I allow the Department to deposit on my behalf a copy of this dissertation in the Oxford University Research Archive ('ORA') where it shall be freely available online for use in accordance with ORA's Terms and Conditions of Use [https://ora.ox.ac.uk/terms_of_use].
3. I understand that this dissertation should not contain material that can be used to personally identify individuals or specific groups of individuals (unless permission has been obtained from the individuals) and that such material should be removed before this dissertation is deposited in ORA.
4. I agree to be bound by the terms of the ORA Grant of Non-exclusive Licence [https://ora.ox.ac.uk/deposit_agreements] and I warrant that to the best of my knowledge, making my thesis available on the internet will not infringe copyright or any other rights of any other person or party, nor contain defamatory material.
5. I agree that my dissertation shall be available for download in ORA in accordance with paragraphs 2, 3 and 4 above.

Signed [an electronic signature is sufficient]:	Hanlu Duan

Date:	21 Aug. 2024

Effects of Mobile-Assisted Language Learning on Proficiency and Motivation for Chinese ESL Students' Learning French on Duolingo

Acknowledgement

First and foremost, I would like to express my gratitude to my supervisor, Sara Ratner, for her invaluable guidance and insightful feedback throughout the whole year. Your expertise and encouragement have been instrumental in shaping this thesis, and I sincerely appreciate your unwavering support.

I am also grateful to all the participants who generously shared their time and experiences for this study. Your willingness to contribute has been crucial to the success of this research, and I am truly thankful for your cooperation.

To my family, thank you for your love, patience, and financial support. Your belief in me has been a source of strength and motivation that encourages me throughout my life. Love you all, my mom, dad, grandparents and dog.

Finally, I would like to extend my thanks to my friends and boyfriend. Your understanding and love have been a cornerstone of my determination. I am incredibly fortunate to have you by my side. I appreciate your encouragement and the many moments of relief and distraction you provided during this intense period.

Abstract

With the rapid evolution of technologies, mobile devices have become ubiquitous in daily life, offering unprecedented flexibility in language teaching and learning. A variety of language learning apps claim to boost learning outcomes, enhance learners' motivation and make the experiences enjoyable. This research aims to investigate the effectiveness of mobile-assisted language learning (MALL) platforms, focusing on Duolingo as a representative example.

This study adopts a mixed-method quasi-experimental design involving 18 Chinese college students in the UK. Data were obtained from a pretest and a posttest of French proficiency, a motivation survey guided by Self-Determination Theory (SDT), and semi-structured interviews with 5 participants. The students learned French through English instruction exclusively using Duolingo for 4 weeks in uncontrolled conditions. The results showed that the learners' scores significantly increased after the 4-week learning session, but their progress was not correlated with the time spent on Duolingo. Furthermore, novice learners made more progress than intermediate learners with similar learning time. Participants reported positive attitudes on both extrinsic and intrinsic motivation scales, though neither scale correlated with the progress they made. According to the interview data, Duolingo enhanced learners' motivation, perseverance and learning experiences, but it may not be as effective as traditional instruction due to its unsystematic knowledge, mechanical exercises and slow progress, particularly for more advanced learners. Limitations and implications are also discussed, highlighting the need for further research with robust designs to enhance the understanding of the impacts of MALL platforms.

Keywords: mobile-assisted language learning, mobile learning, language learning apps, Duolingo, Self-Determination Theory

Table of Contents

Quick Reference to Abbreviations

ACTFL: American Council on the Teaching of Foreign Languages

AMS:Academic Motivation Scale

AMTB: Attitude/Motivation Test Battery

App: application

CEFR: Common European Framework of Reference for Languages

CUREC: Central University Research Ethics Committee

ESL: English as a Second Language

EFL: English as a Foreign Language

GBP: Great Britain Pound

ICT: information and communication technology

IMI: Intrinsic Motivation Inventory

L1: first language

L2: second language

L2MSS: L2 Motivational Self System

L3: third language

MALL: mobile-assisted language learning

M-learning: mobile learning

SLA: second language acquisition

SDT: Self-Determination Theory

UNESCO: United Nations Educational, Scientific, and Cultural Organization

WebCAPE: Web based Computer Adaptive Placement Exam

XP: experience points (a kind of rewards on Duolingo)

List of Tables

List of Figures

Chapter 1. Introduction

1.1. Background

As technological advancements continue to reshape the way people gain knowledge, technology-enhanced learning remains a focus in educational research. Smartphones, tablets and other portable devices are becoming ubiquitous around the world, providing individuals with access to extensive information related to almost any field or subject. Moreover, the onset of the COVID-19 pandemic, along with associated quarantine policies, accelerated the increase of mobile learning (m-learning) demands due to the temporary closure of schools. The interruption of this historic epidemic is unprecedented: during the most serious outbreak, schools from pre-primary to upper-secondary levels in 151 countries were closed or partially closed, affecting nearly 1.3 billion students (UNESCO, 2022)[1]. Therefore, the classes were forced to move online with the help of telecommunication technologies, m-learning platforms and personal devices. This shift during the pandemic contributed to the integration of technology into education in the post-COVID-19 pandemic digital era (Jeong, 2023).

Following this trend, mobile-assisted language learning (MALL), as a subcategory of m-learning, also attracted learners' and researchers' attention. It benefited from information and communication technology (ICT) which could transcend the constraints of time and space to deliver educational content, enabling people to learn with great flexibility. However, within the domain of MALL research, the diversity of platforms and features makes it challenging to generalize universally applicable conclusions regarding its impacts on language learning. Prior empirical studies may investigate language learning outcomes influenced by built-in mobile features such as photo-taking (e.g., Liu & Chen, 2015) and text messaging (e.g., Derakhshan, 2011); or the impacts of social media such as Facebook (e.g., Fithriani et al., 2019) and Instagram (e.g., Gonulal, 2019). However, these features and platforms are not specifically designed for language learning, nor can they provide teaching materials by themselves. Therefore, this thesis primarily focuses on apps designed for language learning. The experiment is carried out using arguably the most representative one, Duolingo, which delivers structured language courses in a gamified way.

Duolingo has been chosen for its worldwide accessibility, highly-rated reputation

[1] UNESCO: Global monitoring of school closures
https://webarchive.unesco.org/web/20220629024039/https://en.unesco.org/covid19/educationresponse/

and long-standing leading position in the educational app industry. It was founded in 2011 by Luis von Ahn and Severin Hacker with the mission to "develop the best education in the world and make it universally available" (Duolingo, 2020). Upholding this mission, it is now offering 108 language courses to over 300 million learners all over the world (Duolingo, June 2024)[2]. It is also highly rated 4.7 out of 5 in both the Apple Store and Google Play (until March 2024). Its global popularity and extensive content make it the current market leader in the MALL app industry. The latest statistics released by Duolingo revealed its rapid development with a promising future: the total revenues achieved 531.1 million USD in 2023 with 44% increase from the prior year (Duolingo, 2024). It is also expected to be one of the next billion-dollar startups (Forbes, 2019). Meanwhile, the actual effectiveness of Duolingo and other commercial language learning apps has long been questioned.

1.2. Research Background and Objectives

Can mobile apps really be a reliable assistance or even a substitute for traditional language classes? To what extent can they enhance learning outcomes? Numerous commercial materials and research findings supported their effectiveness. For example, Duolingo claims that 'you can learn as much as 5 semesters of university instruction in just 5 sections of Duolingo'[3]. The claims were somewhat supported with empirical evidence: the average progress made by Spanish learners is 8.1 WebCAPE[4] points per hour of study on Duolingo (Vesselinov & Grego, 2012). According to Kittredge et al. (2024), after 4-6 weeks of engagement with the Duolingo app, users can achieve A1 proficiency level according to the Common European Framework of Reference for Languages (CEFR), enabling them to start basic conversations. It is also found to be even more effective than traditional English as a Foreign Language (EFL) classes in boosting various English proficiency (Rodríguez-Fuentes & Swatek, 2023). Exciting as these reports may seem, they are mostly conducted by the research team of Duolingo, or funded by the company and used as advertising content on its website. Less positive or significant outcomes have been reported by some independent researchers, noting the effectiveness of Duolingo may be exaggerated by its commercial materials (Loewen, 2019). Similar situations have also been observed in the research of other language learning apps, which will be further discussed in the literature review chapter.

Furthermore, advertisements of mobile apps also emphasize their motivational

[2] Duolingo Language Courses: https://www.duolingo.com/courses/all

[3] Duolingo Works: https://www.duolingo.com/efficacy

[4] Web based Computer Adaptive Placement Exam: https://emmersion.ai/products/webcape/

features that make language learning more enjoyable and easier to persist. For example, Duolingo claims, “After just 4 weeks of Duolingo, 8 out of 10 new learners are more motivated to learn a language” (Duolingo, see Footnote 3), and Babble states “99% of the participants declared they would continue to use Babble”[5]. However, there is a paucity of studies that systematically investigated users’ motivation with MALL platforms under established theoretical frameworks. Even among the small number of empirical research with rigorous design and solid theoretical foundation, few target the learning process of non-English languages, thus leaving an important gap to be filled. Therefore, this thesis aims to investigate the impacts of MALL on Chinese ESL learners’ achievement and motivation in learning French on Duolingo.

1.3. Research Questions and Outline

This thesis is guided by three research questions and adopts a mixed-method approach. The first two questions are addressed through quantitative methods, while the third one is explored using a qualitative approach.

1. To what extent does the use of Duolingo contribute to Chinese ESL students’ French proficiency?

2. In what ways does Duolingo influence students’ motivation in learning French?

3. How do students perceive their experiences of MALL with Duolingo?

This chapter has briefly introduced the research background and the platform selected to carry out this experiment. In the next chapter, I will analyse the existing studies regarding MALL, Duolingo and motivation in language learning with purposefully selected studies. Then, the methodology chapter will explain the research design, the sampling approach, the Duolingo platform, data collection and analysis methods and ethical considerations in great detail. Results will be presented with learners’ statistics, relations of the variables, and interview data from the experiment, addressing the research questions. Explanations of the results and potential limitations will be discussed in the following chapter. In the end, a brief conclusion will summarize the background, study design, findings and implications of this research.

[5] Research at Babble: https://uk.babbel.com/why-babbel-works

Chapter 2. Literature Review

The objective of this thesis is to investigate Chinese ESL students' achievement and motivation in the MALL experience with Duolingo. A review of previous literature will be provided in the following order: m-learning and MALL, MALL research in language proficiency, MALL research in learners' motivation, and research gaps.

2.1. M-learning and MALL

2.1.1. Concepts of M-learning and MALL

Researchers have attempted to define m-learning from different perspectives. Quinn (2001, p. 1) emphasized the tools used to carry out learning activities, defining m-learning as "elearning through mobile computational devices: Palms, Windows CE machines, even your digital cellphone". O'Malley et al., (2005) defined m-learning from the learners' perspective as a kind of learning that happens when the learner is not fixed at a predetermined location, or when the learner benefits from the learning opportunities offered by mobile technologies. The concept of "mobile" could signify "the learning mediated by mobile devices", "the mobility of learners", or the mobility of learning content that can be accessed anywhere (Taylor, 2006, p. 24). Therefore, the definition of m-learning may vary in individual studies to serve different research aims, and its concept has also evolved due to technological development and the emergence of smaller, more portable devices.

In the early stage of m-learning studies, researchers included a range of portable mediums with the potential to serve learning purposes, such as books, audio cassettes, portable radios and audio CDs (Kukulska-Hulme & Shield, 2008). However, with the advancements in ICT, m-learning studies now focus on using smartphones, tablet computers, e-books and other such wireless Internet-connected devices which can be used to access learning materials. Therefore, the concept of m-learning in this thesis aligns with the more recent definition released by UNESCO in 2013:

> Mobile learning involves the use of mobile technology, either alone or in combination with other ICT, to enable learning anytime and anywhere. Learning can unfold in a variety of ways: people can use mobile devices to access educational resources, connect with others, or create content, both inside and outside classrooms. (West & Vosloo, 2013, p.6)

In this light, the concept of MALL is also formulated, referring to the use of mobile

technology and probably with other ICT in language learning, which enables learners to access learning materials, connect with others or create contents anytime, anywhere. Although MALL is usually regarded as a specialization or subcategory of m-learning that specifically looks into the language learning process (Duman et al., 2014; Shortt et al., 2023), a distinction is specified here. M-learning also encompasses the general technological support for broad educational needs, such as management platforms, administration of school systems, and communications between schools, families and students (West & Vosloo, 2013), while MALL would primarily focus on linguistic skills and language learning experiences. Therefore, the general m-learning platforms which only have administrative or communicative functions, such as learning management systems (e.g., Canvas, MS Teams) would not be considered in this thesis, and the MALL platforms involved will be specified in the following section.

2.1.2. MALL Platforms

Platforms employed by prior empirical MALL studies can be briefly classified into two types, one featured in built-in functions of mobile devices like photo-taking (e.g., Liu & Chen, 2015), audio/video recordings (e.g., Shakarami et al., 2014) and text messages (eg., Alemi et al., 2012), while the other included commercial products like language learning apps, websites and social media. Although both types take advantage of the flexibility and convenience offered by mobile devices, they actually share little in common in terms of learning content and pedagogy. The built-in functions of smartphones or tablets do not create systematic instructional materials by themselves, nor do they provide resources for professional training or practising. Consequently, the influences of built-in mobile features are not language-specified, and are easily transferable to other areas of mobile learning. Therefore, to keep this thesis concise and focused on the language learning process, it would primarily concentrate on MALL with commercial platforms which could potentially provide language-specific instruction or practising opportunities.

A variety of commercial language learning apps have been analyzed to figure out the trends and limitations in the development of the MALL industry. Heil et al. (2016) reviewed the top 50 ranked language learning apps on Google Play and Apple Store. Their results revealed a predominant focus on teaching vocabulary in isolation without contextualized usage, and little explanatory corrective feedback or personalized adaption was provided to individual learners. Based on their findings, only a small part of the top-ranked apps aim to offer a whole package of language learning experiences that could possibly be an alternative to human teachers, while most apps focus on a single language or a specific

language skill, mostly vocabulary, and provide fragmented knowledge for practising. The different focus and characteristics of language learning apps pose a challenge to generalizing the overall effectiveness of MALL, because the findings in a certain app may not be transferable to another MALL platform with different designs or focus. Therefore, further classification of the apps is essential to draw more reliable and generalizable conclusions.

The commercial apps in previous empirical MALL literature can be further categorized according to their primary functions. Rosell-Aguilar (2017) classified apps for language learning into three types, which are illustrated in Figure 1. The first type includes apps that provide whole packages of language learning solutions. These apps have structured language courses with explanatory instructions to grammatical rules and a variety of exercises, or provide opportunities for interaction by matching users with other language learners, tutors or the community through certain algorithms. Duolingo, Babble, Busuu, HelloTalk and Tandem are several successful representatives of this kind. The second type encompasses apps not designed for language learning but useful for language learners. Apps of this kind may not have detailed explanations of grammar or systematic

exercises, but they can potentially improve language proficiency through implicit learning, information input and interaction with others. For example, Memrise and Quizlet target vocabulary building with flashcards; Kindle and BBC News may improve reading comprehension; ChatGPT and Blogger promote writing proficiency; YouTube and TED enhance listening and speaking skills; and social media apps can help with overall competence by boosting communication, interaction and providing potential learning contents. Furthermore, the third type would be dictionaries and translation apps, such as Google Translate and Merriam-Webster.

With this classification in mind, this thesis concerns all three types of language learning apps to provide a comprehensive overview of MALL research, but the first type would be the primary focus as it relates more closely to Duolingo, the platform selected to carry out the experiment. The next section will explore the existing research on the effectiveness of MALL, assessing how well these apps achieve their educational goals and support language acquisition.

2.2. MALL Research in Language Proficiency

2.2.1. Overview

MALL has gained increased attention from researchers thanks to the advancements in technology and the surge in m-learning demands. According to Burston and Giannakou (2022), a total of 1144 experimental implementation studies in MALL were released between 1994 and 2019 across different languages and linguistic features, and 58 overviews or meta-analyses of MALL studies were published between 2006 and 2020. In general, previous literature found the influences of MALL implementations overwhelmingly positive in various aspects of language learning, including vocabulary (Lin & Lin, 2019; Mahdi, 2018), reading comprehension (Klimova & Zamborova, 2020; Gutiérrez-Colón et al., 2023), academic writing (Solikhah & Wirawati, 2020), speaking and listening (Kukulska-Hulme & Shield, 2007). Moreover, the positive effects of empirical studies were supported by impressive effect sizes in meta-analyses. Peng et al. (2021) analyzed 17 MALL literature from 2008 to 2017 with a focus on comparison between MALL treatment and traditional instruction, revealing a large effect size for overall L2 proficiency (d = .95), writing (d = 1.33), listening (d = .99), and a medium effect size for reading (d = .52) and speaking (d = .46).

The meta-analyses conducted by Chen et al. (2020) and Burston and Giannakou (2022) are also valuable as they included a relatively large number of empirical studies across a considerable time span. Chen et al. (2020) reported a medium-to-large effect size

(g = 0.722) for the effectiveness of using mobile tools in language learning achievement based on their analysis of 80 empirical studies between 2008-2018. They also found MALL is more effective than traditional teaching methods. Similarly, a large effect size for both experimental (g = 0.72) and quasi-experimental (g = 1.16) studies was reported by Burston and Giannakou's (2022) meta-analysis of 84 empirical research between 1994 and 2019. However, the problem of publication bias was discussed in both meta-analyses, which could possibly temper the positive results.

These two meta-analyses contributed insightful results and compensated for each other's weaknesses. Burston and Giannakou's (2022) study adopted stricter inclusion criteria only incorporating studies with a reasonable sample size (N > 15) and a minimum treatment duration of 8 weeks, while these criteria were not specified in Chen et al.'s (2020) work. However, the lack of differentiation between smart and non-smart devices, as well as between language-specific platforms and general-purpose apps, may undermine their findings by failing to account for the variations in functionality of mobile tools across the 25-year time span. The results of Chen et al. help to clarify this confusion: the effect size of integrating non-smart devices (g = 0.787) was marginally larger than using smart devices (g = 0.763), and the effect size for educational apps (g = 0.759) was greater than that for general-purpose apps (g = 0.667). After all, both meta-analyses supported the positive effects of MALL on learning achievement with convincing statistics, and valuable pedagogical implications were thus suggested for future language education.

Although limitations of MALL have also been noted, such as small screens, added costs, distractions from fancy designs, reliance on network and charging, as well as the addictive nature of mobile devices (AbuSa'aleek, 2014; Kacetl & Klímová, 2019; Zain and Bowles, 2021), the use of mobile technology and portable tools in education seems an inevitable trend as they are already ubiquitous in daily life with powerful functions and great flexibility. The primary concern of MALL lies in how to use them efficiently, what platforms to use and in what ways could they enhance the learning process and academic gains (Kacetl & Klímová, 2019).

2.2.2. Empirical Research in Apps Designed for Language Learning

Most prior MALL literature integrated mobile devices as an assisting role in classroom learning or formal tuition, while less attention has been paid to the platforms that aim to provide a whole language learning solution (Huang, 2020). In other words, the first type of apps designed for language learning in Roselll-Aguilar's (2017) taxonomy have been under-researched. It is not hard to find commercial materials with very impressive

effectiveness evidenced by research conducted or funded by the app companies and notably displayed on their websites. For example, Babble stated “15 hours of study in a two-month period to cover the requirements for one college semester of Spanish” (Babble, see Footnote 5). Busuu claims “just 22 hours on Busuu was equivalent to studying a semester at college”[6]. Mango Languages reported “15 hours average study time using Mango to cover the requirements of one college semester”[7]. Duolingo also shows “You can learn as much as 5 semesters of university instruction in just 5 sections of Duolingo” (Duolingo, see Footnote 3).

Inspiring as these claims may seem, their effectiveness is quite vaguely described using “college semesters” to quantify learners’ achievement. However, college semester is not a universally accepted measurement and the language proficiency represented by “one college semester” is not yet clear. A closer investigation into the source of these claims found a series of reports released by Vesselinov, Grego and their colleagues over the past fifteen years. They evaluated 10 commercial language learning apps: Rosetta Stone (2009a, 2019), Duolingo (2012), Language Zen (2015), Busuu (2016a, 2021), Babble (2016b), Hello English (2017), Italki (2018), Pimsleur (2019a), Mango Languages (2019b) and LingQ (2023). They utilized very similar methodologies incorporating a pre-intervention test and a post-intervention test. The target language was Spanish in all these efficacy studies except for Hello English which obviously targets English learning. The study of Hello English adopted a controlled experiment which randomly allocated 97 Indian students from grades 8-12 to either the experimental group using the app or the control group not using the app, and the experimental group significantly outperformed the control one after 3-4 months of intervention (Vesselinov & Grego, 2017). Their studies of the other nine apps did not involve a control group but recruited participants from the app users by email, and the treatment duration spanned 2 months. They used WebCAPE (see Footnote 4) to evaluate learners’ reading, listening and grammar proficiency, some of the results for different apps are listed below:

- Duolingo: Participants gained an average of 8.1 WebCAPE points per hour of study on Duolingo, and 34 hours of learning can cover the requirements for the first college semester of Spanish (*N* = 88; Vesselinov & Grego, 2012).
- Language Zen: The average increase is 11 WebCAPE points per hour, and 25 hours

[6] Busuu Research: https://www.busuu.com/en/research

[7] Mango Languages: https://mangolanguages.com/discover-the-impact-of-mango-languages-efficacy-study-results/

can cover one college semester (*N* = 101, Vesselinov & Grego, 2015).

- Babble: The average increase is 12.7 WebCAPE points per hour, and 21 hours can cover one college semester (*N* = 325, Vesselinov & Grego, 2016b).
- Italki: The average increase is 14 WebCAPE points per hour, and 19 hours can cover one college semester (*N* = 102; Vesselinov & Grego, 2018).
- Mango Languages: The average increase is 18 WebCAPE points per hour, and 15 hours can cover one college semester (*N* = 95; Vesselinov & Grego, 2019b).
- Rosetta Stone: The average increase is 21 WebCAPE points per hour, and 13 hours can cover a college semester (*N* = 143; Vesselinov et al., 2019).
- Busuu: The average increase is 5.8 WebCAPE points per hour, and 13 hours can cover a college semester (*N* = 114; Vesselinov et al., 2021).
- LingQ: The average increase is 23.3 WebCAPE points per hour, and 12.6 hours can cover a college semester (*N* = 101; Vesselinov et al., 2023).

The series of efficacy studies using similar methodologies and analytical approaches provided solid evidence to compare different apps. Their random sampling enabled the studies to include a relatively large number of participants from various age groups (18 and over), education levels and professional backgrounds. And they quantified learning outcomes using “points per hour”, which is reasonable and clear. However, there are still several problems we should highlight. The efficacy of Rosetta Stone ranked second among all these platforms, demonstrating an hourly increase of 21 points, while Busuu, with the least hourly progress of only 5.8 points, covers one college semester using the same study time (13 hours for both apps). The results for Busuu in 2021 also contradicted Vesselinov and Grego’s earlier findings in 2016 which reported an hourly increase of 12 WebCAPE points and a total of 22.5 hours to cover a semester (N=196). Meanwhile, a closer look into the testing instrument found the WebCAPE was developed by the expert testing division of Rosetta Stone (see Footnote 4), so it could be much fairer and more convincing if a different testing instrument is adopted that does not have interest conflicts with any of the apps. Furthermore, although the authors stated the data collection and analysis were independently conducted by the research team, all these studies were funded by the respective app companies, and most of them are not peer-reviewed or published in prestigious journals. Therefore, further empirical evidence is needed to validate their reliability.

Results from other experimental research with different testing instruments and target languages are also insightful. Lord (2015) found no significant difference in either oral or written proficiency between three groups of participants learning Spanish through traditional classes, with Rosetta Stone, or a combination of both. Their proficiency was assessed through the Spanish College Level Examination Program and the Versant Automated Oral Proficiency Test before and after the treatment of 16 weeks. However, the very small sample size (N = 12) undermined its findings, and the exact time spent on the app was not reported, which posed a challenge to verify the claim that 13 study hours on Rosetta Stone can cover the first college semester of Spanish learning made by Vesselinov et al. in 2019. However, what we can conclude from this experiment is that the Rosetta Stone group spent less time than the classroom group (45 hours) during the 16-week college semester, but achieved similar learning outcomes.

Furthermore, Loewen et al. (2019) investigated the outcomes of nine adults learning Turkish exclusively on Duolingo. The participants were all experienced language learners with mixed L1s but had no prior exposure to Turkish. They took the Duolingo Progress Test and the Turkish 151 Test after the 12-week intervention and wrote journals during the period. The results revealed that although there was an improvement in Turkish proficiency, the actual effectiveness of Duolingo might be overstated by those commercial materials and Duolingo-funded research. Their finding “even after 34 hours of study, only one participant received a score that would be considered a passing grade in the university’s first semester Turkish course” explicitly contracted Vesselinov and Grego’s (2012) claim that 34 hours of learning on Duolingo could cover the material for the first college semester. However, the small sample involving only nine participants is likely to question the reliability of Loewen et al.’s research.

Kessler et al. (2023) somewhat supported Loewen et al.’s finding in their research on Babble and Duolingo involving 59 adults with mixed L1s. The participants learned Turkish from scratch for 8 weeks and made progress in the posttest designed by the research team. However, they averaged only 43.8% in the posttest with 15.4 hours learning on Babble, and 38.8% with 17.6 hours on Duolingo. The results of Duolingo were largely in line with Loewen et al.’s (2019) finding, whilst the achievements of both groups seem to challenge the findings of Vesselinov’s team, arguing the actual effectiveness has been exaggerated. However, it is also reasonable to note, that both Loewen et al. and Kessler et al. focused on novice learners with no prior Turkish proficiency while Vesselinov’s team involved experienced Spanish learners, and the difference between Turkish and Spanish may also

account for part of the reasons to their different findings.

On the other hand, Jiang et al. (2020) advocated for Dolingo's effectiveness focusing on reading and listening skills. The research involved 225 adult US learners with no prior proficiency in the target language to learn either Spanish (n = 135) or French (n = 90) exclusively from Duolingo. According to the American Council on the Teaching of Foreign Languages (ACTFL) levels for listening and reading proficiency, Spanish learners who completed beginning-level courses on Duolingo reached intermediate low in reading and approached novice high in listening; French learners approached intermediate low in reading and novice high in listening. Their scores are comparable to the proficiency of students in the fourth semester of university studies. This finding largely agrees with Vesselinov and Grego's (2012) report. However, the time spent to complete the beginning-level materials on Duolingo may take hundreds of hours (an average of 141 hours for Spanish and 20 hours less for French) which spans over a few years, the ignorance of possible exposure to other information input in the target language may confound the results. Moreover, the leading author of this study is an internal researcher of Duolingo, which further questions the actual trustworthiness with this conflict of interest.

With the abundance of reports demonstrating the effects of mobile apps are comparable to, or even better than that of traditional instruction, studies with negative results are quite scarce, especially in the recent decade. Publication bias might account for part of the reasons. Usually, studies with statistically significant results and positive outcomes are more likely to be published (Chen et al., 2020; Burston, 2015). Furthermore, many of the studies that reported very positive results from MALL treatment integrated the mobile apps as an additional tool in comparison to classroom-only instruction (e.g., Alfuhaid, 2021; Vesselinov &Grego, 2017). In other words, the experimental group received the same classroom instruction as the control group, but also took advantage of the MALL instruction during the experiment. In this way, the experimental group received more educational input and devoted more time to learning activities, which renders the comparison unfair and problematic. It could be obviously reasonable that if another additional instruction or practice is integrated, such as reading books or writing essays in the language, the students will still outperform their classroom-only counterparts due to increased input and effort. Therefore, experimental designs should be more cautious with the control of conditions, and keep the comparisons fair.

Therefore, the few studies with fair comparison between experimental and control groups, and reported negative results of MALL intervention should be exceptionally valued.

Ratzlaff (2015), in their master thesis, found the participants taking the entry-level university courses (n = 10) significantly outperformed those learning with Duolingo (n = 5) after the controlled experiment spanning one college semester. The testing instrument was rigorously developed by the researcher after consulting extensive resources and includes a total of nine sections evaluating different aspects of language skills and knowledge. Although the Duolingo group spent less time (average 29.6 hours) than the classroom group (54 class hours and 27 after-class hours), it seems not possible to support the claim that 34 hours of Duolingo study can cover one college semester (Vesselinov & Grego, 2012).

Although substantial literature exists analyzing the effectiveness of language learning apps, there has been a significant gap between the results reported by app-related research teams and those obtained by independent researchers. What we can conclude so far is that mobile apps can teach foreign languages to some extent, but their actual efficacy still needs further investigation. In addition to learning outcomes, the motivational effects of language learning apps also remain another critical dimension of MALL research, which is discussed in the following section.

2.3. MALL Research and Learners' Motivation

2.3.1. Introduction to Motivation Theories and Development

Motivation means the "direction and magnitude" of people's actions, involving why they choose to do something, how long they would keep doing it and how much effort they put into it (Dörnyei & Ushioda, 2021, p4). Motivation research relating to second language acquisition (SLA) emerged in the 1950s and has been generating insightful studies. Dörnyei (2005) divided the development of L2 motivation research into three distinct phases: the Social Psychological Period (1959-1990), the Cognitive-Situated Period (the 1990s) and the Process-oriented period (the 2000s onwards).

Social Psychological Period (1959-1990). The first research in motivational factors and language achievement was done by Gardner and Lambert in 1959. They measured the language aptitude, verbal intelligence and various attitudinal and motivational characteristics of 75 high school students learning French as L2. The results showed two independent factors, linguistic aptitude and motivation, were equally related to learners' achievement in French. This study attracted scholars' attention to motivation research in SLA and generated numerous replicating studies and similar investigations across different countries and languages (e.g., Gardner & Lambert, 1972). Based on their work, it is recognized that motivation and attitudes are independent of language aptitude, and each of

these factors uniquely influences the outcomes in language learning (Gardner, 1988).

The Socio-educational Model was also proposed and developed during this period by Gardener and his associates. It aims to account for the influences of individual differences, including intelligence, language aptitude, anxiety and motivation, on L2 proficiency (Gardner, 1988). They also formulated the Attitude Motivation Test Battery (AMTB) to investigate a range of variables, such as integrativeness, attitudes towards the learning situation, language anxiety and motivation (Gardner, 2019). The Socio-educational Model laid the foundation for the subsequent motivation theories in language acquisition, and the format of AMTB has been frequently adopted in empirical studies.

Cognitive-Situated Period (the 1990s). This period is featured by examining motivation in more situated social and personal contexts, drawing upon cognitive theories in educational psychology. A range of motivation theories were applied to guide the research during this time. For example, the SDT (Deci & Ryan 1985) explains the role of motivation from intrinsic and extrinsic perspectives and highlights the importance of basic psychological needs in personal growth. Attribution Theory (Weiner, 1985) suggests the ways in which people attribute the causes to their success or failure in turn affect their motivation and future behaviour. Moreover, the Three-level Framework (Dörnyei, 1994), drew on components from the socio-educational model and the SDT to offer a structured approach to understanding the factors that influence motivation from the language level (integrative and instrumental motivation), the learner level (desire to achievement and self-confidence) and the learning situation level (course, teacher and group variables).

Process-Oriented Period (the 2000s onwards). Motivation research in this stage is characterized by Dörnyei, Ushioda and their colleagues, emphasizing an interest in the dynamic nature and temporal variation of motivation, as well as its relationship to learners' behaviour and classroom performance (Dörnyei, 2005). Against the evolving sociolinguistic landscape of globalization in the late 1990s and early 2000s, language use has become increasingly connected with culture, identity, ethnicity and other sociocultural factors. Therefore, a need for reconceptualizing and retheorizing L2 motivation research has been identified (Ushioda & Dörnyei, 2009). Notably, the L2 Motivational Self System (L2MSS) emerged as a crucial framework in response to this shift during this period. It is shaped by personal and social ideologies, suggesting the discrepancy between one's current state and future image can motivate learners to work towards the desired destination (Dörnyei, 2009). The L2 Motivational Self System (L2MSS) has been one of the most prominent frameworks for investigating learners' motivation in SLA (Al-Hoorie, 2018).

Among the various great frameworks across the three periods in motivation research, SDT was selected to design the motivation survey of this study because it highlights learners' autonomy, which is essential in self-paced learning with mobile apps. Moreover, SDT involves both intrinsic and extrinsic dimensions that meet the need to explore the enjoyable and social features of Duolingo. The next section will provide a more detailed introduction to this framework.

2.3.2. The Self-Determination Theory

The key concept of SDT lies in the difference between amotivation and motivation, as well as the distinction between autonomous motivation and controlled motivation (Gagné & Deci, 2005). Figure 2 illustrates the relationships and differences between the components. Amotivation means an absence of motivation or a lack of intention to do something. People may feel amotivated because they do not recognize the value of performing an action, feel unable to do it or do not expect any outcomes from the action (Ryan & Deci, 2000). In contrast to amotivation, both autonomous and controlled motivation are intentional. Intrinsic motivation is driven by people's inner enjoyment and good feelings in the action itself, which is undoubtedly autonomous, while the types of extrinsic motivation are somehow driven by external factors, and their autonomy can vary in the degree to which they are autonomous or controlled (Gagné & Deci, 2005). In summary, SDT provides a continuum that ranges from amotivation (no self-determination at all) to intrinsic motivation (totally self-determined), and in the middle ground along the scale are the four types of extrinsic motivation.

Looking at the components of the SDT continuum, each type of motivation reflects the degree to which the values and regulations of a certain behaviour are internalized (Ryan & Deci, 2000). Behaviour with external regulation shows the least autonomous motivation, merely to meet external requirements or avoid punishments. Introjected regulation motivates people to perform an action in order to keep their self-esteem or to be respected by others, and possibly to avoid embarrassment, anxiety, guilt or other negative emotions in social contact, which is somewhat external and still controlling. Identified regulation is more autonomous and self-determined, reflecting that a person recognizes the values of the behaviour. If identified regulation is fully internalized into people's broader self-concept and overall ideology, it would become integrated regulation which represents the highest level of internalization of extrinsic motivation. However, even the the most internalized form of extrinsic regulation still differs from intrinsic motivation, because it aims at the instrumental value of an action and expects positive outcomes from the behaviour, whereas

intrinsic regulation is inherently driven by the enjoyment of the action itself and is fully autonomous (Gagné & Deci, 2005).

SDT was originally proposed in research on human behaviour and work motivation, but because of its broad concepts and inclusive nature, it has been applied to a range of research areas, such as education (e.g., Niemiec & Ryan, 2009), sports (e.g., Pelletier et al., 2001), and psychology (eg., Deci & Ryan, 2008). This study specifically looks into its application to MALL research as the guideline for designing motivation surveys and investigating learners' experiences.

2.3.3. Application of SDT in MALL Research

Most language learning apps claim to boost learners' motivation, however, there's a paucity of studies that integrated SDT to systematically investigate this issue. Although some of the research resorted to Gardner's theory (e.g., Habibie, 2020) or L2MSS (e.g., Vesselinov & Grego, 2019a, 2021) in designing the surveys, most studies simply did not adopt any motivation theory at all. Below are some of the notable literature among the small number of MALL empirical studies with SDT that are related to this research.

Vesselinov (2009b) measured the attitude and motivation of 164 Rosetta Stone users randomly recruited through a newspaper advertisement. They found a significant increase in both intrinsic and extrinsic motivation in learning Spanish. The measuring

instrument was meticulously designed based on three sets of scales, the Academic Motivation Scale (AMS), the Intrinsic Motivation Inventory (IMI), and the adaption of Gardner's AMTB. Among them, both AMS and IMI were derived from SDT to measure learners' motivation with multi-dimensional subscales. The data were collected through 7-point Likert-Scale items, and the results indicated that after 1-month learning with Rosetta Stone, the participants maintained or increased all aspects of intrinsic and extrinsic motivation in learning a foreign language. The finding advocates for the motivational value of the mobile app, highlighting its role in enhancing learners' attitudes and learning experiences. Moreover, the comprehensive and rigorous instrument is also insightful to the design of the subsequent motivation studies.

Akman and Karahan (2023) investigated 110 students' perceptions toward MALL through questionnaires and written interview questions, suggesting an overall positive attitude towards using mobile apps for language learning. The participants were recruited through convenience sampling from the English Language Teaching Department in Turkey, and the survey and interview questions were rigorously designed focusing on multiple dimensions, including acceptance of m-learning tools, learners' autonomy, and components of SDT. Their findings highlighted the benefits of MALL in boosting learners' motivation and autonomy, providing substantial support to the integration of ICT tools into the process of language teaching and learning in this highly digital world.

Motivational benefits of MALL were also reported in Zeng and Fisher's (2024) 6-week research with 20 Chinese junior school students learning English through Duolingo. They hypothesized that the gamified elements of Duolingo could enhance intrinsic motivation, and the intrinsic motivation generated by Duolingo can facilitate their general intrinsic motivation in learning English, which they thereby defined as the "transfer of motivation". Similarly, it is also hypothesized that learners' autonomy and competence developed from Duolingo could also be "transferred" to general L2 autonomy and competence. According to questionnaire results and group interviews, all their hypotheses were confirmed, highlighting the potential of MALL in stimulating L2 motivation. However, the findings could be more substantial if a larger sample was recruited.

In terms of theoretical innovation related to SDT, Chen and Zhao (2022) proposed a new framework combining SDT and the Technology Acceptance Model, and examined this framework with 277 Chinese college students using 5 gamified English vocabulary-learning apps. They explored the relations between autonomous and controlled motivation, as well as their relations to learners' perceived usefulness and ease of use of the apps. Data were

obtained through a survey involving Likert-Scale items and self-reported frequency and duration of the app usage. Results indicated that people choose to use the apps mostly because of autonomous motivation rather than the controlled one, and autonomous motivation positively affects the perceived usefulness and ease of use while controlled motivation only affects the former. Innovative as it is in terms of the theoretical framework, the sole reliance on self-reported survey data without further investigation into the underlying causes may fail to provide robust data or in-depth findings, thus challenging the significance of the outcomes.

While the results of MALL on learners' motivation and perceptions seem predominantly positive, negative effects were also reported. Adara (2020), through a controlled experiment with 69 Indonesian students in either a traditional English class or a MALL-integrated class, found negative impacts of MALL on learners' autonomy and motivation. The experiment had 14 sessions, each lasting about 1.5 hours, and questionnaires were administered before and after the treatment. Semi-structured interviews were also conducted at last. However, their findings are quite questionable as the researcher failed to explain the experiment sample and instruments sufficiently. The age group of the participants were not mentioned, moreover, the MALL platform, the design of the questionnaire and the procedures of treatment were not illustrated in detail. Therefore, further research is still needed with rigorously designed experiments to explore the comparison between MALL and traditional approaches in learners' motivation.

These studies applied SDT to empirical MALL research in various contexts with different focuses, providing insightful results to enhance understanding. However, they shared common shortcomings for highly relying on self-reports and interviews for data collection, while few investigated learning outcomes with standard tests, school grades or more objective measures. Furthermore, existing literature only sheds light on learners' motivation in learning English as the L2, with little attention paid to other languages. Therefore, important gaps exist in MALL motivation research with more comprehensive and objective research designs and targeting non-English languages, especially in the light of SDT. The next section will detail the gaps that this thesis aims to address.

2.4. Research Gaps

2.4.1. Target Language and Language of Instruction

According to Burston and Arispe (2022), the dominant target language within MALL research is undoubtedly L2 English, representing nearly 90% of the experimental

implementations, while the second place is Chinese which only accounts for 3%, and all other languages together take up a marginal proportion. Although commonly taught European languages, such as Spanish, German, Italian and French gained some attention in Duolingo studies, their proportion was rather small. English is still the one that attracted most academic attention which takes up nearly 40% of the empirical research on Duolingo (Shortt et al., 2023). Therefore, the MALL learning processes and outcomes targeting non-English languages have been largely under-researched, so this thesis would focus on learning French as an additional language for Chinese ESL students.

Notably, in this experiment, the participants were learning an L3 through L2 English instruction rather than their native language, because English and French both belong to the Indo-European languages and share many common linguistic features, while Chinese belongs to the Sino-Tibetan languages which are quite distant from the target language. Typological similarity or overlap between languages positively influences L3 learnability (Hammarberg, 2001; Schepens et al., 2016), so the researcher chose the English-instructed French course as the research material, hoping to facilitate learning outcomes. As all the participants were studying in UK universities with above-intermediate English proficiency (Overall IELTS Score 6.0 and above), they were unlikely to experience problems with the English instructions on Duolingo. Little previous MALL literature took the language of instruction into consideration, but the inclusive portfolio of Duolingo makes it feasible by offering courses delivered in different languages. However, it is by no means possible to draw substantial conclusions with just this master thesis, but it may contribute to a new research area in MALL that deserves attention. Future research could also make comparisons between using different instruction languages to learn the same language on the same platform, which could possibly be valuable against the trend of globalisation and the increasing number of multilingual speakers.

2.4.2. Theoretical Framework

Burston and Giannakou (2022) pointed out that MALL research is largely atheoretical, calling for more theory-based practices to enhance awareness and produce insightful findings. According to Shortt et al.'s (2023) review of 35 Duolingo empirical studies, although more than 90% of them included one or more theories, over 60% are design-focused which investigates the design of different interventions or Duolingo features rather than how the app actually impacts on the learners or learning experiences. Moreover, as discussed in the previous section, there is a paucity of studies exploring the motivational effects of MALL in the light of SDT. To bridge this gap, the current research investigates

learner's motivation under the guidance of SDT, aiming to enhance understanding of learners' cognitive and psychological needs in the MALL experience from the perspectives of intrinsic and extrinsic motivation. It is also expected to address the gap in SDT's application to the acquisition of a non-English language.

2.4.3. Research Methodology

To avoid subjectivity which highly depends on personal experiences, and to enhance the depth and comprehensiveness of the research, a mixed-method approach integrating a pretest, a posttest, a motivation survey and optional interviews was designed. Both quantitative and qualitative data were analysed to explore the learners' achievement and motivation in the MALL process. Although tests, surveys and interviews are all common measures in plenty of MALL research, most prior experiments just adopted one or two of them separately without combined analysis. Therefore, some important information might be overlooked, thus the empirical studies may fail to reveal the whole story. This research aims to address this gap by providing a comprehensive understanding of the MALL experience with balanced objectivity and subjectivity.

Given the valuable contributions and imitations of current MALL literature discussed above, this study aims to address the gaps by focusing on the effects of using Duolingo to learn French, a less commonly researched language in this context, through the lens of SDT. The following chapter details the research design, participant selection, platform of instruction, methods of data collection and analysis, and ethical considerations in this study to achieve these objectives.

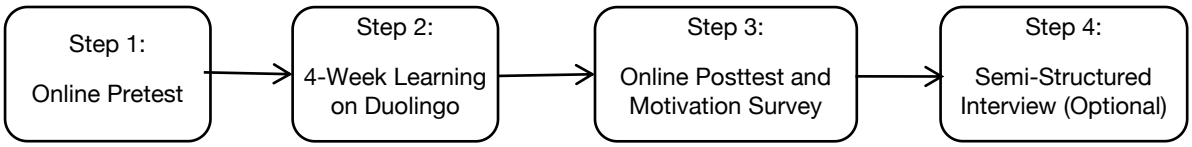
Chapter 3. Methodology

3.1. Research Procedure

As illustrated in Figure 3, this mixed-method research is conducted following these four steps: pretest - 4-week learning session - posttest and survey - optional semi-structured interview. Total time spent on Duolingo, pretest and posttest results, survey feedback and interview recordings were collected for data analysis.

Figure 3

Procedures of the Experiment



Guidelines and introductions to Duolingo settings and operations were sent to the participants through online communication before the experiment officially started, in order to make sure they downloaded the correct version and chose the right course to work on. The participants were also encouraged to follow each other on Duolingo to experience the social features. They are also free to ask the researcher any questions regarding the app and the experiment.

3.2 Participant Selection

3.2.1. Sampling

The participants were recruited based on voluntary participation through convenience sampling by personal contact, social media posts and WeChat group messages. The inclusion criteria are clearly listed:

1) University students at least 18 years of age
2) Native Chinese speakers and learned English as L2
3) Not advanced learners of French
4) Base in the United Kingdom or have access to the UK version of Duolingo

5) Willing to learn French using only Duolingo for 4 weeks and complete online tasks.

The restrictions of the geographical location or the versions of the app are imposed because Duolingo may have different designs and functions in other countries due to local policies. The participants have to download the correct app version before the beginning of the experiment, but they do not have to stay in the UK during the whole research period. They would not receive any financial incentive or compensation, but a one-month Duolingo subscription (worth GBP 9.99) was provided to unlock all the premium features and avoid wasting time on advertisements.

3.2.2. Sample Size and Attrition

The initial sample of this experiment involved a total of 21 Chinese ESL students aged 18 and above who are international students studying in UK universities. Among them, 7 are males (33%) and 14 are females (67%). Their English proficiency ranged from intermediate to advanced levels (IELTS Score 6.0 and above), and some had experiences in learning additional languages including Korean, German, Hungarian and Spanish. The participants have various majors, such as education, applied linguistics, engineering, social science and computer science. As Duolingo only offers French courses from A1 to B2, all the participants would be at or below B2 French proficiency according to the CEFR.

During the 4-week learning session, 3 of the participants (2 male and 1 female) dropped out from the experiment during the first week, so they were excluded from data analysis and subsequent qualitative research. One of them indicated the reason for losing interest in learning French, and one quit because of family issues, and the other did not provide any specific reason. All other 18 participants successfully completed the experiment, and all their data were valid for further analysis.

The final sample includes a total of 18 participants including 13 females and 5 males. Among them, 8 were at the A1 level (44%), 5 were at the A2 level (28%), 4 were at the B1 level (22%), and only 1 was at the B2 level (6%) according to the pretest results. The difference between learners with different proficiency levels will be explored in more detail in the results section.

3.3. The Learning Platform

The participants would learn through all the content provided by Duolingo. This app adopted a gamified approach and user-friendly operations to attract users. People can start learning with a few taps on the screen even without professional guidance. Just like playing

a game, they can collect Gems through completing daily quests (Figure 4) and achievements. They also get 'XP' (experience points) upon completion of each lesson, story, and challenge (Figure 5). Gems could unlock more challenges and activities, while weekly promotions or demotions to the higher or the lower league are determined by the learners' XP ranking among other users (Figure 6). Moreover, Duolingo also has a cute mascot, a green owl called Duo to interact with the users, which makes the learning experiences less boring (the green character in Figure 5).

The courses on Duolingo include various exercises that are logically structured. Each unit focuses on a particular topic (e.g. greetings, ordering food, telling time) and involves several key tips in vocabulary and grammar. Each unit has a guideline introducing the basic lexical, grammatical or phonological knowledge (Figure 7), and the lessons would focus on practising the knowledge in the forms of exercises, conversations and games. The exercises mainly include speaking, listening, gap-filling, matching and translation, emphasizing pronunciation, vocabulary, grammar and comprehension. Some of the examples are shown in Figure 8-10. The courses and exercises are progressively organized according to the complexity and frequency of language features, allowing learners to start from an appropriate level after taking a placement quiz.

Duolingo also allows for social interactions with the community and friends to enrich learning experiences. People can search for contacts, follow their friends and study together. Friends can send gifts to each other and complete ‘Friends Quests’ to win prizes (Figure 11). They can also send congratulations to friends upon their completion of achievements (Figure 12).

In this experiment, Super Duolingo subscriptions were offered to each of the participants to enable ad-free exploration of the personalized practices, and to ensure learning time. According to Duolingo, subscribed learners are 4.2 times more likely to finish their course (Duolingo advertisement on the app, Figure 13).

To minimize the interference caused by other learning materials, the participants were required not to learn French in any other ways during the experiment, including formal teaching, online classes, other language apps and self-study through textbooks. However, they were allowed to practice what they learned on Duolingo or speak French with friends in daily life.

3.4. Data Collection

Quantitative data of this research include learning time on Duolingo, test results and scalar ratings in the motivation survey, while qualitative data were collected from the interviews and the open question in the survey.

3.4.1. Learning Time

Participants were asked to report the learning time every week. They should send the researcher a screenshot of either the 'Screen Time' in the Apple setting or the 'Duolingo Weekly Report' sent by automatic email from the app, in order to prove the exact time they spent on Duolingo. All the participants were required to provide their weekly learning time on the Monday of the next week.

The researcher suggested that users spend at least an hour per week using this app, but it was not a mandatory requirement. Participants had the freedom to decide how much time they allocated to this learning activity. During the learning session, the researcher did not interfere much with the learning activities. The students had the autonomy to decide when to use the app and what features to explore, in order to make sure the researcher did not impose external control on their learning behaviour. This learning pattern is highly self-paced, which is expected to reveal learners' true motivation and autonomy. Their data would be valid as long as they reached the minimum requirement of learning time (60+ minutes during the 4 weeks). This aims to reveal the flexibility and convenience of MALL, allowing people to learn whenever and wherever they want, which reflects the typical usage pattern the Duolingo users in most cases.

3.4.2. Pretest and Posttest

The pretest and posttest were the same offered by the ESL website[8]. This test was selected because it is suitable for learners with elementary or intermediate French competence (from A1 to B2) and is simple to administer with no registration or subscription needed. It includes 40 multiple-choice questions targeting vocabulary, grammar and comprehension. Each question has four options, and only one of them is correct. The participants had to choose one of the options and then click the "Answer Question" button to proceed. The test-takers were required to only refer to their own knowledge in the tests, which means dictionaries, discussion with others and online searching were not allowed. It normally takes about 20 minutes to complete, but as it was conducted on open websites, the participants could do it at their own pace as long as they finished all of the questions. Upon completion of all the 40 questions, the results and proficiency level will be automatically shown on the screen. The participants were asked to send the screenshot with their score and proficiency level to the researcher.

3.4.3. Motivation Survey

[8] ESL: https://www.esl-languages.com/en/online-language-tests/french-test/start-test

The motivation survey includes a total of 13 questions. Among them, Questions 1-10 used five-point Likert-Scale items where 1 represents “strongly disagree” and 5 indicates “strongly agree” to investigate learners’ intrinsic and extrinsic motivation. Question 11 and Question 12 asked whether the participants would be willing to continue using Duolingo after the experiment, and the reasons why they would (or would not). The last question is optional and is totally open to participants where they could write down whatever they wanted to share related to the research, the experiment and the app. The survey is attached in Appendix 1.

The design of this survey was guided by Ryan and Deci’s (2000) SDT which divided learners’ motivation into Amotivation, Extrinsic Motivation and Intrinsic Motivation. As the survey targeted those who had already successfully completed a 4-week learning session, they were assumed to show some extent of motivation in language learning, so the “amotivation” state seemed inappropriate to be included here. Therefore, the motivation survey only focused on learners’ extrinsic and intrinsic motivation in language learning with 5 Likert-Scale items in each dimension. The survey was administered through Microsoft Forms, so the participants simply needed to finish it online.

3.4.4. Semi-Structured Interviews

After the post-test and the survey, 5 participants were purposively sampled to take part in a follow-up interview based on their gender, test performance, French proficiency, time spent on the app, survey feedback and willingness to contribute. The interviews were semi-structured and lasted about 30 minutes. The participants can use either English or Chinese to better express themselves, and the conversations were recorded for qualitative analysis. Related information of the interviewees is listed in Table 1.

The interviewees could choose to have the interview online through Teams or in person. The interview questions focused on learners’ language learning experiences both in traditional classes and with mobile apps, especially with Duolingo. They were encouraged to share opinions about the strengths and weaknesses of the app and to make comparisons between traditional learning and MALL. Some of the features and functions of Duolingo were mentioned to help them reflect on the learning experiences. The interview questions are presented in Appendix 2 for reference. As the interviews were semi-structured, the participants could also talk about any other personal thoughts about learning different languages and using other mobile apps, as long as they felt these were related to the research topic. The participants were informed in advance that all the conversations would be recorded, and some of their ideas would be quoted in the thesis.

Although the interviews were unlikely to include any sensitive or offensive questions, they were free to leave at any time if they felt uncomfortable to continue. Interview recordings would be transcribed and translated verbatim into English since most participants used Chinese for better communication.

Table 1

General Information of the Interviewees

Interviewee	Gender	Pretest Score	Pretest Level	Posttest Score	Posttest Level	Progress	Learning Time
A	M	19	B1	18	A2	-1	294
B	F	38	B2	38	B2	0	451
C	F	19	B1	20	B1	1	1244
D	F	11	A2	27	B1	16	150
E	F	8	A1	12	A2	4	329

Note. Learning Time is measured in minutes

3.5. Data Analysis

The software used for analyzing quantitative data is IBM SPSS Statistics 29.0. Quantitative data include pretest scores, posttest scores, total time spent on Duolingo (measured in minutes), learners' progress (progress = posttest score - pretest score), and scalar ratings in extrinsic and intrinsic motivation. As the experiment included a small sample size (N = 18), the Shapiro-Wilk Test was adopted to investigate the normality of the variables, and the results with skewness and kurtosis statistics are provided in Table 2. Among the six variables, pretest score (W = .754, p < .001) and learning time (W = .832, p = .005) are not normally distributed, therefore, non-parametric tests were performed to analyse the difference between pretest and posttest scores, as well as the correlation between learning time and progress. The correlations between motivation variables and learners' progress were analyzed through the Pearson Test as they are all normally distributed. The Histograms and Q-Q plots of the variable distribution are attached in Appendix 3 for reference.

Table 2

Normality Check of Quantitative Variables

Variables	Shapiro-Wilk		Skewness		Kurtosis	
	Statistic	Sig. *p*	Statistic	Std. Error	Statistic	Std. Error
Pretest Score	.754	<.001	1.969	.536	4.521	1.038
Posttest Score	.901	.061	1.412	.536	2.949	1.038
Progress	.918	.121	1.060	.536	1.946	1.038
Learning Time	.832	.005	1.436	.536	1.612	1.038
Extrinsic Motivation	.981	.963	.58029	-.109	.536	-.327
Intrinsic Motivation	.925	.161	-.767	.536	.567	1.038

Note. Sig. = Significance*;* Std. = Standard

Braun and Clark's (2006) 6-step approach for thematic analysis was applied to exploring the qualitative data. Firstly, the Chinese interview recordings were accurately transcribed and translated verbatim into English, and all the transcripts were read through several times to have an overview of the interviewees' feedback. Subsequently, the data were systematically coded, and the codes were then categorized into several themes with related quotes from the transcripts organized under each theme. After carefully reviewing and defining the themes, the data were woven together for interpretation and report. The key findings will be presented in the results chapter.

3.6. Ethical Considerations

The experiment was approved by the Central University Research Ethics Committee (CUREC) of the University of Oxford with the reference number C1A24087 (Appendix 4). The students were informed of the research procedures, data collection and data protection methods in advance through the Participant Information Sheet (Appendix 5) and Written Consent Form (Appendix 6). Before the research officially started, all the participants signed the Written Consent Form to explicitly confirm their willingness and awareness to take part in this experiment. Although this research was unlikely to cause risks to the participants or include any sensitive data, they were free to quit at any time if they did not want to continue. All their identifiable data were removed or de-identified with participant numbers, and the quotes from interviews would not include directly identifiable information.

With the ethical considerations meticulously addressed and all necessary consents obtained, the study proceeded smoothly, ensuring participants' privacy and data integrity. The following chapter presents the results derived from the data collected, offering an in-depth investigation into the participants' learning outcomes, motivation and experiences.

Chapter 4. Results

To ensure coherence and ease of understanding, the results of each research question are reported separately in sequence.

4.1. Learners' French Proficiency

4.1.1. Overview

Table 3 shows the participants' statistics in this experiment. The full mark of both the pretest and post-test is 40, with each multiple-choice question representing one point, and a learner's progress is calculated by subtracting the pretest score from the posttest score. The total pretest score is 234, and the total posttest score is 298. The mean of progress is 3.56, indicating that, on average, each participant increased their posttest score by 3.56 points. Additionally, 77.8% of the participants improved their scores in the posttest. Two students got the same scores in both the pretest and the posttest, while two students even experienced a decrease in their scores. The most significant improvement observed was an increase of 16 points over the four weeks, while the largest decline was a decrease of 4 points.

In terms of learning time, the most dedicated participant spent a total of 1380 minutes learning French on Duolingo, averaging nearly 50 minutes per day, and three students engaged with the app for more than 1000 minutes during the experiment. Most learners devoted 200-600 minutes to using the app, while the minimum learning time among the participants was 71 minutes.

Table 3

Overview of Learners' Statistics

Participant Number	Pretest Score	posttest Score	Progress	Learning Time
1	20	21	1	441
2	19	18	-1	294
3	38	38	0	451
4	7	8	1	563
5	19	20	1	1244
6	7	10	3	475

Participant Number	Pretest Score	posttest Score	Progress	Learning Time
7	7	15	8	378
8	11	27	16	150
9	8	12	4	329
10	10	13	3	305
11	10	6	-4	244
12	8	15	7	472
13	14	17	3	1016
14	22	23	1	1380
15	7	12	5	535
16	7	15	8	568
17	12	12	0	435
18	8	16	8	71
Sum	234	298	64	9351
Mean	13	16.56	3.56	519.50
Min.	7	6	-4	71
Max.	38	38	16	1380

Note. Min. = Minimum; Max. =Maximum

4.1.2. Difference Between Pretest and Posttest Scores

With a group of non-normal data, the Wilcoxon Signed Rank Test is adopted to investigate the difference between the pretest and posttest scores, and the results are shown in Table 4. The posttest scores (*Mdn* = 15) were significantly higher than the pretest scores (*Mdn* = 10) after 4-week learning on Duolingo ($Z = 2.884$, $p = .004$, $r = .48$). The effect size is calculated: $r = \frac{Z}{\sqrt{N}} = 0.48$. According to Cohen (1988), the effect size $r = .1$, $r = .3$ and $r = .5$ respectively indicates a small, medium and large effect. So the $r = .48$ here shows a medium-to-large effect, which means the observed value is not only statistically significant but also has a notable practical impact.

Table 4

Results of Related-Samples Wilcoxon Signed Rank Test

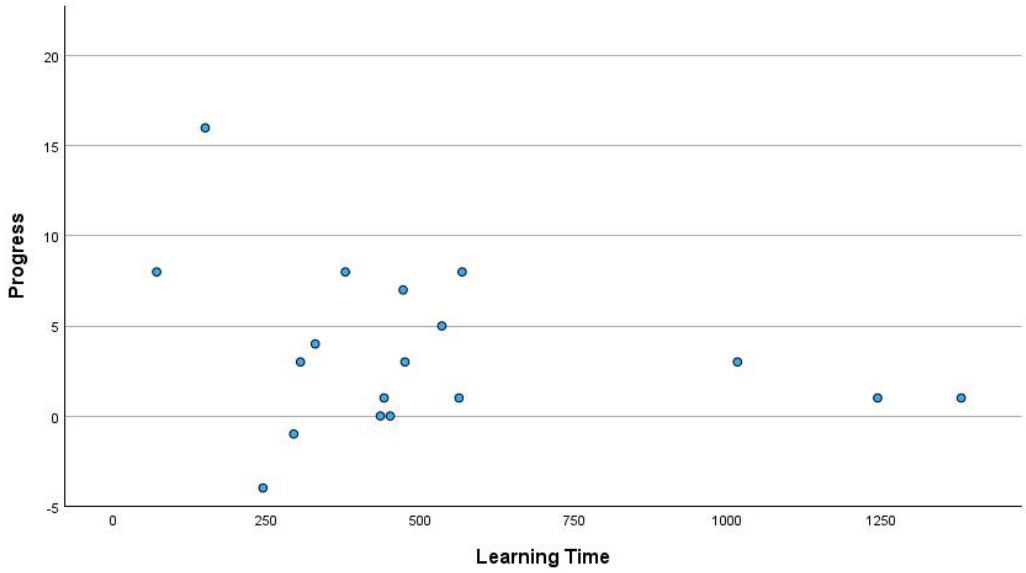
Total *N*	18
Test Statistic	123.500
Standard Error	19.245
Standardized Test Statistic (*Z*)	2.884
Asymptotic Significance (2-sided test, *p*)	.004

4.1.3. Correlation Between Learning Time and Progress

Following the significant difference identified between pretest and posttest scores, it is wondered if there exists a correlation between the amount of time spent on Duolingo and the progress made by the learners in French. Therefore, a scatter plot was generated (Figure 14), which did not show any linearity between the two variables.

Figure 14

Simple Scatter of Progress by Learning Time



Note. Learning Time is measured in minutes

A closer investigation was done through Spearman's rho because only the progress data are normally distributed (p = .121), while the learning time variable is not normal (p = .005). The results are shown in Table 5, confirming the observation from the scatter plot. The Spearman's correlation coefficient is -.093, indicating a very weak and negative

association. The time people spent on Duolingo was not significantly correlated with their progress in the French test (p = .714). Bootstrap analysis supported this finding, with the bias-corrected and accelerated (*BCa*) 95% confidence interval (*CI*) ranging from -.599 to .538, encompassing zero and further suggesting no reliable association between the two variables. This finding suggested that the participants did not necessarily get better grades as they spent more time learning French on Duolingo, which might contradict people's common sense. The reasons and results from prior literature will be explained in the discussion chapter.

Table 5

Correlation Between Learning Time and Progress

			Learning Time	Progress
Learning Time	Correlation Coefficient		1.000	-.093
	Sig. (2-tailed, *p*)		.	.714
	N		18	18
	Bootstrap	Bias	.000	.016
		Std. Error	.000	.287
		BCa 95% *CI*	.	[-.599, .538]
Progress	Correlation Coefficient		-.093	1.000
	Sig. (2-tailed, *p*)		.714	.
	N		18	18
	Bootstrap	Bias	.016	.000
		Std. Error	.287	.000
		BCa 95% *CI*	[-.599, .538]	.

Note. Bootstrap results are based on 1000 Bootstrap samples.

4.1.4. Comparison Between the Novice and Intermediate Learners

After carefully reviewing the progress data from the experiment, the researcher noticed that those who increased their scores by more than 5 points were mostly novice learners with little French proficiency before the experiment, while those who made little or no progress mostly achieved A2 and above proficiency levels in the pretest. Therefore, it is hypothesized that novice learners are likely to make more progress than intermediate

learners through Duolingo instruction. A further investigation is performed by dividing all the participants into either the Novice Group with A1 French level in the pretest (n = 8), or the Intermediate Group with A2 and above initial proficiency (n = 10). As shown in Table 6, the progress (W = .723, p = .002) and learning time (W = .819, p = .025) of the Intermediate Group are not normally distributed, so the Mann-Whitney U Test is performed to explore the differences between the two groups.

Table 6

Normality Check of Progress and Learning Time in Novice and Intermediate Groups

	Group	Shapiro-Wilk			Skewness		Kurtosis	
		W	*df*	Sig. (*p*)	Statistic	Std. Error	Statistic	Std. Error
Progress	Novice	.879	8	.185	-.599	.752	-1.026	1.481
	Intermediate	.723	10	.002	2.301	.687	6.575	1.334
Learning Time	Novice	.839	8	.073	-1.563	.752	2.567	1.481
	Intermediate	.819	10	.025	.998	.687	-.674	1.334

Note. W = Shapiro Wilk statistics*; df* = degree of freedom

Table 7

Mann-Whitney U Test Comparing the Results of Novice and Intermediate Learners

	Novice Group			Intermediate Group			*U*	*Z*	*p*	*r*
	Mdn.	*M.*	*SD.*	*Mdn.*	*M.*	*SD.*				
Progress	6	5.5	2.673	1	2	5.312	12.500	-2.468	.012	-.58
Learning Time	473.5	423.88	166.06	438	596	444.8	37.000	-.267	.829	-.063

Note. Mdn. = Median; *M.* = Mean; *SD*. = Standard Deviation; *U* = Mann-Whitney U; *Z* = standardized test statistic; *p* = significance; *r* = effect size

Table 7 shows the output of the Mann-Whitney U Test with statistics of both groups. Distributions of the progress scores for the two groups were significantly different (U = 12.5, Z = -2.468, p = .012), indicating the novice learners (Mdn = 6) made significantly more

progress than the intermediate learners (*Mdn* = 1) during the four weeks, and the effect size *r* = -.58 suggests a large effect. However, no significant difference was found between the learning time of the two groups (*U* = 37, *Z* = - .267, *p* = .829, *r* = -.063).

4.2. Students' Motivation in Learning French

4.2.1. Reliability of the Motivation Survey

To assess the internal consistency of the motivation survey, Cronbach's Alpha (Cronbach, 1951) for both intrinsic motivation and extrinsic motivation was presented in Table 8. Each of the two motivation scales includes 5 Likert-Scale items with ratings ranging from 1 to 5 points. The Cronbach's Alpha values suggest that the survey items are reliable for intrinsic motivation (α = .822) and acceptable for extrinsic motivation (α = .645) constructs. These results support the use of these scales in this thesis and further analysis to explore the motivations of Duolingo users. However, future studies should also revise the items, especially for the extrinsic motivation components to increase its consistency and reliability.

Table 8.

Reliability Check of the Motivation Survey

	Cronbach's α	*Cronbach's α Based on Standardized Items*	*Number of Items*
Intrinsic Motivation	.822	.830	5
Extrinsic Motivation	.645	.670	5

4.2.2. Overview of Learners' Motivation

The descriptive statistics of the motivation survey are shown in Table 9, suggesting a higher level of intrinsic motivation (*M* = 4.03 ± 0.16) than the extrinsic one (*M* = 3.56 ± 0.14).

Looking closer at the statistics of each Likert-Scale item (Table 10), the highest-rated statement is Question 3: "Compared with learning through textbooks and classes, using Duolingo is more interesting and enjoyable" (*M* = 4.5 ± 0.146); while the least agreed one is Question 6: 'Duolingo can help me get better grades in the language' (*M* =

3.278 ± 0.195). The mean value of all the items exceeded 3 points, indicating the participants have a positive view towards the motivational impacts of Duolingo compared with traditional learning methods.

Table 9

Statistics of the Motivation Survey

	N	Minimum	Maximum	Mean		Std. Deviation	Variance
				Statistic	Std. Error		
Intrinsic Motivation	18	2.60	5.00	4.0333	.16108	.68342	.467
Extrinsic Motivation	18	2.40	4.60	3.5556	.13677	.58029	.337

Table 10

Statistics of Each Likert-Scale Item

Items	*N*	Minimum	Maximum	Mean		Std. Deviation	Variance
				Statistic	Std. Error		
1	18	2.00	5.00	3.778	.207	.878	.771
2	18	2.00	5.00	4.056	.235	.998	.997
3	18	3.00	5.00	4.500	.146	.618	.382
4	18	2.00	5.00	3.722	.240	1.018	1.036
5	18	2.00	5.00	4.111	.212	.900	.810
6	18	1.00	5.00	3.278	.195	.826	.683
7	18	2.00	5.00	3.722	.211	.895	.801
8	18	2.00	5.00	3.778	.173	.732	.536
9	18	2.00	5.00	3.444	.217	.922	.850
10	18	2.00	5.00	3.556	.258	1.097	1.203

Notably, Question 11 and 12 of the survey investigate the learners' willingness to continue using the app after the experiment and the reasons why they would (or would not) use it. 88.9% of the learners (16 out of 18 participants) indicated "Yes", and their reasons

are presented in Figure 15. Each participant could select as many options as they wanted. The most important benefit of Duolingo to keep students' intended perseverance lies in its flexibility and convenience to learn anywhere, followed by the interesting and enjoyable learning experience.

Figure 15

Reasons to Continue Using Duolingo

Only two participants indicated they would not use Duolingo any more. One of them complained about the unsystematic learning materials and noted "You can't see your progress with relation to a whole picture. It's more like an after-class activity". The other mentioned the unhelpful learning materials and slow progress. They wrote in the open question "Sometimes I do prefer the language learning way in China where teachers stuff you with loads of grammar and practices".

4.2.3. Correlations Between Learners' Motivation and Progress

As the progress, intrinsic motivation and extrinsic motivation variables are all normally distributed, the Pearson Correlation Test is applied to analyzing these data. As shown in Table 11 and 12, no significant correlations were found either between the progress and intrinsic motivation (r = -.32, p = 0.898) or between the progress and extrinsic motivation (r = .005, p = 0.983). Therefore, learners' progress in French proficiency is not significantly related to their motivation levels, which contradicts most prior motivation research in the language learning field. The reasons that may account for this phenomenon will be explored in the discussion chapter.

Table 11

Pearson Correlation Between Progress and Intrinsic Motivation

		Progress	Intrinsic Motivation
Progress	Pearson Correlation (*r*)	1	-.032
	Sig. (2-tailed, *p*)		.898
	N	18	18
Intrinsic Motivation	Pearson Correlation (*r*)	-.032	1
	Sig. (2-tailed, *p*)	.898	
	N	18	18

Table 12

Pearson Correlation Between Progress and Extrinsic Motivation

		Progress	Extrinsic Motivation
Progress	Pearson Correlation (*r*)	1	.005
	Sig. (2-tailed, *p*)		.983
	N	18	18
Extrinsic Motivation	Pearson Correlation (*r*)	.005	1
	Sig. (2-tailed, *p*)	.983	
	N	18	18

4.3. Learning Experiences With Duolingo

The semi-structured interviews not only focused on the experience of learning French through Duolingo, but also broadly investigated the learners' attitudes in learning other languages, using other apps and some classroom activities. Through the semantic analysis of the interview data, four basic themes were identified regarding learning motivation, users' experience, comparison between MALL and traditional methods, and limitations of Duolingo. Each of the themes was explained in great depth with supporting quotes from the interviewees' feedback.

4.3.1. Participants' Motivation in MALL

When asked about their motivation in learning a foreign language, the interviewees primarily reported extrinsic motivation that is related to instrumental needs, such as social communication, future work and academic requirements. Intrinsic motivation was just briefly mentioned but they soon switched to the discussion of extrinsic motivation.

> I think French is quite a romantic language, so I just wanted to learn it. I've been learning French continuously because I've already invested so much time and effort. I want to learn it better, at least reach the same level as my English, when communicating with others, I won't be stumbling over my words. (Interviewee B)

> I am considering an internship in France in the future. I think that would be an opportunity to practice, because I kind of want to go to the United Nations or universities. (Interviewee C)

> In fact, I don't have much interest in French because I feel I won't need it. Even for travelling, English is usually enough. But I like Cantonese because I love Hong Kong. I feel Cantonese has a unique charm. I have considered going to Hong Kong for work in the future. (Interviewee D)

> When we were learning a second foreign language at school, we had options of French, Japanese, or Russian, and I chose French. I felt it was right because I was somewhat familiar with French culture and interested in it. Also, I thought French and English were quite similar, so it might be easier to learn. The most important thing is to pass the exam first, right? (Interviewee E)

Moreover, the interviewees highlighted some features of Duolingo that positively influenced their motivation. The day streak challenge enhanced their perseverance, the short lessons allowed people to learn flexibly at any time they want, and the knowledge has been integrated into real-life scenarios to keep users motivated.

> I’m still using Duolingo, one lesson per day just to keep the streak. It's the motivation for learning that can push you a bit, otherwise you might not even remember to study, and then procrastinate again. (Interviewee A)

> It can also keep your motivation to continue learning because each lesson is short. You can use it anytime, so it's very convenient and suitable for busy people like us. (Interviewee A)

> When it comes to motivation, I feel Duolingo deliberately incorporates related vocabulary into its storytelling sections. It tries to integrate words related to everyday life. Sometimes, when I'm walking in France and see a word I've learned, I get quite excited. It makes daily practice more meaningful to me. (Interviewee C)

However, Interviewee B who achieved Level B2 in the pretest pointed out the demotivating factors of Duolingo, especially for more advanced learners. As the learners approach a higher level, the content offered by the app becomes repetitive, which slows down the progress and demotivates learners from using it continuously.

> Initially, I used Duolingo frequently, but as I progressed to around B2, its pace slowed down, which became frustrating. I lost motivation and eventually stopped using it. Upon reaching the more advanced levels, it kept repeating earlier material rather than new knowledge, which made the learning slow. Personally, I need intensive language input in a short period to memorize information and progress quickly to a certain level. (Interviewee B)

4.3.2. Users' Experience with MALL

Both positive and negative comments on the experiences of MALL were reported by the participants. The additive nature of Duolingo and the gamified elements such as daily check-in, XP rewards and weekly rankings were highly valued, especially by Interviewee C. Its convenience and beginner-friendly operations were also noted.

> When I was in France, we were still walking outside late in the night, until after 11 o'clock. Suddenly I realized I hadn't used Duolingo today. So I just read aloud on the streets in France, it was really crazy. My friends around me felt so embarrassed that I was reading in public. But I kept reading. I quite like this kind of daily check-in and repetition. (Interviewee C)

> I was really obsessed with it. At the time of the weekly assessment of rankings, if I happened to be at the threshold of promotion, I would go crazy for it, just do everything I could to earn XPs, pushing myself to rank in the top three. (Interviewee

C)

Duolingo is really fun! I especially like the Match Madness Challenge (a timed challenge where the users match the words with their meanings). I think it eventually becomes a kind of muscle memory. When you see a word, you don't really think about it, just quickly pick the meaning. I find that pretty fun. (Interviewee C)

I find that it really helped me to remember certain grammar and words through repetition. I don't consciously memorize the words or anything like that, but they just appeared multiple times on the app. I feel that this repetitive method is quite helpful. (Interviewee C)

The layout of its pages, the interface is simple to use, and the cartoon designs are quite adorable. The operation is also good, very convenient. (Interviewee B)

If you start learning from scratch using this method, it would be interesting. Learning a completely unfamiliar language in this way could provide a basic understanding and make it engaging, which makes it a good way to explore new languages. (Interviewee E)

However, the participants also reported some negative feedback, especially participant B, who complained about the slow progress. Just contrary to the responses from C in the previous paragraphs, they disliked the embarrassing oral practices. Participant E also pointed out the problem of unsystematic grammar instruction.

I would never read aloud (using Duolingo) in public places. I'd rather skip speaking exercises, especially when it's inconvenient. I feel it’s stupid, so sometimes I just skip those oral practices. (Interviewee B)

I initially wanted to learn Korean with Duolingo, but found it inefficient. So, I started looking for formal online Korean courses instead. (Interviewee B)

> I prefer learning where I can quickly grasp everything I need to know at once. I want to absorb everything in one go rather than having information drip-fed to me slowly. I feel (using Duolingo) I have to complete one thing before moving on to the next. (Interviewee B)

> Fragmented time might be more suitable for vocabulary learning or listening practice, however, when it comes to learning grammar, I need a longer period of focused time or at least a more structured course setup (Interviewee E)

4.3.3. Comparison Between Traditional Learning and MALL

When asked about the differences between MALL and learning in formal classes either online or in person, the most notable advantage of mobile apps is the convenience that enables people to learn anytime and anywhere.

> It takes fragmented time to study, which is very suitable for people like me who always want to learn but struggle to realize, or who have no time to take formal language classes. (Interviewee A)

> I often use Duolingo while walking or when I'm in bed. Using Duolingo is the first thing I do when I wake up. I spend around 40 minutes on it in my bed. This kind of learning isn't confined to a specific place or time. It's not formal studying, so I don't see it as a burden. Sometimes, it’s just a way to kill time. (Interviewee C)

Although the convenience and pleasure of the mobile app were noted, most of the participants still believed classroom learning or online formal classes were more efficient to learn a language well. Interactions among classmates, intensive information input, systematic knowledge, customized learning experiences and reliable speaking exercises are the advantages of traditional learning valued by the participants.

> I definitely prefer learning in a classroom setting because I enjoy the interactivity. Studying with peers and classmates is more engaging and interactive. Although Duolingo has friend quests and some interactive features, they are too limited. There are not many activities, at least for me. (Interviewee A)

> When we were learning languages, our teachers would develop various fun games

to make the learning enjoyable. In terms of efficiency, classroom is definitely better because the period is longer and more structured, allowing for more contents to be taught. However, the benefit of apps like Duolingo is that you can learn in short, fragmented periods. But it’s reasonable if you spend less time, you'll definitely learn more slowly. (Interviewee A)

I think Duolingo is more suitable for those beginners learning purely for fun. Once you reach a more advanced level, you need something more professional. For me, spending five minutes on Duolingo might be less effective than reading a 5-minute text on my own. (Interviewee B)

Duolingo feels very mechanical. But the traditional methods, I mean one-on-one lessons, classroom learning, or self-study by reading books, it’s more flexible, and if you have a tutor you can customize the learning to your needs. (Interviewee D)

Another problem is the speaking practices on the app. It checks your pronunciation based on algorithms, which can be unreliable. There's no face-to-face interaction, making it very mechanical. I took IELTS courses with a foreign tutor and spent over ten hours in one-on-one sessions. I found these very effective, especially for speaking skills. The foreign tutor really made a big difference in the later stages of my preparation. (Interviewee D)

Apps like Duolingo, with many gamified features, might be more effective for children. But if I really want to learn something, I'd probably prefer doing more traditional learning or exercises. (Interviewee E)

4.3.4. Limitations of Duolingo

Besides the shortcomings mentioned in comparison to traditional formal language learning, other limitations of Duolingo were also discussed by the interviewees. Participant B disliked the day-streak features that push learners to keep learning, which were actually favoured by other participants.

If I care too much about such indicators, like day streak, rankings and learning

> statistics, it might give me additional pressure, and I might not want to do anything at all in the end. Just don't push me, or I'll go crazy. (Interviewee B)

Moreover, participant C doubted the actual applicability of the knowledge offered by the app, because some of the example sentences are quite strange in daily communication. However, as they just learned for 4 weeks on Duolingo, their current stage only provides limited vocabulary to make sentences, and more practical examples and resources are provided by the app in the later stages.

> I think it's because the app's vocabulary and resources are quite limited, it's difficult to express what I really want to say. If I want to introduce myself, I definitely can't say "Je suis un cheval" (I am a horse), right? I need to say who I am, where I'm from in China, and what I'm studying at a tertiary education level. but the app hasn't offered the vocabulary so far, and I feel it's kind of hard to apply what I have learned in real life. (Interviewee C)

In summary, although most interviewees appreciated the convenience and motivational incentives of Duolingo, such as the day-streak challenge and rankings, significant drawbacks were also highlighted, including the repetitive exercises, fragmented knowledge and limitations in advanced stages of learning. The findings indicate that Duolingo can be an entry-level tool for language learners, however, it is unlikely to be a qualified alternative to formal classroom instruction. The next chapter will discuss the findings of this research in great depth and detail, with substantial literature from predecessors, seeking to address the existing disputes over the MALL efficacy, motivation and potential limitations of this study.

Chapter 5. Discussion

In this chapter, the findings of each research question are discussed separately, and the limitations of this whole study will be elaborated subsequently, focusing on the sample, duration, method of data collection and assessments.

5.1. Learners' French Proficiency

Although the results found a significant difference between the participants' pretest scores and posttest scores, however, we should be very cautious to claim causal impacts of Duolingo on learners' academic gains, because this study adopted a quasi-experimental design involving only one group of participants. We could only conclude that the participants' French scores were significantly improved after the 4-week intervention, but a more rigorous experimental design with random allocation of the participants to either the treatment group or the control group should be conducted to further explore the causal effects of Duolingo on learners' achievement.

According to Spearman's rho, the participants' progress was not significantly correlated with the time spent on Duolingo. This finding might contradict our common sense that the more time you spend on learning, the more progress you are likely to make. However, it agrees with the results of most prior research. Loewen et al (2019) reported no strong correlation between study time (in hours) and Duolingo Progress Test scores in Turkish (N = 9) based on a 12-week intervention. Sudina and Plonsky (2024) found the total minutes of Duolingo exposure within the 6-month intervention only had weak to moderate associations with proficiency scores in Spanish and French (N = 287). Similarly, Kesser et al. (2023) reported weak-to-moderate but nonsignificant associations between the study time on either Babble (n = 27) or Duolingo (n = 32) and the Turkish test scores. Moreover, the non-linear nature of the MALL effect was also noted by Vesselinov et al. (2021), claiming that learning between 2-3 hours per week on Busuu is the most effective for reading, grammar and oral proficiency gains, less or more learning time would improve the achievement at a lower rate.

However, not many studies specifically measured the exact learning time on the mobile apps tested. Most prior research simply described the duration of the experiment as "the intervention lasts XX weeks", or described the frequency of treatment as 'the participants receive the treatment X times per week". These descriptions are quite general and may fail to show the actual time of exposure to the apps. The engagement time is crucial in evaluating the efficacy of MALL platforms and learners' achievement, and

therefore, more investigations are still needed to further explore the relations between study time and learners' gains.

Moreover, prior research in MALL also found advanced learners are likely to progress slower than beginners when learning through mobile apps (Vesselinov and Grego, 2012; Vesselinov and Grego, 2016b; Vesselinov and Grego, 2017). This claim is supported by the feedback from Interviewee B in this research. The significant difference between the progress of novice and intermediate groups revealed by the Mann-Whitney U Test also confirmed prior findings. Although the most dedicated learners who spent more than 1000 minutes during the experiment were all from the intermediate group, there was not a significant difference between the learning time of the two groups, and the novice learners made significantly more progress than the intermediate learners. However, as this research only involved beginners and intermediate French learners, it was not possible to further investigate the differences between intermediate and advanced learners. Future research could employ participants with various proficiency levels to explore the efficiency of mobile app instructions.

5.2. Students' Motivation in Learning French

Although learners generally showed positive feedback in both intrinsic and extrinsic motivation towards MALL compared with traditional learning, intrinsic motivation was rated higher than the extrinsic motivation scale. This finding is in line with prior motivation research related to MALL and technology-enhanced language learning (eg. Chen, 2020; Zeng & Fisher, 2023). The gamified elements of Duolingo and the rewarding systems can make the learning activity enjoyable, which is also supported by the interview data from the participants. However, One of the possible reasons for the lower ratings for extrinsic motivation lies in the design of the motivation survey, as shown by Cronbach's Alpha, the internal consistency of extrinsic motivation items is not as good as that of the intrinsic motivation constructs. Future studies should improve the design of the Likert-Scale items to make the survey more rigorous. Moreover, the constructs of extrinsic motivation in the SDT continuum are more complicated and instrumental, and it by nature expects positive outcomes from the learning activity. However, the insignificant correlation between learning time and progress revealed that the effort devoted to learning with the app may not necessarily lead to enhanced learning achievement, which could potentially frustrate the learners and impair their extrinsic motivation.

On the other hand, neither intrinsic nor extrinsic motivation significantly correlates with their progress in French proficiency. This result is not surprising because motivation is

just one of the numerous potential factors contributing to the success of language learning. Other variables, such as language aptitude, English proficiency, and time of exposure to the target language, were not strictly controlled or investigated in this research. In addition, motivation studies in language learning mostly focus on the influences of different teaching methods on learners' motivation through self-reports or subjective assessments, while the relationship between learners' motivation and actual achievement in tests has been rarely explored through robust objective data. Therefore, if future research seeks to investigate the potential correlation between motivation and language progress, the participants' demographic characteristics and learning conditions should be strictly controlled, and more objective measures should be integrated as well.

5.3. Interview about Learners' Experiences with MALL

According to the interviews with representative participants, in-depth feedback to MALL and Duolingo was collected. The interviewees mostly chose to learn a language driven by extrinsic motivation which focused on academic requirements and professional aspirations. Therefore, it is reasonable to assume that the instrumental functions of Duolingo as a learning tool are more important than its recreational features that just enhance learning experiences. The negative feedback from the interviewees largely supported this assumption, highlighting problems of slow progress, unsystematic knowledge, inefficient oral practice and a lack of interaction between users. Although they appreciated the adorable design, great convenience and interesting gamified features, they still believed formal learning in traditional ways is more effective in improving language proficiency, especially for advanced learners. MALL can be a good after-class activity to practice, but it is not a qualified alternative to formal teaching.

However, it is also important to note that the interviewees were purposefully, rather than randomly selected, aiming to provide more comprehensive and in-depth feedback. Therefore, the only two participants who selected 'NO' in the survey Question 11 "Will you continue to use Duolingo after the experiment ends" were both interviewed, and one of the only two participants who retrogressed in the posttest was interviewed as well. Therefore, the interview data would likely seem more negative than the overall users' feedback. Actually, the scalar ratings of the motivation survey showed both intrinsic and extrinsic motivation were rated positively (Mean scores reached 3 points and above) compared to traditional learning methods.

Furthermore, none of the interviewees fully explored the socializing functions of Duolingo. Although the researcher suggested that they could follow each other to work on

“friend quests” together and win rewards, few of them actually engaged in these activities. They simply followed the progress of the main menu and learned by themselves, which made the learning experience less interactive or cooperative.

5.4. Limitations of the Research

Although this research made valuable contributions to MALL and motivation in language learning by bridging the research gaps in target language, instruction language, cognitive theoretical framework and comprehensive methodologies, limitations still exist. The small sample size, short duration, self-reported data and the lax assessment methods would be the most prominent weaknesses.

5.4.1. Experiment Sample and Duration.

The small and biased sample has been one of the primary concerns of MALL experiments. Creswell (2014, p. 164) recommended a sample of approximately 15 participants in each experiment group, 30 participants for correlational research and 350 participants for survey studies. Based on this standard, the sample in this research barely met the standard for a single-group treatment while still far below the requirement of reliable surveys. The small sample size may reduce the statistical power and increase the risk of chance. Moreover, the sample was approached based on personal contact and social media posts, so they were all relatively good language learners who already learned L2 English successfully and took an interest in learning French. Some of the participants even learned several additional foreign languages before. So the sample may not be representative enough to reveal the whole story of the general population.

The short duration is another problem of the experiment. According to Burston and Arispe (2022), language learning takes time, and any intervention that spans less than an academic term is unlikely to produce meaningful outcomes. A treatment duration of at least 8 weeks in empirical MALL research is recommended, because short intervention is likely to produce overly positive results due to the novelty effect of innovative technologies. This problem has also been frequently discussed in computer-assisted language learning studies, indicating that learners’ initial enthusiasm towards a new learning tool can significantly influence their motivation, but the passion may be temporary and may not necessarily transfer to long-term effects (Chwo et al., 2018; see also Stockwell & Hubbard, 2013; Murray & Barnes, 1998). Peng et al.’s (2021) meta-analysis of 17 MALL studies further supported the existence of the novelty effect. They found the longer an experiment duration lasts, the less effective its mobile technology is, because the boredom and fatigue

phenomena may occur after the initial excitement, which could slow down their progress. However, this research fail to meet the 8-week minimum requirement due to limited time for the master thesis, so the test results and motivation ratings are likely to be less positive if a longer learning session is adopted. Future MALL research should take the novelty effect into consideration, adopt a larger and more representative sample, and span a longer treatment duration to produce more rigorous results.

5.4.2. Data collection

Another limitation that might undermine the reliability of the research lies in the self-reported data. Although the participants were asked to provide screenshots of their screen time or weekly reports to prove their time of exposure to Duolingo, there are possibilities that they could simply open the app and leave it there without engaging in the learning activities. Some researchers adopted other variables to quantify learners' engagement, such as the number of lessons completed, or the XPs earned on Duolingo (e.g. Loewen et al., 2019). These measures might work better in ensuring learners' engagement, but also have their own limitations. The number of lessons only counts the main learning activities but ignores other additional activities such as reviewing personal mistakes, socializing with other users and participating in extra challenges on the apps. Meanwhile, the XPs can be earned in various ways, including daily check-ins and sharing links, that could not accurately reflect learners' efforts.

Furthermore, the pretest and posttest scores were also collected through the screenshots of the test results. As the tests were not conducted in controlled conditions, it was hard to prohibit cheating or seeking help from others. Although the participants were told to finish the tests by themselves and only refer to their own knowledge, it was not possible to strictly control their actual activities. Also, the pretest and the posttest were completely the same, so if some learners checked and memorized the answers after the pretest, their posttest scores would be greatly improved, which is likely to make the results overly positive. Therefore, it is suggested that MALL research should quantify learners' engagement with multiple indicators, rather than solely relying on learning time or just one other variable, and the assessment should be conducted in a controlled laboratory condition.

5.4.3. Assessment Instrument

This research mainly included two instruments, the proficiency test and the motivation survey. However, it is hard to claim the rigour of any of them. The proficiency

test was chosen primarily because it caters to learners from A1 to B2 levels which meets the needs of the sample, and it does not require registration, personal information, or a test fee. The test provides only a brief reference to the estimated overall competence in the target language rather than a certified proof of linguistic proficiency, and the reliability was not tested in terms of its consistency with other well-established and authoritative proficiency tests. Furthermore, it covers a broad range of proficiency levels with just 40 questions, making it difficult for the learners to make a notable difference within just 4 weeks. The tiny progress may only result in an increase of one or two points in the posttest, which is easily influenced by random guess or chance. One interviewee also noted this issue in the interview:

> For that test, I didn't feel much difference when I did it in the pretest or the posttest. Maybe because it was quite broad and general, but I only studied a few pieces of knowledge for a short period of time. So, it's like you go only from three to five, but it’s not reflected by the test. (Interviewee A)

The design of the motivation survey was guided by SDT, and its reliability was checked through Cronbach’s Alpha. However, the internal consistency of the items, especially the constructs of extrinsic motivation, was not so desirable. It had not been realized until it was too late that more extensive resources should be carefully consulted in designing the survey. Several valuable studies with more rigorous motivation questionnaires (e.g. Akman & Karahan, 2023; Chen & Zhao, 2022; Vesselinov, 2009b) were found when this thesis was being composed, which could be really helpful in improving this instrument. However, as there has not been a universally accepted motivation survey in MALL practice, more efforts could be made to revise and improve the components to make it more meaningful in future studies.

Chapter 6. Conclusion

Mobile learning has attracted researchers' attention over the recent decades thanks to technological advancements and increased demands for the flexibility and convenience of learning activities. Against this trend, mobile-assisted language learning (MALL) platforms have proliferated, generating numerous studies to investigate their impacts on learning outcomes, learners' motivation and experiences. Although prior research explored the apps' effectiveness in boosting academic achievement, there has been a large discrepancy between the claims of commercial materials or app-funded research and the results reported by independent researchers. Moreover, although most language learning apps claim to boost learners' motivation, insufficient research has been rigorously designed to systematically explore this issue, especially with the guidance of SDT.

To address the limitations of the current literature, this research seeks to make valuable contributions by investigating the Chinese ESL students' achievement, motivation and experiences in learning French through English instruction on Duolingo. To achieve this goal, a mixed-method approach incorporating a pretest, a posttest, a motivation survey and optional semi-structured interviews is adopted to collect robust data. The research recruited 18 college students through convenience sampling and spanned a duration of 4 weeks. The participants were allowed full autonomy and freedom to manage their learning schedule and activities, in order to truthfully reflect the usage scenario. This study targets three research gaps in current MALL research: (1) Non-English languages have been under-researched, especially through L2 instruction; (2) Theoretical frameworks have been insufficiently applied to MALL motivation research, especially the SDT; (3) Most MALL research are not rigorously designed with comprehensive methodologies or objective instruments.

The results revealed that the participants made significant progress after learning with Duolingo for 4 weeks, but their progress was not correlated with the number of minutes they spent on the app. Among all the test-takers, the novice learners significantly outperformed the intermediate learners in their progress of French proficiency, but the learning time of both groups did not differ significantly. The students also exhibited positive attitudes to Duolingo compared with traditional learning in both intrinsic and extrinsic motivation, but neither of the two motivation scales correlated with their progress in the proficiency tests. Their comments and complaints about MALL and Duolingo were also analyzed in great detail. Most interviewees valued the convenience and enjoyable learning experiences, noting Duolingo improved their perseverance and willingness to learn French.

Meanwhile, some participants pointed out Duolingo's problems, including unsystematic grammar knowledge, mechanical speaking exercises and slow progress for advanced learners.

These findings provide valuable implications for language learners, educators, researchers and mobile app companies. First of all, the gamified features and convenience of language learning apps make them useful tools to acquire and practice a language for beginners, but they might be less efficient for more advanced learners. Moreover, educators are suggested to integrate MALL as an assisting role to formal instruction because it can enhance students' learning outcomes, perseverance and motivation. However, formal instructions are still important as they provide more systematic and intensive knowledge in a more interactive and cooperative way. Researchers should carry out more rigorous and unbiased studies to investigate the impacts of MALL against the rapid evolution of technologies and the emergence of new MALL platforms. Finally, mobile app companies should value the research findings and the users' feedback, upgrade their products and enhance the design of courses, making it more competitive by delivering more systematic language courses and catering to the needs of more advanced learners.

However, due to limited time, resources and personal ability, shortcomings still exist in this research, undermining its rigour and reliability. The small sample size, short duration, self-reported data and loosely-designed instruments are the most notable problems. More attempts should be made in future MALL research to improve the research conditions and methodologies to produce more meaningful and robust results.

References

AbuSa'aleek, A. O. (2014). A review of emerging technologies: Mobile assisted language learning (MALL). *Asian Journal of Education and e-learning*, *2*(6), 469-475. https://www.ajouronline.com/index.php/AJEEL/article/view/1995

Adara, R. A. (2020). Negative Effects of MALL on the Improvement of Learners Autonomy and Motivation. *The International Conference on Innovations in Social Sciences and Education*, *1*(1), 613-622. https://conference.loupiasconference.org/index.php/ICoISSE/article/view/84/84

Akman, E., & Karahan, P. (2023). ELT Students' Perceptions Toward Mobile-Assisted Language Learning (MALL): Exploring its Effects on Motivation and Learner Autonomy. *International Journal of Educational Researchers (IJERs)*, *14*(2), 1-20. https://doi.org/10.29329/ijer.2023.565.1

Alfuhaid, S. R. (2021). The Utilisation of Duolingo to Enhance the Speaking Proficiency of EFL Secondary School Students in Saudi Arabia. *English Language Teaching*, 14(11), 9-15. https://doi.org/10.5539/elt.v14n11p9

Al-Hoorie, A. H. (2018). The L2 motivational self system: A meta-analysis. *Studies in Second Language Learning and Teaching*, *8*(4), 721-754. https://doi.org/10.14746/ssllt.2018.8.4.2

Alemi, M., Sarab, M. R. A., & Lari, Z. (2012). Successful learning of academic word list via MALL: Mobile Assisted Language Learning. *International Education Studies*, *5*(6), 99-109. https://doi.org/10.5539/ies.v5n6p99

Babble. (n.d.). *Research at Babble*. https://uk.babbel.com/why-babbel-works

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Burston, J. (2015). Twenty years of MALL project implementation: A meta-analysis of learning outcomes. *ReCALL*, *27*(1), 4-20. https://doi.org/10.1017/S0958344014000159

Burston, J., & Arispe, K. (2022). Experimental MALL research in SLA: Past, present, and future. In N. Ziegler & M. González-Lloret (Eds.), *The Routledge handbook of second language acquisition and technology* (1st ed., pp. 326-340). Routledge. https://doi.org/10.4324/9781351117586

Burston, J., & Giannakou, K. (2022). MALL language learning outcomes: A comprehensive meta-analysis 1994–2019. *ReCALL*, *34*(2), 147-168. https://doi.org/10.1017/S0958344021000240

Busuu. (n.d.) *Busuu Research.* https://www.busuu.com/en/research

Chen, C. H. (2020). AR videos as scaffolding to foster students' learning achievements and motivation in EFL learning. *British Journal of Educational Technology*, *51*(3), 657-672. https://doi.org/10.1111/bjet.12902

Chen, Y., & Zhao, S. (2022). Understanding Chinese EFL learners' acceptance of gamified vocabulary learning apps: an integration of self-determination theory and technology acceptance model. *Sustainability*, *14*(18), Article 11288. https://doi.org/10.3390/su141811288

Chen, Z., Chen, W., Jia, J., & An, H. (2020). The effects of using mobile devices on language learning: A meta-analysis. *Educational Technology Research and Development*, 68(4), 1769-1789. https://doi.org/10.1007/s11423-020-09801-5

Chwo, G. S. M., Marek, M. W., & Wu, W.-C. V. (2018). Meta-analysis of MALL research and design. *System*, *74*, 62-72. https://doi.org/10.1016/j.system.2018.02.009

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). L. Erlbaum Associates. https://doi.org/10.4324/9780203771587

Creswell, J. W. (2014). *Educational research : Planning, conducting and evaluating quantitative and qualitative research* (4th ed.). Pearson Education Limited.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334. https://doi.org/10.1007/bf02310555

Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of research in personality*, *19*(2), 109-134. https://doi.org/10.1016/0092-6566(85)90023-6

Deci, E. L., & Ryan, R. M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian Psychology*, *49*(3), 182-185. https://doi.org/10.1037/a0012801

Derakhshan, A. (2011). The impact of text-messaging on EFL freshmen's vocabulary learning. *The EUROCALL Review*, 19, 47-55. https://doi.org/10.4995/eurocall.2011.16271

Dörnyei, Z. (1994). Understanding L2 motivation: On with the challenge! *The Modern Language Journal*, *78*(4), 515-523. https://doi.org/10.2307/328590

Dörnyei, Z. (2009). The L2 motivational self system. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (1st ed., Vol. 36, pp. 9-42). Multilingual Matters. https://doi.org/10.21832/9781847691293-003

Dörnyei, Z., & Ushioda, E. (2021). *Teaching and researching motivation* (3rd ed.). Routledge. https://doi.org/10.4324/9781351006743

Dörnyei, Z. (2005). *The Psychology of the Language Learner: Individual Differences in Second Language Acquisition* (1st ed.). L. Erlbaum. https://solo.bodleian.ox.ac.uk/permalink/44OXF_INST/35n82s/alma991022194263007026

Duman, G., Orhon, G., & Gedik, N. (2014). Research trends in mobile assisted language learning from 2000 to 2012. *ReCALL*, *27*(2), 197-216. https://doi.org/10.1017/S0958344014000287

Duolingo. (2020). *duocon 2020 highlights*. https://www.duolingo.com/duocon2020

Duolingo (2024, June). *Duolingo Language Courses*. https://www.duolingo.com/courses/all

Duolingo. (2024). *Duolingo Reports 65% DAU Growth, 45% Revenue Growth and Record Profitability in Fourth Quarter 2023.* https://investors.duolingo.com/static-files/f456546b-5c18-4e4b-966d-59534089ef02

Duolingo. (n.d.). *Duolingo Works*. https://www.duolingo.com/efficacy

Fithriani, R., Dewi, U., Daulay, S. H., Salmiah, M., & Fransiska, W. (2019). Using Facebook in EFL writing class: Its effectiveness from students' perspective. *KnE Social Sciences*, *3*(19), 634–645. https://doi.org/doi.org/10.18502/kss.v3i19.4892

Forbes. (2019 July 16). *Next Billion-Dollar Startups 2019.* https://www.forbes.com/sites/amyfeldman/2019/07/16/next-billion-dollar-startups-2019/

Gagné, M., & Deci, E. L. (2005). Self-determination theory and work motivation. *Journal of Organizational behavior*, *26*(4), 331-362. https://doi.org/10.1002/job.322

Gardner, R. C. (1988). The socio-educational model of second-language learning: Assumptions, findings, and issues. *Language Learning*, *38*(1), 101-126. https://doi.org/10.1111/j.1467-1770.1988.tb00403.x

Gardner, R. C. (2019). The socio-educational model of second language acquisition. In M. Lamb, K. Csizér, A. Henry, & S. Ryan (Eds.), *The Palgrave handbook of motivation for language learning* (pp. 21-37). Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-030-28380-3_2

Gardner, R. C., & Lambert, W. E. (1959). Motivational variables in second-language acquisition. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *13*(4), 266-272. https://doi.org/10.1037/h0083787

Gardner, R. C., & Lambert, W. E. (1972). *Attitudes and motivation in second-language learning*. Newbury House Publishers.

Gonulal, T. (2019). The use of Instagram as a mobile-assisted language learning tool. *Contemporary Educational Technology*, *10*(3), 309-323. https://doi.org/10.30935/cet.590108

Gutiérrez-Colón, M., Frumuselu, A. D., & Curell, H. (2023). Mobile-assisted language learning to enhance L2 reading comprehension: A selection of implementation studies between 2012–2017. *Interactive Learning Environments*, *31*(2), 854-862. https://doi.org/0.1080/10494820.2020.1813179

Habibie, A. (2020). DUOLINGO AS AN EDUCATIONAL LANGUAGE TOOL TO ENHANCE EFL STUDENTS'MOTIVATION IN WRITING. *British (Jurnal Bahasa dan Sastra Inggris)*, *9*(1), 13-26. https://journal.umgo.ac.id/index.php/British/article/view/502

Hammarberg, B. (2001). Roles of L1 and L2 in L3 production and acquisition. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *Cross-Linguistic Influence in Third Language Acquisition : Psycholinguistic Perspectives* (pp. 21-41). Multilingual Matters. https://doi.org/10.21832/9781853595509-003

Heil, C. R., Wu, J. S., Lee, J. J., & Schmidt, T. (2016). A review of mobile language learning applications: Trends, challenges, and opportunities. *The EuroCALL Review*, *24*(2), 32-50. https://doi.org/10.4995/eurocall.2016.6402

Huang, Z. (2020). *Thirteen years since the first iPhone: A systematic review on the effectiveness of language learning apps on smart devices*. [Master's Thesis, University of Oxford]. Oxford University Research Archive. https://solo.bodleian.ox.ac.uk/permalink/44OXF_INST/35n82s/alma991025233671007026

Jeong, K.-O. (2023). Integrating technology into language teaching practice in the post-COVID-19 pandemic digital age: From a Korean English as a foreign language context. *RELC Journal*, *54*(2), 394-409. https://doi.org/10.1177/00336882231186431

Jiang, X., Rollinson, J., Plonsky, L., & Pajak, B. (2020). *Duolingo efficacy study: Beginning-level courses equivalent to four university semesters* (DRR-20-04). Duolingo. https://duolingo-papers.s3.amazonaws.com/reports/duolingo-efficacy-whitepaper.pdf

Kacetl, J., & Klímová, B. (2019). Use of smartphone applications in english language learning—A challenge for foreign language education. *Education Sciences*, *9*(3), Article 179. https://doi.org/10.3390/educsci9030179

Kessler, M., Loewen, S., & Gönülal, T. (2023). Mobile-assisted language learning with Babbel and Duolingo: comparing L2 learning gains and user experience. *Computer*

Assisted Language Learning, 1-25. https://doi.org/10.1080/09588221.2023.2215294

Kittredge, A., Peters, R., Neumann, F., & Jiang, X. (2024). *Duolingo learners can start a conversation after 4-6 weeks of app use* (DRR-24-01). Duolingo. https://duolingo-papers.s3.amazonaws.com/reports/Duolingo_whitepaper_language_conversation_2024.pdf

Klimova, B., & Zamborova, K. (2020). Use of mobile applications in developing reading comprehension in second language acquisition—A review study. *Education Sciences*, *10*(12), Article 391. https://doi.org/10.3390/educsci10120391

Kukulska-Hulme, A., & Shield, L. (2007). *An overview of mobile assisted language learning: Can mobile devices support collaborative practice in speaking and listening*. T. O. University. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=089e6c20970b6503da4bb9f6b5177015247c2f42

Kukulska-Hulme, A., & Shield, L. (2008). An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction. *ReCALL*, *20*(3), 271-289. https://doi.org/10.1017/S0958344008000335

Lin, J.-J., & Lin, H. (2019). Mobile-assisted ESL/EFL vocabulary learning: A systematic review and meta-analysis. *Computer Assisted Language Learning*, *32*(8), 878-919. https://doi.org/10.1080/09588221.2018.1541359

Liu, P. L., & Chen, C. J. (2015). Learning English through actions: a study of mobile-assisted language learning. *Interactive Learning Environments*, 23(2), 158-171. https://doi.org/10.1080/10494820.2014.959976

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

Lord, G. (2015). " I don't know how to use words in Spanish":" Rosetta Stone" and Learner Proficiency Outcomes. *The Modern Language Journal*, *99*(2), 401-405. https://www.jstor.org/stable/43650036

Mahdi, H. S. (2018). Effectiveness of mobile devices on vocabulary learning: A meta-analysis. *Journal of educational computing research*, *56*(1), 134-154. https://doi.org/10.1177/0735633117698826

Mango Languages. (n.d.) *Discover the Impact of Mango Languages: Efficacy Survey Results*. https://mangolanguages.com/discover-the-impact-of-mango-languages-efficacy-study-results/

Murray, L., & Barnes, A. (1998). Beyond the “wow” factor—evaluating multimedia language learning software from a pedagogical viewpoint. *System*, *26*(2), 249-259. https://doi.org/10.1016/S0346-251X(98)00008-6

Niemiec, C. P., & Ryan, R. M. (2009). Autonomy, competence, and relatedness in the classroom: Applying self-determination theory to educational practice. *Theory and research in Education*, *7*(2), 133-144. https://doi.org/10.1177/1477878509104318

O'Malley, C., Vavoula, G., Glew, J., Taylor, J., Sharples, M., Lefrere, P., Lonsdale, P., Naismith, L., & Waycott, J. (2005). *Guidelines for learning/teaching/tutoring in a mobile environment*. https://hal.science/hal-00696244

Pelletier, L. G., Fortier, M. S., Vallerand, R. J., & Briere, N. M. (2001). Associations among perceived autonomy support, forms of self-regulation, and persistence: A prospective study. *Motivation and emotion*, *25*, 279-306. https://doi.org/10.1023/A:1014805132406

Peng, H., Jager, S., & Lowie, W. (2021). Narrative review and meta-analysis of MALL research on L2 skills. *ReCALL*, *33*(3), 278-295. https://doi.org/10.1017/S0958344020000221

Quinn, C. (2001). *mLearning: Mobile, wireless, in-your-pocket learning*. https://www.researchgate.net/profile/Clark-Quinn/publication/343083598_mLearning_Mobile_Wireless_In-Your-_Pocket_Learning/links/5f15dfcba6fdcc3ed718d7bf/mLearning-Mobile-Wireless-In-Your-Pocket-Learning.pdf

Ratzlaff, N. J. (2015). *A cross comparison and efficacy study of DuoLingo and an entry-level German 1A course* [Master’s Thesis, California State University]. Fresno. https://scholarworks.calstate.edu/downloads/zk51vj33z

Rodríguez-Fuentes, R. A., & Swatek, A. (2023). *A Comparison Between Classroom and MALL Instruction with Duolingo: Learning English at the A2 CEFR Level*. https://duolingo-papers.s3.amazonaws.com/reports/RodriguezFuentes_etal_whitepaper_language_proficiency_Columbia_2023.pdf

Rosell-Aguilar, F. (2017). State of the app: A taxonomy and framework for evaluating language learning mobile applications. *CALICO journal*, *34*(2), 243-258. https://www.jstor.org/stable/90014690

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist*, *55*(1), 68-78. https://doi.org/10.1037110003-066X.55.1.68

Schepens, J. J., Van der Slik, F., & Van Hout, R. (2016). L1 and L2 distance effects in learning L3 Dutch. *Language Learning*, *66*(1), 224-256. https://doi.org/10.1111/lang.12150

Shakarami, A., Khajehei, H., & Hajhashemi, K. (2014). Tech-assisted language learning tasks in an EFL setting: Use of hand phone recording feature. *International Journal of Applied Linguistics & English Literature*, *3*(5), 100-104. https://doi.org/10.7575/aiac.ijalel.v.3n.5p.100

Shortt, M., Tilak, S., Kuznetcova, I., Martens, B., & Akinkuolie, B. (2023). Gamification in mobile-assisted language learning: A systematic review of Duolingo literature from public release of 2012 to early 2020. *Computer Assisted Language Learning*, *36*(3), 517-554. https://doi.org/10.1080/09588221.2021.1933540

Solikhah, I., & Wirawati, D. (2020). Mobile Assisted Language Learning and Its Insights in Teaching Academic Writing: A Literature Review. *Asian EFL Journal*, *27*(2.2), 102-118. http://eprints.uad.ac.id/id/eprint/38458

Stockwell, G., & Hubbard, P. (2013). Some emerging principles for mobile-assisted language learning. *The International Research Foundation for English Language Education*, *2013*, 1-15. https://www.tirfonline.org/wp-content/uploads/2013/11/TIRF_MALL_Papers_StockwellHubbard.pdf

Sudina, E., & Plonsky, L. (2024). The effects of frequency, duration, and intensity on L2 learning through Duolingo: A natural experiment. *Journal of Second Language Studies*, *7*(1), 1-43. https://doi.org/10.1075/jsls.00021.plo

Taylor, J. (2006). *Evaluating mobile learning: What are appropriate methods for evaluating learning in mobile enviroments*. Kaleidoscope. https://www.cin.ufpe.br/~mlearning/intranet/m-learning/Big%20Issues%20in%20Mobile%20Learning.pdf#page=25

UNESCO. (2022, June 29). *Education: From disruption to recovery*. https://webarchive.unesco.org/web/20220629024039/https:/en.unesco.org/covid19/educationresponse/

Ushioda, E., & Dörnyei, Z. (2009). Motivation, language identities and the L2 self: A theoretical overview. In *Motivation, language identity and the L2 self* (Vol. 36, pp. 1-8). Multilingual Matters. https://doi.org/10.21832/9781847691293

Vallerand, R. J., Pelletier, L. G., Blais, M. R., Briere, N. M., Senecal, C., & Vallieres, E. F. (1992). The Academic Motivation Scale: A measure of intrinsic, extrinsic, and amotivation in education. *Educational and Psychological Measurement*, 52(4), 1003-1017. https://doi.org/10.1177/0013164492052004025

Vesselinov, R. (2009a). *Measuring the Effectiveness of Rosetta Stone*. http://comparelanguageapps.com/documentation/EffectivenessStudy2008_RS.pdf

Vesselinov, R. (2009b). *Measuring the Attitude and Motivation of Rosetta Stone Users*. http://comparelanguageapps.com/documentation/MeasuringTheAttitudeandMotivationofRSUsers.pdf

Vesselinov, R., & Grego, J. (2012). *Duolingo effectiveness study*. https://www.languagezen.com/pt/about/english/Duolingo_Efficacy_Study.pdf

Vesselinov, R., & Grego, J. (2015). *Efficacy of Language Zen*. http://comparelanguageapps.com/documentation/Language_Zen_Efficacy_Study.pdf

Vesselinov, R., & Grego, J. (2016a). *The busuu Efficacy Study*. http://comparelanguageapps.com/documentation/The_busuu_Study2016.pdf

Vesselinov, R., & Grego, J. (2016b). *The Babbel efficacy study*. http://comparelanguageapps.com/documentation/Babbel2016study.pdf

Vesselinov, R., & Grego, J. (2017). *Hello English Efficacy Study*. http://comparelanguageapps.com/documentation/HelloEnglish_2017Study.pdf

Vesselinov, R., & Grego, J. (2018). *The italki Efficacy Study*. https://www.gettingsmart.com/wp-content/uploads/2018/01/italki2018FinalReport.pdf

Vesselinov, R., & Grego, J. (2019a). *Pimsleur Efficacy Study*. http://comparelanguageapps.com/documentation/Pimsleur_EfficacyStudy2019.pdf

Vesselinov, R., & Grego, J. (2019b). *Mango Languages Efficacy Study*. http://comparelanguageapps.com/documentation/Mango_Languages_FinalReport2019.pdf

Vesselinov, R., Grego, J., Sacco, S. J., & Tasseva-Kurktchieva, M. (2019). *The 2019 Rosetta Stone efficacy study*. http://comparelanguageapps.com/documentation/The2019_RS_FinalReport.pdf

Vesselinov, R., Grego, J., Tasseva-Kurktchieva, M., & Sedaghatgoftar, N. (2021). *THE BUSUU EFFICACY STUDY 2021*. http://comparelanguageapps.com/documentation/Busuu_Efficacy_Study_2021.pdf

Vesselionv, R., Grego, J., & Tasseva-Kurktchieva, M. (2023). *LingQ Efficacy Study*. http://comparelanguageapps.com/documentation/LingQ_Efficacy_2023.pdf

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological review*, *92*(4), 548. https://doi.org/10.1007/978-1-4612-4948-1_6

West, M., & Vosloo, S. (2013). UNESCO policy guidelines for mobile learning. https://unesdoc.unesco.org/ark:/48223/pf0000219641

Wolpers, M. (2010). *Sustaining TEL: From innovation to learning and practice : 5th European Conference on Technology Enhanced Learning, EC-TEL 2010, Barcelona, Spain, September 28 - October 1, 2010, proceedings*. Springer. https://doi.org/10.1007/978-3-642-16020-2

Zain, D. S. M., & Bowles, F. A. (2021). Mobile-assisted language learning (MALL) for higher education instructional practices in EFL/ESL contexts: A recent review of literature. *Computer-Assisted Language Learning Electronic Journal*, *22*(1), 283-307. https://callej.org/index.php/journal/article/view/332

Zeng, C., & Fisher, L. (2024). Opening the “black box”: how out-of-class use of Duolingo impacts Chinese junior high school students’ intrinsic motivation for English. *ECNU Review of Education*, *7*(2), 283-307. https://doi.org/10.1177/20965311231171606

Appendix 1 Survey of Duolingo Users' Learning Motivation

Likert-Scale items

1. Compared with learning through textbooks and classes, Duolingo can help me gain more knowledge

2. Compared with learning through textbooks and classes, using Duoling helps me feel more fulfilling and accomplished

3. Compared with learning through textbooks and classes, using Duoling is more interesting and enjoyable.

4. Compared with learning through textbooks and classes, using Duolingo makes me fell easier to learn the language well.

5. Compared with learning through textbooks and classes, using Duolingo makes me love the language more

6. Duolingo can help me to get better grades in the language.

7. Duolingo helps me better communicate and get along with others using the language.

8. Using Duolingo to learn a language helps to exhibit my ability and makes me more respected by my friends, family or others.

9. Learning a language through Duolingo will make me more competitive in academic or career

10. Duolingo helps to reduce my frustration, misunderstanding, embarrassment or other potential disadvantages in language learning and using.

Other questions

11. Do you think you will continue to use Duolingo to learn a language after the experiment ends?

Yes

No

12a. If Yes, why?

Interesting and enjoyable learning experience

Keep the rank or streak

Useful learning materials and practices

Flexibility and convenience to learn anywhere

Simple interface and operations

Play with friends

Other: ____

12b. If No, why

Subscription fee

Materials are not helpful

Practices are repeated and monotonous

Not interactive enough

Do not like the interface or design

Do not need to learn a language

Other: ___

13. Do you have any other feedback to share? (Optional)

Appendix 2 Semi-Structured Interview Questions

(1) What motivates you to learner French and participate in this research?

(2) How do you feel about the designs and functions of Duolingo?

(3) How do you think Duolingo impacted on your motivation in learning French?

(4) What features do you find the most helpful about Duolingo?

(5) What do you think are the major limitations of Duolingo?

(6) What are the differences between traditional learning methods and MALL?

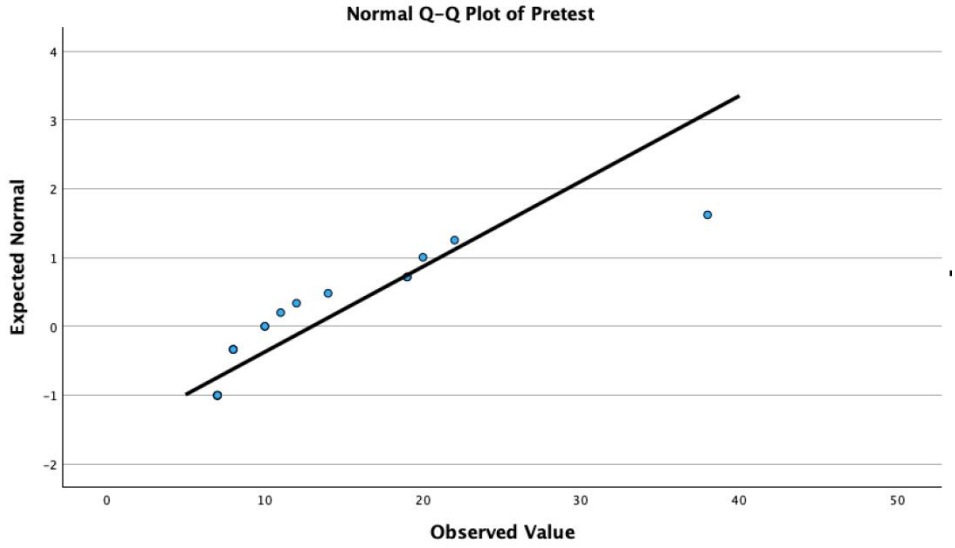
(7) (To some participants) What do you think are the reasons why you performed exceptionally well (or badly?) during the experiment period?

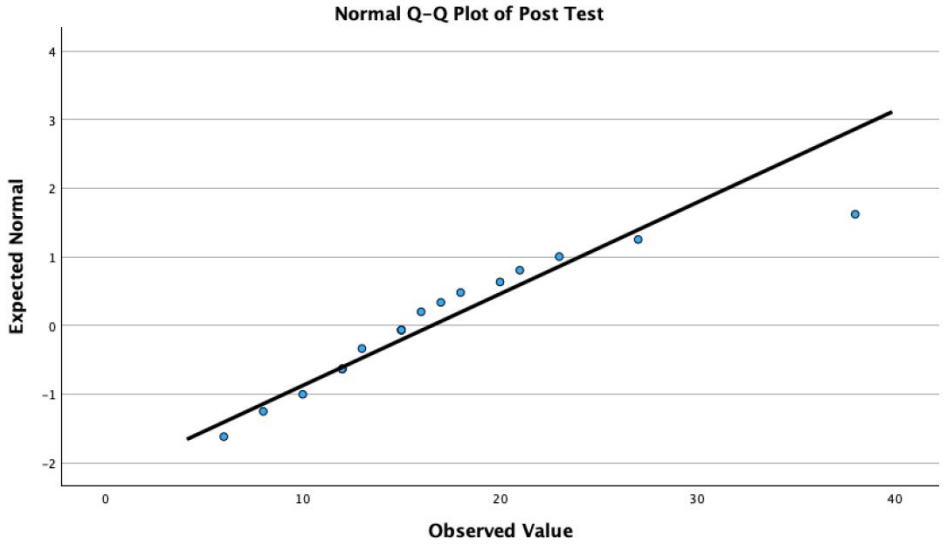
(8) Have you learned other languages with Duolingo?

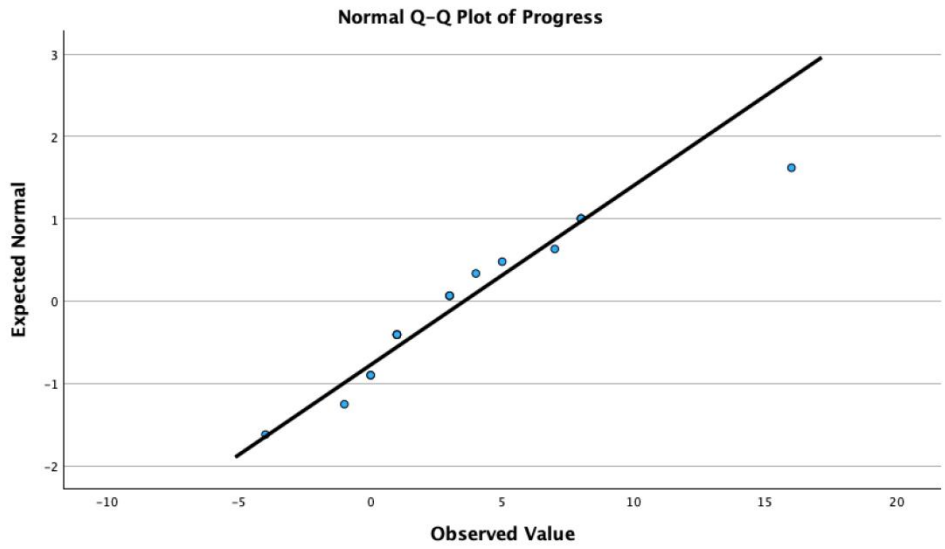
(9) Have you used other mobile apps to learn languages? how do you feel about them?

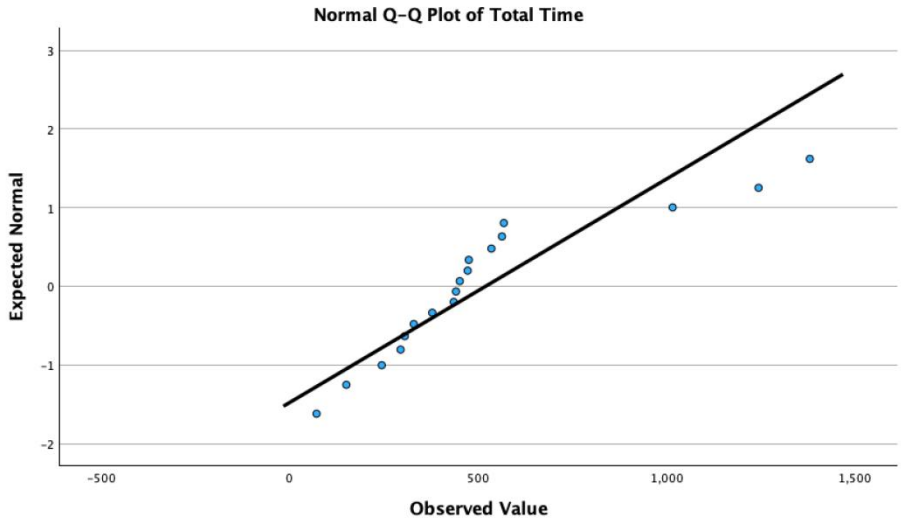
(10) Do you have something else to share about Duolingo or your language learning experiences?

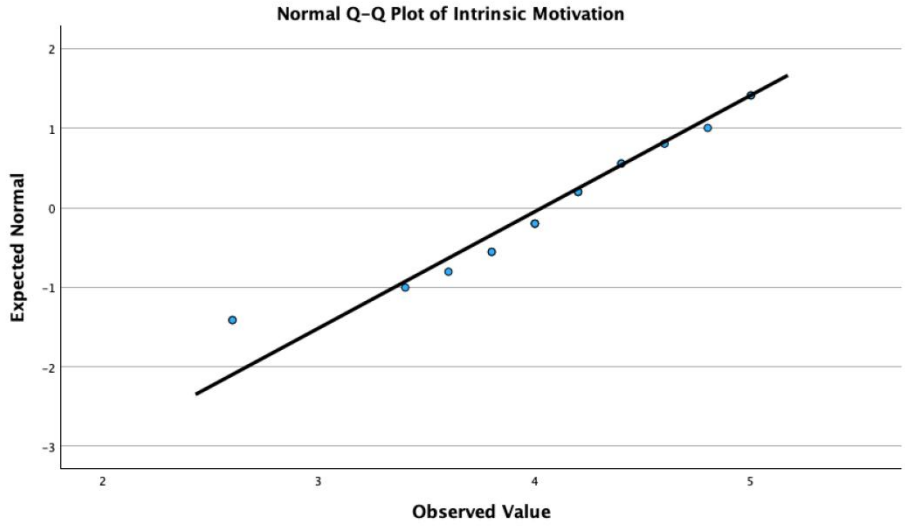
Appendix 3 Q-Q Plots and Histograms Showing the Distribution of Variables

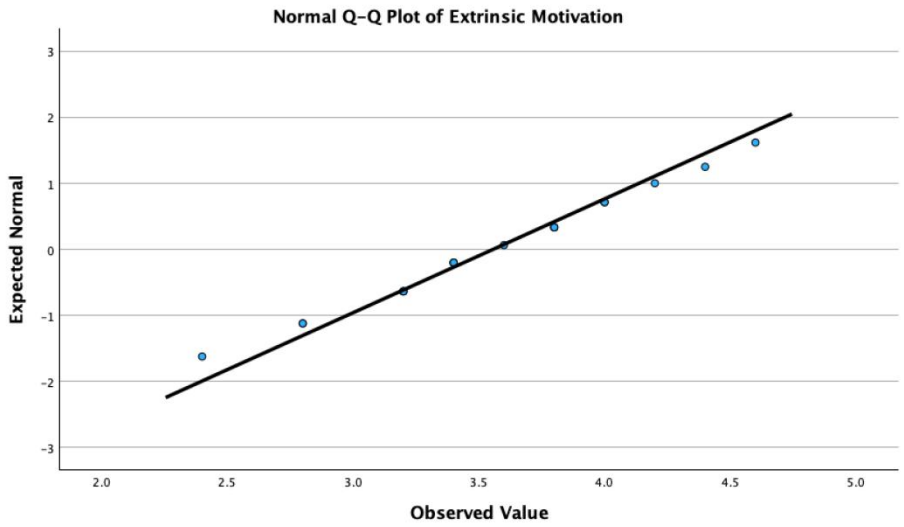


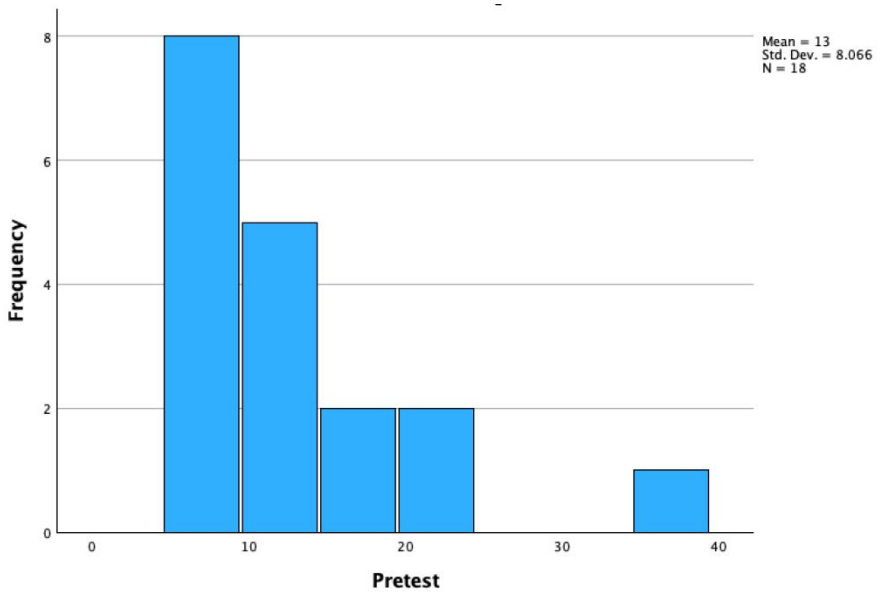


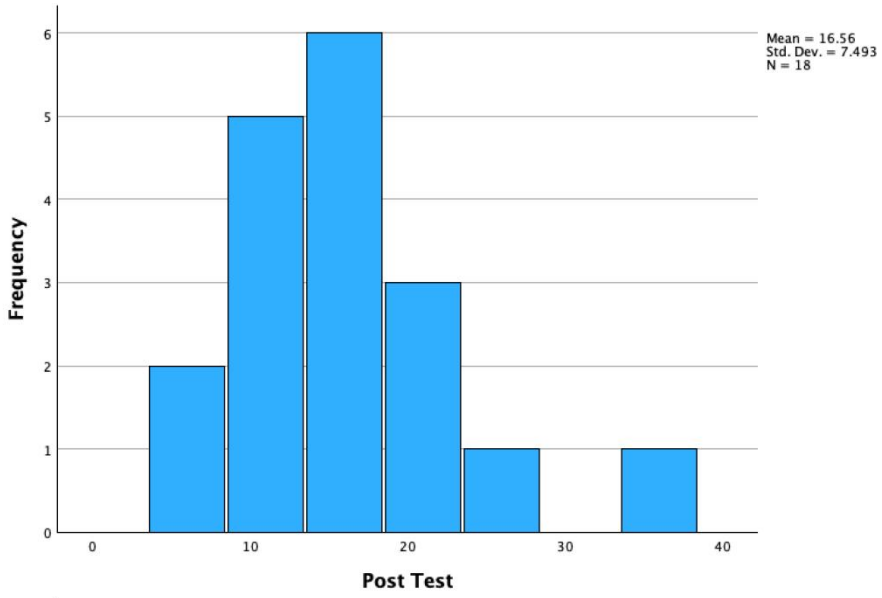


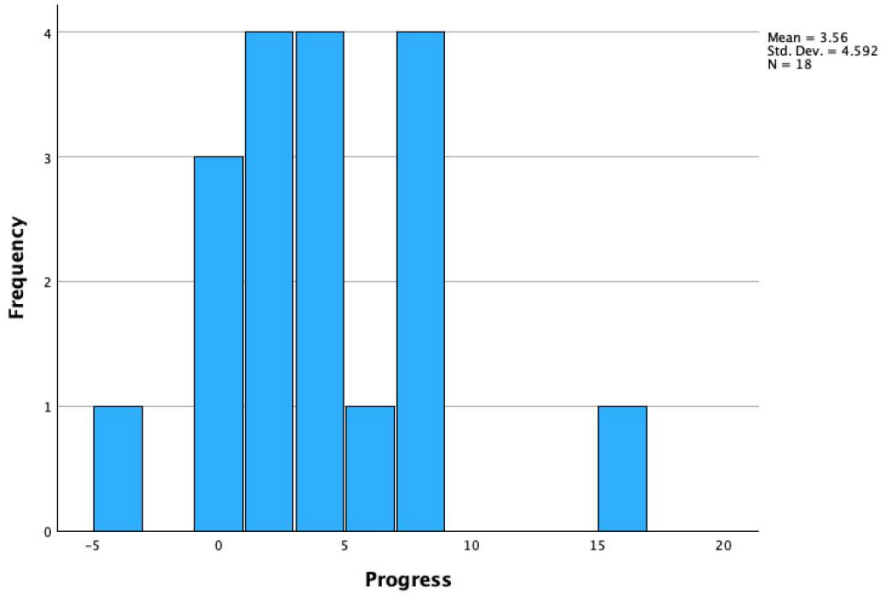

Normal Q-Q Plot of Total Time
Expected Normal
Observed Value
3
2
1
0
-1
-2
-500
0
500
1,000
1,500

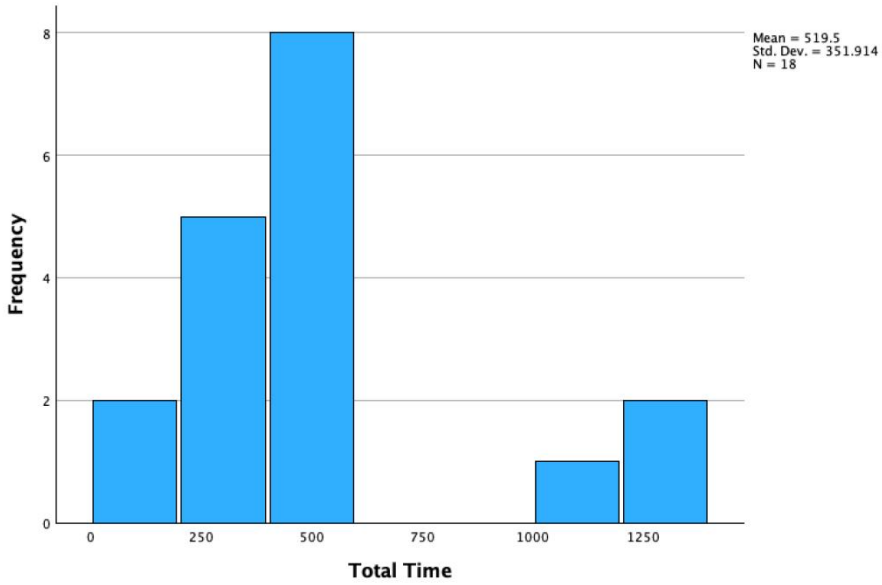

Normal Q-Q Plot of Intrinsic Motivation
Expected Normal
Observed Value
2
1
0
-1
-2
-3
2
3
4
5

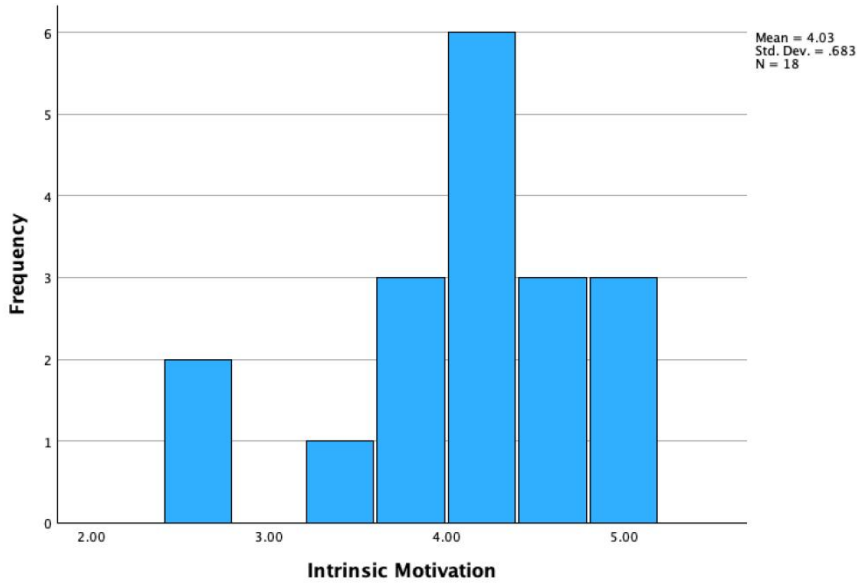

Normal Q-Q Plot of Extrinsic Motivation
Expected Normal
Observed Value
3
2
1
0
-1
-2
-3
2.0
2.5
3.0
3.5
4.0
4.5
5.0

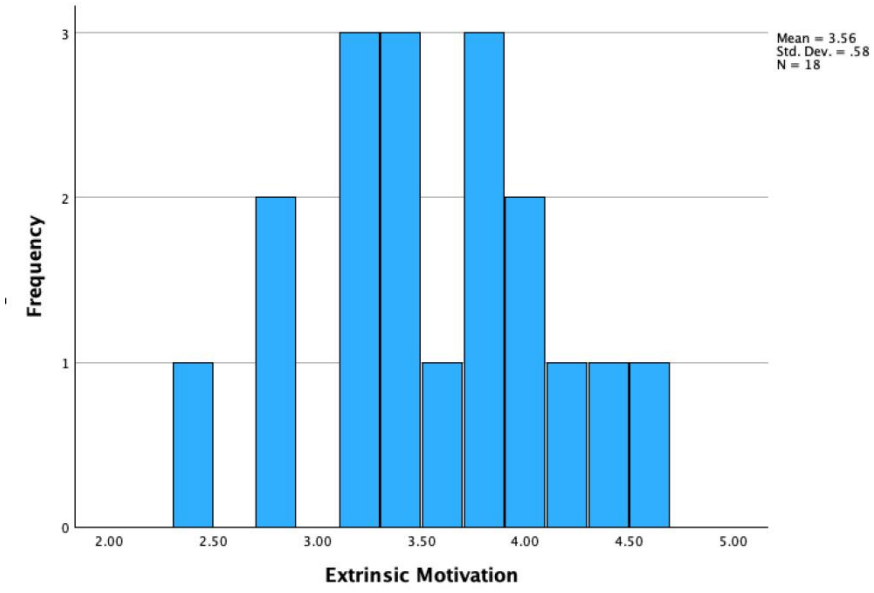

Mean = 13
Std. Dev. = 8.066
N = 18
Frequency
Pretest


Mean = 16.56
Std. Dev. = 7.493
N = 18
Frequency
Post Test


Mean = 3.56
Std. Dev. = 4.592
N = 18
Frequency
Progress


Mean = 519.5
Std. Dev. = 351.914
N = 18
Frequency
Total Time
0
250
500
750
1000
1250


Mean = 4.03
Std. Dev. = .683
N = 18
Frequency
Intrinsic Motivation
2.00
3.00
4.00
5.00


Mean = 3.56
Std. Dev. = .58
N = 18
Frequency
Extrinsic Motivation
2.00
2.50
3.00
3.50
4.00
4.50
5.00

Appendix 4 CUREC Approval With Identifiable Information Removed

SOCIAL SCIENCES & HUMANITIES
INTERDIVISIONAL RESEARCH ETHICS COMMITTEE
DEPARTMENTAL RESEARCH ETHICS COMMITTEE

Department of Education
15 Norham Gardens, Oxford OX2 6PY
student.curec@education.ox.ac.uk; staff.curec@education.ox.ac.uk

[redacted]
Department of Education, Social Sciences Division
University of Oxford

29.02.24

[redacted]

Research ethics approval

Research title: Effects of Mobile Assisted Language Learning on Chinese ESL Students' Proficiency and Motivation: Using Duolingo to Learn French

Research ethics reference: C1A 24 087

The above application has been considered on behalf of the Education Departmental Research Ethics Committee (DREC) in accordance with the University's procedures for ethical approval of all research involving human participants.

I am pleased to confirm that, on the basis of the information provided to the DREC, ethics approval has now been granted for this study.

Please note the following:

Personal data: It is the responsibility of the PI to ensure that all personal data collected during the project is managed in accordance with the University's guidance and legal requirements.

In-person activities: Any data collection involving in-person interactions with participants must have an up-to-date fieldwork risk assessment in place; further guidance is available from the Safety Office's website.

Amendments: Please notify the committee if you intend to make any amendments to the information in your ethics application as submitted at date of this approval, as all changes must receive ethical approval prior to implementation. The amendment form is available on the SSH IDREC webpage.

We welcome feedback on your experience of the ethical review process and suggestions for improvement. Please email any comments to staff.curec@education.ox.ac.uk / student.curec@education.ox.ac.uk or ethics@socsci.ox.ac.uk.

Yours sincerely

Robert Klassen

RKlass

Appendix 5 Participant Information Sheet With Identifiable Information Removed

Department of Education
15 Norham Gardens, Oxford OX2 6PY

[redacted]

Exploring the impact of Duolingo on Chinese ESL learners' French Proficiency and Motivation

PARTICIPANT INFORMATION SHEET

Central University Research Ethics Committee Approval Reference: [C1A 24 087]

1. **Introductory paragraph**

You are being invited to take part in a research project. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether you wish to take part.

2. **Why is this research being conducted?**

This study explores the influence of Duolingo on Chinese ESL learners' French proficiency, motivation, and learning experiences. The research questions will include:

1. To what extent does the use of Duolingo contribute to ESL students' French proficiency?

2. In what ways does Duolingo influence students' motivation in learning French?

3. How do students perceive their experiences of mobile assisted language learning compared with traditional ways of learning languages?

With the advancements of technologies, mobile devices are reshaping the way people learn languages, and the effectiveness of language-learning apps has been supported by many studies. This research aims to contribute valuable insights to mobile assisted language learning and its impact on language acquisition.

3. **Why have I been invited to take part?**

This research will involve about 30 participants aged 18 and above. All the participants should be Chinese-speaking college students, fluent in English as a second language and have little or no French proficiency. Potential participants are expected to take an interest in learning French and have stable access to Duolingo application.

4. **Do I have to take part?**

No. It is up to you to decide whether to take part. You can withdraw yourself from the research, without giving a reason by advising the researcher of this decision. The deadline by which you can withdraw any information you have contributed to the research is June 1, 2024. All your data collected will be deleted from personal devices and the institution storage server, and will be excluded from data analysis if you decide to withdraw

5. **What will happen to me if I take part in the research?**

You are invited to participate in a multi-task online study which will last for a total of six weeks. Your written consent will be collected once you have read this participant information sheet.

During the research, You will be excepted to:

1. Take an online pretest to assess your French proficiency at the beginning of the study, which may take about 20 minutes

2. Engage in a 4-week learning session using Duolingo to learn French. You should spend a minimum of 100 minutes on Duolingo every week, and your learning time will be collected through "Duolingo Weekly Report" sent by email.

3. Complete a post-test and a survey to assess your learning outcome, which may take 30 minutes in total. The post-test will include 40 multiple choice questions and the survey will include 10 questions about your motivation.

4. Have a one-on-one interview with the researcher to talk about your previous language learning experiences and your feedback about learning through Duolingo. The interview will be semi-structured, last 30-45 minutes and will be recorded through Teams. The interview is optional, so you can refuse to be interviewed without giving any reason.

6. What are the possible disadvantages and risks in taking part?

This research is unlikely to cause any risks or discomfort to participants, but if you feel any discomfort, you can end your participation immediately and withdraw your data.

7. Are there any benefits in taking part?

You will improve your French proficiency by engaging in the 4-week learning session with Duolingo, and explore the functions and materials provided by the app. You are likely to find an enjoyable and fun way to enhance your language learning experiences during the experiment.

8. Expenses and payments

Each participant will receive one-month "Super Duolingo" subscription which is worth £9.99.

9. What information will be collected and why is the collection of this information relevant for achieving the research objectives?

The researcher (Hanlu Duan) and the supervisor (Sara Ratner) will have access to the research data.

During the research, below information will be collected:

1. Written consent - to make sure participants agree to take part in the research

2. Results of pretest and post test - to explore learns' gains in French proficiency

3. Survey results - to explore learners' motivation

4. Interview recordings and transcription - to collect in-depth qualitative data

5. participants' contact details - to communicate with participants

Identifiable data (including consent forms) will be stored securely in Nexus 365 OneDrive for Business file, which is approved by the University of Oxford. Video recordings will be deleted immediately after transcription. Other research data will be stored for 3 years after publication or public release of the work of the research. Personal data will not be shared with any other institutions.

10. Will the research be published? Could I be identified from any publications or other research outputs?

The findings from the research will/may be written up in a master's dissertation, academic publications or conference presentations. All the identifiable information will be removed from quantitative data analysis, and interviews will be pseudonymised. Participants can also choose not to take part in the interview.

NB: For doctoral students or other qualifications where a thesis or dissertation needs to be deposited in the Oxford University Research Archive, include the following: A copy of my thesis/ dissertation will be deposited both in print and online in the Oxford University Research Archive where [it will be publicly available to facilitate its use in future research/ its access will be restricted].

1. Data Protection

The University of Oxford is the data controller with respect to your personal data, and as such will determine how your personal data is used in the research. The University will process your personal data for the purpose of the research outlined above. Research is a task that is performed in the public interest. Further information about your rights with respect to your personal data is available from the University's Information Compliance web site at https://compliance.admin.ox.ac.uk/individual-rights.

2. Who has reviewed this research?

This research has received ethics approval from a subcommittee of the University of Oxford Central University Research Ethics Committee. (Ethics reference: C1A 24 087)

3. Who do I contact if I have a concern about the research or I wish to complain?

If you hav[redacted] about any aspect of this research, please contact the researche[redacted]d I will do my best to answer your query. I will acknowledge your concern within 10 working days and give you an indication of how it will be dealt with. If you remain unhappy or wish to make a formal complaint, please contact the Chair of the Research Ethics Committee at the University of Oxford who will seek to resolve the matter as soon as possible:

Social Sciences & Humanities Interdivisional Research Ethics Committee
Email: ethics@socsci.ox.ac.uk

Address: Research Services, University of Oxford, Boundary Brook House, Churchill Drive, Headington, Oxford OX3 7GB

4. Further Information and Contact Details

If you would like to discuss the research with someone beforehand (or if you have questions afterwards), please contact:

Department of Education, University of Oxford
15 Norham Gardens, Oxford OX2 6PY
University ema

Appendix 6 Written Consent Form With Identifiable Information Removed

Department of Education
15 Norham Gardens, Oxford OX2 6PY

UNIVERSITY OF OXFORD

Consent to take part in Effects of Mobile Assisted Language Learning on Chinese ESL Students' Proficiency and Motivation: Using Duolingo to Learn French

Central University Research Ethics Committee (CUREC) approval reference: [C1A 24 087]

Purpose of Study:
This study aims to explore the impact of mobile apps on Chinese ESL learners' French proficiency, motivation and learning experiences, seeking to make a valuable contribution to further research of language acquisition and mobile assisted language learning

	Please initial each box if you agree with the statement
I confirm that I have read and understand the information sheet for the above research. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	
I understand that my participation is voluntary and that I am free to withdraw at any point until **01/Jun/2024**, without giving any reason.	
I understand who will have access to personal data provided, how the data will be stored and what will happen to the data at the end of the project.	
I understand that I will not be identifiable from any publications, reports, presentations, videos or websites.	
I consent to being video recorded.	
I understand how videos recordings will be used in research outputs.	
Use of quotations: Please indicate your preference (select *one* option): a) I do not wish to be quoted. **or** b) I agree to the use of quotations in research outputs if I am not identifiable.	
I give permission for you to contact me again to clarify information.	
I understand how to raise a concern or make a complaint.	

I agree to take part.	☐
I hereby assign to the researcher all copyright in my contribution for use in all work stemming from this project and future projects.	☐
Optional: I agree that my personal contact details can be retained in a secure database so that the researchers can contact me about future studies.	YES / NO

_____	dd / mm / yyyy	_____
Name of participant	Date	Signature
_____	dd / mm / yyyy	_____
Name of person taking consent	Date[1]	Signature

[1] To be signed and dated in the presence of the participant. Once this has been signed by both parties the participant should receive a copy of the signed and dated participant consent form. The original signed and dated consent form should be kept with the project's main documents, which must be kept in a secure location.